Standards, Reputational Costs and Market Share: Empirical Evidence from the Credit Rating Industry

Puneet Prakash
Missouri State University, Springfield

Nanda K. Rangan
Virginia Commonwealth University

We examine the joint effect of average default risk and uncertainty on behavior of rating agencies. We find when average default risk rises in the economy standards weaken but a simultaneous increase in uncertainty causes the standards to tighten more. The net effect is the documented conservatism in ratings. A trade-off between short term gains in market share and long term costs to reputation provides a consistent explanation for these findings

Keywords: credit, ratings, risk

INTRODUCTION

Recent spate in literature (see Goel and Thakor, 2015; Jeon and Lovo 2013) on economic behavior of rating agencies highlights the issues associated with them in the wake of the 2008 financial crisis. The market frictions addressed include cyclicality in raters' behavior, contingent regulation, asset complexity, labor market imperfections, and myopic nature of agencies. The maintained assumption is raters reduce information asymmetry on default risk between investors and borrowers (White, 2010). However, these studies do not account for the increase in uncertainty in default risk (measured by dispersion in probabilities of default)[1] and the subsequent impact on rating agency behavior. Our empirical study adds uncertainty dimension to the line of inquiry above to bridge this gap and provides an explanation for the increase in rater conservatism reported in the literature (see Baghai, Servaes, Tamayo, 2014).

We choose issuer ratings of S&P to demonstrate the role uncertainty plays in the rating process. Data on issuer or firm ratings is readily and publicly available. Additionally, issuer ratings can be considered plain vanilla compared to complex assets, and they reflect probability of default in contrast to expected loss in default for issue ratings. Our study finds that when mean default risk increases in the economy, the uncertainty increases too with the correlation being +0.94. Such high correlation implies that effect of average default risk and uncertainty needs to be analyzed jointly.

Our findings indicate that the average probability of default has a direct relation with loosening of rating standards while uncertainty has an inverse relation. That is, when average default risk rises in the economy, the standards *weaken*, but simultaneously the increase in uncertainty makes rating agencies *tighten* their standards. However, the marginal effects of uncertainty are greater in magnitude. Thus, the joint effect of the average default risk and uncertainty is that rating agencies appear to tighten their standards when default risk increases, and loosen it when it decreases.

If very high average default risk is identified with busts, then our findings indicate the joint effect of risk and uncertainty is that ratings standards tighten during busts leading to increase in downgrades of existing firms and lower initial ratings of new ones. The opposite is true for periods of economic booms. Our findings on the net effect of default risk distribution on ratings conform to the procyclical nature of ratings predicted in many theoretical studies earlier (Bar-Isaac and Shapiro,2012; Opp, Opp, Harris, 2012; Bolton, Freixas, Shapiro, 2013).

However, our result can be puzzling at first glance because we find rating standards decline when *average* default risk increases in the economy. However, since we account for the concurrent *dispersion* in default risk in our study, the net effect we discover is the same as predicted in the academic studies noted above. Nonetheless, a consistent explanation must be provided. An explanation for our findings is given below but a more comprehensive comparison of our paper's underlying arguments to previous studies follows in the literature review section.

Raters provide an ordinal ranking of borrowers. When average default risk increases, the rater who does not lower standards shall have to assign a lower letter grade to a new borrower, in comparison to a competitor who does lower its rating standards. As a result, the rater with inflexible standards shall lose revenue in the issuance market. In order to preserve revenue share the rater lowers its standards. Becker and Milbourn (2011) find that both S&P and Moody's lowered its rating standards as Moody's gained market share over the period 1995-2006, thus providing a basis for our argument that rating standard alterations are tied to market share.

However, as the average default risk increases so does the uncertainty about credit quality. In a cross section of issuers, the rater finds it increasingly difficult to distinguish borrowers' quality with as much precision as prior to the increase in default risk. In the event a highly rated borrower defaults the rater bears substantial reputational costs. So, in order to preserve reputation it tightens the standards and assigns a low rating. Our empirical findings can be thus be explained by the trade-off between revenue share and reputational costs. From a broader perspective, this trade-off is a classic case of balancing short term gains and long term costs.

At this point, we consider it imperative to distinguish the nature of our explanation of the findings from those contained in previous studies. The theoretical papers (see Jeon and Lovo, 2013 for a review) develop their predictions based upon one of the following two assumptions. One, that the true underlying quality of a firm can be determined precisely by the rater at finite costs, however high. Or, the rater can identify a good firm as good (and a bad one as bad) with a fixed probability. But, what if with given budget and time constraints true quality cannot be determined precisely, or the probability with which the good firm can be separated from bad keeps changing over time? Many theory papers (Strausz, 2005) argue that absent a monopoly, raters' accuracy is likely to decrease. Our explanation takes the latter route.

Additionally, in accordance with prior finance literature that finds raters increased their standards over time (see Blume, Lim, and Mackinlay, 1998), we also separate the secular effects of time on rating standards from effects that risk and uncertainty in credit markets have on them,. That is, we model rating standards as a function of time, risk, and uncertainty. We do not find evidence that over the time period of our study (1986 – 2005) there is either a secular tightening or lowering of rating standards. In fact, there is a non-linear effect of time where the relation between standards and time displays quadratic effects with opposite signs on the linear and squared terms. This makes it impossible to conclude that standards either tightened or weakened monotonically over the study period.

Even though there is no secular trend in standards, since the increases in uncertainty lead to tightening of standards and the magnitude of its marginal impact is greater (relative to the marginal impact of average default risk), the net effect leads to tightening of perceived credit rating standards over time as evidenced in previous studies. This is more likely if S&P adjust its ratings in an overly conservative manner due to reputational concerns in response to increases in risk and uncertainty.

Overall, using the credit rating industry as a crucible, our study adds to two strands of certifier industry literature. To the question how certifiers attempt to balance reputational costs with market share concerns over time, our study identifies that it is the result of marginal adjustments to standards when default risk and uncertainty increase. The overall tightening of rating standards reported in prior studies can be ascribed to reputational cost concerns of the rater when uncertainty increases, and explains the unwarranted conservatism of rating agencies documented in the literature (Baghai et al., 2014). In our study, we find that on the margin these reputational cost concerns outweigh the benefit of loosening the standard to gain market share for an established rater in a competitive certification market. At this point we note that we employ vanilla issuer ratings in this study, in contrast to studies that invoke asset complexity to explain observed changes in rating standards (for example, Skreta and Veldkamp, 2009). To the specific financial question on the relation between rating standards and business cycles we find that it is driven by uncertainty (dispersion) about credit quality too than by mean default risk alone.

The implications of our paper are pertinent to at least two sets of players, investors and regulators. For investors like pension funds, with committed belief in ratings in the spirit of Boot and Milbourn (2006), our study points out that the investments foregone solely due to lowered ratings during economic downturns may be a value losing proposition. The reason being that firm quality may not alter fundamentally, but the rater tightens standards due to reputational concerns leading to downgrades. For investors that deal in complex debt instruments, our study harkens that rate of change in ratings on vanilla products signals the uncertainty that raters face in the market. Therefore, ratings of concurrent complex instruments may not be as precise as expected. For the regulator, our study suggests that prescribing a rigid investment grade (IG) boundary for public fund investors in vanilla products may not be ideal. Some flexibility in this boundary especially during downturns may be warranted. Such flexibility becomes even more relevant if a firm's value is destroyed upon getting downgraded from investment to non-investment grade (NIG) due to feedback effect of ratings on value as argued in Manso (2013).

The rest of the paper is organized as follows. In the next section we undertake a brief review of the literature related to our paper and provide the economic framework of our study. In section 3 we discuss the research design, followed by a discussion of empirical methodology, data and sample in Section 4. Section 5 contains the results and robustness checks. Finally, we present our conclusions in Section 6.

THE UNDERLYING FRAMEWORK AND REVIEW OF RELATED LITERATURE

The credit rating markets are dominated by the big three in the United States of America, namely, Standard and Poor's (S&P), Moody's and Fitch, even though there are ten Nationally Recognized Statistical Rating Organizations (NRSROs), so designated by the Securities Exchange Commission. Since the big three follow an issuer pays model of revenue, moral hazard driven behavior may lead to fee-for-good rating. However, the raters themselves argue that long term reputational costs are high enough for them to indulge in myopic profit maximizing behavior. No evidence of such behavior has been found.

The view we take in this paper is that the rater observes a signal of default risks of all firms and truthfully constructs a distribution of the risk of default in a cross section using that signal. Then it decides on cutoffs, or thresholds, to assign letter ratings to borrowers depending upon the observed signal of probability of default. In Opp, Opp, Harris (2012), cross sectional variation in these thresholds are a function of asset complexity and degree of competition. By focusing on issuer ratings over the period 1986-2005 (more on the choice of this period in the empirical methodology, data and sample section), asset complexity ceases to be an issue in our sample. Moreover, the issuer rating market over this period is dominated by primarily two players S&P and Moody's limiting the effect of competition (see White, 2010; Bolton, Freixas, Shapiro, 2013) on thresholds. Indeed several researchers have found there was very little difference between the two raters as far as vanilla instruments like straight corporate bonds were concerned. Like the Opp et al. (2012) study, the variation in rating thresholds in the Bolton et al. (2013) model also takes place around exogenously defined boundaries set by the regulator due to the regulatory advantage of being rated, say, AAA or IG. Rating contingent regulation is a necessary condition in the Bolton et al. (2013) paper to account for variation in the thresholds even if it is time-series variation. Reputation costs are exogenous in Bolton et al. (2103). However authors do recognize that as reputation costs rise, the quality of rating improves (see page 106). Similarly, Bar-Isaac and

Shapiro (2011) analyze the analyst incentives at the rating agencies for cyclicality in ratings, but also treat reputation as exogenous. Moreover, in Bar-Isaac and Shapiro (2011) study, economic fundamentals remain constant over time. In a related paper, Bar-Isaac and Shapiro (2013) endogenize reputation effects but lean on labor-market frictions to draw their conclusions. The paper argues that a monopoly rater invests more in quality during recession as it becomes easier to hire quality analysts due to labor market conditions. For a competitive market, however, the authors offer (page 73, point (2)), when forecasts of growth/economic conditions are better, ratings quality should be higher. This is because reputation building is needed for milking in good times. The authors write that in their model tight labor markets in booms can bring about either procyclical or countercyclical accuracy in ratings (page 71) if credit rating industry is competitive.

Studies by Opp et al. (2012), Bolton et al. (2013), and Bar-Isaac and Shapiro (2011, 2012) explain ratings inflations and resulting inaccuracy during economic booms and busts. The frictions they employ to arrive at their results are different from the straight trade-off between market share and reputation cost we employ to explain findings of our paper. The maintained assumption in these papers is ratings receive a signal about true firm quality that can always be uncovered at a cost (as in Opp et al., 2012; Bar-Isaac and Shapiro, 2011), or with a precision that is unchanging and known (Bar-Isaac and Shapiro 2012; Bolton et al. 2013). In contrast, our explanation relies on the fact that it may be very expensive to uncover the true value of a varying signal of firm quality with fixed precision, so that it may be rendered economically infeasible to do so. The precision thus varies over time. We also assume that ratings are truthfully revealed hence accurate to that extent.

To that extent our paper is more in conformance with Manso (2013). The author assumes raters always prefer accurate ratings to inaccurate ones and proves that small shocks to fundamentals can lead to multi-notch downgrades. The results of this paper suggest that if the fee is small relative to the reputational concerns of rating agencies it only introduces small distortions while inducing rating agencies to make gradual changes in credit ratings. Our paper differs from Manso (2013) in that the rating thresholds are exogenous and invariant in that study's framework but in our paper they are determined within the empirical model and vary.

Doherty, Kartasheva, Phillips (2012) analyze the effect of a new entrant into a monopoly rater market. The study models the variance in the quality of firms, but assumes it can be uncovered perfectly by the rater so that the rater captures all information. The authors then proceed to model the disclosure policy of the rater in the spirit of Lizzeri (1999). Instead, we argue that the rater uncovers true quality not perfectly but with a degree of imprecision, and this imprecision is time variant. That is, the rater fails to capture all information. Although the analysis in Doherty et al. (2012) refers to the reputational cost of the new entrant that determines its behavior, the study does not, however, explicitly account for the association between risk and uncertainty that we do in this paper (see footnote 23, Doherty et al., 2012).

Our papers findings are also related to papers that provide behavioral explanations for ratings inflation during business cycles based on raters' myopia. Mathis, McAndrews and Rochet (2009) build a dynamic model in which the credit rating agency (CRA) builds up reputation by engaging in grading inflation only to cash it afterwards. However, this leads to a crisis of confidence in ratings, leading to reputation cycles. The paper theorizes that if rater income comes from sources other than rating complex products then reputation is a good disciplining device for CRAs, a proposition consistent with our findings in this paper. Similarly, Holden, Natwik, Vigier (2014) show how booming market conditions may lead to ratings inflation. Contrary to ratings inflation and deflation, Fisher (2015) argues ratings are smoothed over time to increase financial stability and that ratings are always noisy in a competitive market. In a related study, Goel and Thakor (2015) point out ratings are coarse to preclude ratings inflation and that excessive competition raises coarseness of ratings leading to fewer rating categories and loss of information. For a review of why screening falls in credit markets during economic booms, the reader is referred to studies by Berger and Udell (2004), Ruckes (2004).

Finally, in relation to sovereign ratings Ferri, Liu, and Stiglitz (1999) find procylicality in ratings driven by economic fundamentals. The authors' propose an endogenous rationale to explain why rating agencies became excessively conservative after having made blatant mistakes in predicting the East Asian crisis. Specifically, rating agencies would have an incentive to become more conservative so as to recover from the damage these mistakes caused to them and to rebuild their own reputation capital.

EMPIRICAL RESEARCH DESIGN

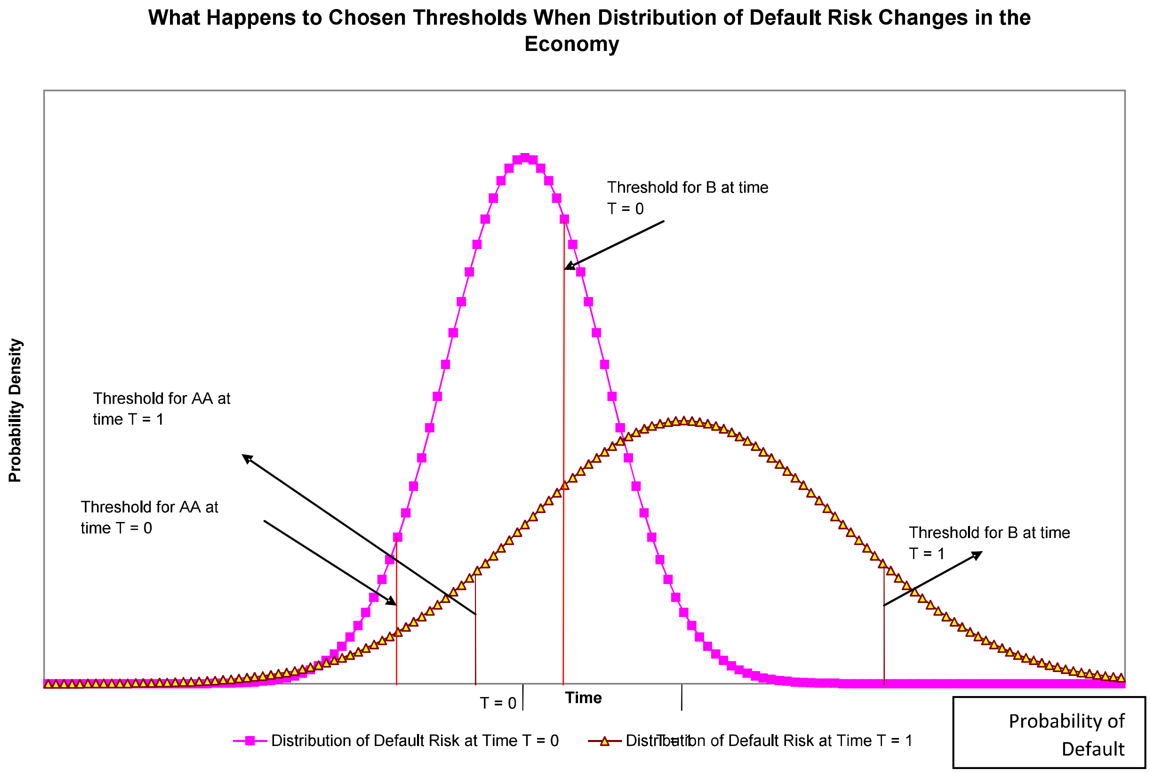
The seminal paper measuring rating standards is by Blume, Lim and MacKinlay (1998) that concludes that rating standards became stricter during the period 1976 to 1995. Blume et al. (1998) measure rating standards using the intercept term in an ordered probit model. Subsequent studies, for example, Amato and Furfine (2004), and Baghai et al. (2014) find a similar trend in the intercept term as Blume et al. (1998) leading to similar conclusion. Baghai et al. (2014) also associate the metric to rating agency's unexplained conservative behavior over the period 1995-2006. We too find a similar trend in the intercept term as reported in Blume et al. (1998) and Baghai et al. (2014), but employ a different approach based on rating thresholds to analyze rating standards, to avoid the shortcomings associated with the intercept term alone. We explain this approach below.

In any ratings process, a rating agency observes a signal of the underlying credit quality of all debt issuers and assigns a letter grade to each. The action of the rating agency thus reduces to estimating the probability of default (PDs) for the universe of firms and choosing cutoff (or threshold) values of estimated PDs to assign letter grades. For example, firms with a PD lower than 10 basis points (bps) may be assigned AAA, while those between 10 bps and 100 bps are assigned AA, and so on. More stringent standards imply higher letter ratings have very low cutoffs. Thus, under a stricter rating standard the cutoff for AAA may be 8 bps and 75 bps for AA. These rating thresholds thus provide an alternative way to capture rating standards.

The rating standards based on thresholds are obtained from the ordered probit model for different rating categories. We allow for time-varying slopes and analyze all thresholds obtained from period-by-period regressions. The cutoff values reflect the choice, and hence the standards, of the rating agency.

Figure 1 describes the empirical framework of the study. Credit quality distributions are constructed by the rating agency that decides on different thresholds to assign ratings, which in turn reflects the agency's standards. Two distributions are shown in Figure 1, one for time T = 0 and another for T = 1. It is assumed that the underlying distribution changes at T = 1, hence the rating agency decides on new thresholds different from those at T = 0. We test the relationship of thresholds to the average (or location) and dispersion (uncertainty) of default risk distribution.

FIGURE 1
EMPIRICAL FRAMEWORK FOR THE PAPER



This figure depicts the empirical framework designed to answer the question, what happens to rating standards over time? At time T = 0, the credit quality of firms has a distribution as shown below. The rating agency chooses the thresholds to categorize firms into different rating categories. At time T = 1, the distribution changes. How do the standards (that is, the thresholds) change?

Our analysis is premised on the positive correlation of the mean and dispersion in default risk (labeled uncertainty). When average quality declines (i.e., the location of the distribution of PDs moves towards 1), the dispersion in quality increases. In terms of observable data, the phenomenon manifests in credit spreads. During economic downturns, spreads on high quality bonds and low quality bonds increase asymmetrically, with the increase being larger in magnitude for low quality firms. Thus the dispersion in credit spreads (difference in borrowing costs of a firm rated AAA versus one rated CCC) [2] increases during downturns.

Economic rationale for this positive correlation between dispersion in default risk and average default risk is provided in several theoretical studies. For example, Eichengreen et al. (2001) has documented that following the Asian crisis of 1997, non-sovereign borrowers (in the Asian region) were crowded out in the financial markets by their more creditworthy counterparts, the sovereign borrowers, due to “flight to quality.” Thakor (2015a) observes the spreads change based on the type and value of collateral offered, and non-collateralized borrowing (typically rated lower than collateralized one, ceteris paribus) costs increasing significantly more during a crisis. Thakor (2015b) argues that when unexpectedly high defaults occur, investors’ beliefs in rater’s ability changes and more weight is placed on luck driven outcomes. In the limit one can foresee that for low quality firms all weight is placed on luck and none on rater ability thus increasing borrowing costs for low quality even more than for high quality firms. Regardless of the causative latent factor underlying the decline in quality, it is expected to have a more severe adverse impact on lower-quality firms than higher-grade firms. For these reasons, the dispersion in the credit quality distribution is likely to increase concurrently with a decline in average quality.

Contrary to the effect of location of credit quality distribution on standards, the rating agency makes its standards more stringent when dispersion increases. Consider two firms rated AAA and BB. If dispersion of underlying quality distribution increases, AAA declines to a marginally lower level but BB declines sharply in quality. Because the rating agency lowers its standards in response to a decrease in average quality, the agency may retain the AAA rating for the first firm. However, if it also reaffirms the BB rating for the second firm, then it is more likely that firms in the BB rating class default more than before, causing the actual default rate amongst BB rated firms to increase substantially. This leads to an erosion of investor confidence in the BB classification of a rating agency and, by extension, to other rating categories. The aggregate effect is the loss of reputational capital for the rating agency, which is of value to the rating agencies. To avoid this outcome, increases in the dispersion of credit quality are accompanied by a tightening of standards, and ensuing downgrades. This is empirically borne out (see Carvalho, Laux, Pereira, 2014) as rating changes are more intense during economic downturns.

So, in this paper we analyze the joint effect of the location and dispersion parameters of underlying credit quality distribution on rating standards.

EMPIRICAL METHODOLOGY, DATA AND SAMPLE

The first step involves modeling the distribution of the signal of credit quality that can be observed by the rater for each period. We employ the Merton (1974) model of default and the optimal hedge equation from Ronn and Verma (1986) to obtain the parameters of the distribution of credit quality.[3] Subsequently, we model the decision-making of the credit rating agency using the ordered probit which is the standard model utilized for this purpose (see Greene, 2003). We use the output from both these models to verify if the functional relationship between the thresholds, which capture rating standards obtained from the ordered probit model, and the credit quality distribution parameters (obtained from the Merton model), hold as predicted. The main functional relation we intend to test is:

$$Rating\ standard = f(\ Mean\ probability\ of\ default, dispersion\ in\ probabilties) \quad (1)$$

and the expected signs are opposite for mean and dispersion.

We use a structural model of default (Merton's, 1974) over statistical models as the focus of this study is on long-term issuer ratings. Two studies also corroborate and suggest that Merton's structural model is a better fit for predicting long-term default compared to other models. Leland (2004) reports that the structural models predict longer-term default frequencies better. Hillegeist et al. (2004) compared the accounting-based Z-score and O-score models with the Merton model and conclude that Merton model "provides significantly more information than either of the two accounting-based measures"[4] and their findings are robust to modifications in the accounting-measure based models. More recently, Bernoth and Pick (2010) use the model to forecast the fragility of banking and insurance sectors. Finally, Bharath and Shumway (2008) find that despite limitations of Merton's model, the probability of default obtained from it adds predictive power to other statistical models of default.

Merton (1974) Model of Default

In Merton's (1974) model, default occurs when the value of the company's assets falls below the level of outstanding debt. This model assumes that the asset value $V(t)$ of the firm at time t, follows a lognormal stochastic process of the form

$$dV(t) = \mu V(t)dt + \sigma V(t)dz(t), \quad (2)$$

where μ and σ are constants, and $z(t)$ is a standard Wiener process. The firm's liabilities have a face value of M maturing at date T. Equity is modeled as a call option on the firm's underlying assets. The boundary condition at time T is given by $S(T) = Max[V(T) - M, 0]$.

The call option on firm assets at time t can be calculated using the Black-Scholes formula:

$$S(V(t),t) = V(t)N(d_1) - Me^{-r(T-t)}N(d_2) \tag{3}$$

where T is time of maturity, t is the current time, r the risk-free rate of interest, M is the strike price for the call option on assets, the argument d_2 equals $\frac{\left[\ln\left(\frac{V(T)}{M}\right)+(r-0.5\sigma^2)\right]}{\sigma\sqrt{T-t}}$, and d_1 equals $d_2 + \sigma\sqrt{T-t}$. $N(.)$ denotes the standard cumulative normal distribution and $S(V(t),t)$ the value of equity holders call option. With this formulation, Merton's risk-neutral probability of default (RNPD) is the probability that the equity holder's call option is out of the money at maturity under a risk-neutral distribution, and is given by RNPD = $1 - N(d_2)$.

Equation (3) above cannot be estimated directly because it contains two unknowns, namely, $V(t)$ and σ. Ronn and Verma (1986) use the optimal hedge equation to make this model operational.

$$\sigma_s = N(d_1) \times \sigma \times [\frac{V(t)}{S(t)} \tag{4}$$

Equation (4) links the volatility of equity σ_s to the volatility of assets σ. $V(t)$ and σ can thus be estimated simultaneously using equations (3) and (4) (Two equations in two unknowns). We adopt this approach to estimate the probability of default.[5]

The risk-neutral default probability, i.e., $1 - N(d_2)$, needs to be adjusted for risk to obtain the risk-adjusted probability of default, referred to as PD henceforth. The formula for PD is given by $1 - N(d_2')$, where

$$d_2{}' = \frac{\left[\ln\left(\frac{V(T)}{M}\right)+(\mu-0.5\sigma^2)\right]}{\sigma\sqrt{T-t}} \tag{4'}$$

and $\mu = r + \beta_A[E(R_m) - r]$. $E(R_m)$ is the expected return on the market, and β_A is the asset beta of the firm. We substitute the actual return on the market obtained in the prior quarter as the expected return on the market. Both R_m and r are obtained from Professor Kenneth French's website.[6] The construction of asset betas and the implementation of the Merton model are outlined in the Appendix A.

A Rating Agency Model

We utilize the same model of ratings as has been used in prior studies on rating standards (Blume *et al.*, 1998; Amato and Furfine, 2004), namely, the ordered probit model. The dependent variable is the rating category y_{it}, where i refers to the firm and t refers to current period. The discrete-ordered variable y_{it} is assumed to be linked to an underlying continuous variable y_{it}^*, which is unobserved and has the following relationship with the observed letter rating:

$y_{it} = y_0 \; if \; y_{it}^* \in (-\infty, \mu_0),$

…

$y_{it} = y_n \; if \; y_{it}^* \in [\mu_{n-1}, \infty),$

where $y_0, \ldots, y_n$ are consecutive integer values corresponding to the observed letter rating categories. Note that the rating categories are based on the thresholds $\mu_0, \ldots, \mu_{n-1}$. The latent continuous variable is modeled as follows.

$y_{it}^* = \alpha_t + \beta_t X_{it} + \varepsilon_{it},$ where

$$E(\varepsilon_{it}) = 0; \text{and } E(\varepsilon_{it}^2) = [\exp(\gamma W_{it})]^2 \tag{5}$$

where X_{it} is the vector of independent variables which determine ratings, and W_{it} is a variable used to correct for heteroskedascticity. The model assumes heteroskedasticity in errors similar to Blume et al. (1998). We use the square root of total assets for W_{it}.

Blume et al. (1998), Baghai et al. (2014) and Amato and Furfine (2002) employ time dummies α_t to capture dynamic rating standards while keeping the slope coefficient fixed. We allow the slope coefficients to vary over time.[7] Given the intercept term summarizes all omitted variables it is susceptible to the omitted variable bias critique. To alleviate this bias, we only test for the relationship between rating standards and the economy-wide credit quality distribution by analyzing thresholds obtained from a period-by-period ordered probit regression (equation (5) above). This execution provides a panel of thresholds $(\mu_{0t}, \dots, \mu_{n-1t})$, which can be analyzed with respect to their relationship to the underlying credit quality distribution. A secondary advantage of running period regressions is that the weights given to different components of the independent variable vector X_{it} can be allowed to vary, thus controlling for time variation in the information content of these variables.

Data and Sample

S & P began providing issuer (firm level) ratings publicly only from the last quarter of 1985. Hence, our period of study begins in 1986. In the wake of the Sarbanes-Oxley Act, and after the Committee of Sponsoring Organizations of the Treadway Commission (COSO) released its Enterprise Risk Management (ERM) Framework in 2004, a systematic shift in ratings criterion occurred. S&P began to incorporate ERM in its ratings and its first notification affecting all financial institutions was released in November, 2005. The same year, early warnings of impending financial crisis were sounded by Professor Robert Shiller to the Office of the Comptroller of Currency, and Professor Raghuram Rajan at the Jackson Hole Economic Symposium. The median home prices decreased by 3.3% between the fall of 2005 and first quarter of 2006, bringing the booming housing sector growth to an abrupt halt.[8] For reasons of systematic change in ratings criterion and the beginnings of financial crises in 2006, our data period ends in 2005 resulting in eighty quarters of data. A period of 20 years is similar to the one employed by Blume et al. (1998). The data necessary to calculate PD require merging both the CRSP daily stock files and the COMPUSTAT North American Industrial Quarterly files. The data for equity risk premiums to convert RNPDs to PDs comes from the daily data files of Professor Kenneth French's website. We use CRSP files to calculate the volatility of the asset process and employ COMPUSTAT Quarterly Files to calculate the strike price and the financial ratios employed in the econometric model of ratings. The monthly risk-free rate of interest is manually collected from the Wall Street Journal; we use Treasury STRIPS to proxy for the spot rates for different maturities. The ratings considered are long-term domestic issuer credit ratings from S & P.

Observations for which valid long-term domestic issuer credit ratings are not available are deleted. Observations for which firm-specific independent variables are missing are also deleted. Given the findings in Loviscek and Crowley (1990) we exclude CUSIP numbers corresponding to government and public administrative bodies as factors like demographic characteristics affect ratings in this category (for example, of municipal bond ratings). Finally, we winsorize the sample based on variables employed in the agency rating model to remove outlier firms. The final sample is thus an unbalanced panel of quarterly data with 47,004 observations on 2,499 firms.

EMPIRICAL TESTS AND RESULTS

We describe the sample in Table 1, and then present the results of our empirical analysis in two parts. The first part pertains to the constructed market variable obtained from Merton model, namely, the probability of default (PD). These results for PD are presented in Tables 2 and 3. The second part discusses the rating agency model variables and the results obtained from ordered probit (Tables 4 and 5). Finally, we combine the two parts together, and present the result of rating standards analysis in Table 6.

Sample Description

Table 1 presents the details of our sample by rating category and period. We categorize the ratings into seven classes (namely, AAA, AA, A, BBB, BB, B, and CCC and below) and for reporting purpose, divide the sample years into four periods of five years each. The letter rating classes are numerically coded for empirical analysis (that is, AAA = 1, AA = 2, A = 3, BBB = 4, BB = 5, B = 6, and CCC & below = 7). The mean and median rating over the period is BBB (i.e., 4). There is a downward trend in the percentage of highly rated firms (i.e., AAA and AA) over the sample period. While the bulk of the firms (89.67%) are rated in the middle letter rating categories (i.e., rated between A to B), 76% of the firms are rated between A and BB. Firms rated B or below form 14.94% of the sample, and the number of firms in the sample with a rating below B is quite smaller. Firms with a rating of CCC or below make up 0.86% of the sample, and AAA firms constitute 1.54% of the entire sample. Between the years 1986 and 2005, 62.17% of the observations belong to rating categories between AAA and BBB, which is referred to as the "investment grade" category.

TABLE 1
NUMBER OF OBSERVATIONS PER LETTER RATING CATEGORY BY PERIOD

Rating	1986-1990	1991-1995	1996-2000	2001-2005	Total
AAA	275	218	170	60	723
AA	1280	1201	939	305	3725
A	3150	3038	3624	1959	11771
BBB	2202	2848	4508	3445	13003
BB	1747	2195	3578	3239	10759
B	1753	1273	1980	1612	6618
CCC & below	172	71	61	101	405
Total	10579	10844	14860	10721	47004

This table shows the distribution of observations by rating category and period. The sample period 1986-2005 is equally divided into four periods of five years each. The firms rated with high letter ratings (that is, AAA and AA) show a downward trend, while firms rated BB show an upward trend. The total number of observations for the sample period is 47,004. The total number of quarters is 80 for Yr1986q1 to Yr2005q4. The letter rating classes are numerically coded as AAA = 1, AA = 2, A = 3, BBB = 4, BB = 5, B = 6, and CCC & below = 7.

Results on Merton's PD

Our first set of analyses pertains to the constructed variable PD because PDs should reflect the underlying credit quality before they can be utilized to analyze rating standards. The summary statistics of the distribution of PDs for the period 1986-2005 is presented in Table 2. The jump in central value in 1987 corresponds to the stock market crash and is indicative of the market revising its expectations of the default risk in the economy. There is a rising trend in both the mean and median from 1995 onward; quarterly analysis reveals that the increasing trend begins in the last quarter of 1995, with the mean PD attaining a maximum in 2000. It decreases slightly during 2001 but rises again in 2002, only to start declining steadily from 2003. A non-parametric visual comparison of PDs with the credit spreads shows that the computed PDs reflect the trend in spreads.[9]

The mean PD is higher than the median values, reflecting the skewed nature of the distribution. Consequently, even though we report results with mean as measure of central tendency and standard deviation as measure of dispersion, we carry out robustness check using other measures as well. In particular, we employ the median and 60th percentile for location, and inter-quartile range and mean absolute deviation for the dispersion in the credit quality distribution. These measures of central tendency and dispersion are independent of the shape of the distribution. The results are qualitatively similar to those obtained with mean and standard deviation.

TABLE 2
SUMMARY STATISTICS OF THE CONSTRUCTED VARIABLE MERTON PD BY YEAR

Fiscal Year	Mean	Median	Std Dev	Minimum	Maximum
1986	0.0660	0.0048	0.1429	0	0.9802
1987	0.1325	0.0156	0.2178	0	0.9737
1988	0.0734	0.0040	0.1503	0	0.9915
1989	0.0573	0.0008	0.1446	0	1.0000
1990	0.0795	0.0043	0.1645	0	0.9751
1991	0.0709	0.0033	0.1523	0	0.9708
1992	0.0715	0.0033	0.1513	0	1.0000
1993	0.0690	0.0036	0.1472	0	1.0000
1994	0.0607	0.0030	0.1383	0	1.0000
1995	0.0659	0.0013	0.1495	0	0.9800
1996	0.0729	0.0032	0.1566	0	0.9910
1997	0.0859	0.0061	0.1681	0	0.9849
1998	0.1427	0.0339	0.2129	0	0.9881
1999	0.1595	0.0609	0.2097	0	1.0000
2000	0.2001	0.1091	0.2270	0	1.0000
2001	0.1696	0.0755	0.2093	0	0.9726
2002	0.1949	0.1240	0.2115	0	0.9666
2003	0.1095	0.0261	0.1704	0	0.8795
2004	0.0802	0.0105	0.1464	0	0.9333
2005	0.0626	0.0063	0.1265	0	0.8994
N Obs	47004				

This table shows the summary statistics of the constructed variable risk-adjusted probability of default (that is, Merton PD). The time to maturity of liabilities used to solve the Merton model is constructed as the Macaulay duration, with current liabilities assumed to have a maturity of 6 months and long-term liabilities assumed to have a maturity of 10 years. Merton PD is computed by solving the Merton (that is, equity as a call option) equation and the optimal hedge equation simultaneously, as described in Ronn and Verma (1986). An outline of the computations is provided in the Appendix A of the paper. The study relies on the capacity of the Merton PD to rank firms in terms of default risk, like ratings, rather than produce actual (i.e., cardinal values of) probabilities of default, and the usage is consistent with Bharath and Shumway (2008). This is indeed the case, as shown in Table 3. The mean and median values point to a skewed distribution.

Our focus is on the ability of the PDs from the Merton model to rank-order firms with respect to their credit quality. The usage of PDs in this fashion is consistent with Bharath and Shumway (2008) and Brockman and Turtle (2003). Table 3 displays the ability of the model to rank-order firms according to their credit quality. The average PD over the entire time period is approximately 10.04%, while the median is 1.15%. The standard deviation is 17.83%, with probabilities ranging from 0 to 1.

An analysis of the mean and median PD across rating categories is provided in Panel A of Table 3. The t-tests (second to last column) for the differences in mean PDs between one rating category and the next reveal that the mean PD in AAA is not significantly different from that for AA, but they are significantly different across all other categories. The last column provides the Mann-Whitney-Wilcoxon z-statistic for the differences between median PDs across rating categories and shows that medians follow the expected ranking of default risk. That is, lower letter ratings have higher mean and median PDs. Moreover, volatility in PD is higher as one traverses down the letter rating scale, thus implying greater dispersion in quality for lower-level categories. The mean and median for investment grade firms are 4.45% and 0.15%, respectively, while they are 19.23% and 10.75%, respectively, for non-investment grade firms.

Even though this study relies upon the ability of the PDs to rank-order firms according to their default risk, the computed PDs are comparable to long-term cumulative default rates (CDR) noted in the literature/practice. Considering bonds issued between 1971 and 1987, Altman (1989) reported a CDR of 31.17% for bonds rated CCC and below, 11.53% for BB-rated bonds and 2.33% for those rated A. Correspondingly, for bonds issued between 1970 and 2006, Moody's 5-year CDRs for equivalent categories are 52.66% for their Caa-C rating, 26.79% for their B rating, and 10.21% for their Ba rating.[10]

Second, since our focus is on long term domestic issuer credit ratings, we analyze the maturity structure of liabilities utilized in the Merton model and verify that it follows the pattern predicted by theory. Diamond (1991) argues that firms with very high or very low ratings are expected to have shorter maturity structures.

The overall average time to maturity for liabilities for the Merton model is 5.22 years. Specifically, the average value for investment grade firm is 4.98 years, and for non-investment grade firm is 5.61 years. Table 3 Panel B provides differences-in-means analysis for time to maturity for different rating categories. We find that the maturity structure of liabilities follows a non-monotonic shape with respect to default risk as predicted. Mean time to maturity for CCC and below (5.02 years) is significantly lower than the mean time to maturity for B-rated firms (5.75 years); meanwhile, mean time to maturity for AAA (3.81 years) is significantly lower than mean time to maturity for AA rated firms (4.38 years). The constructed time to maturity (i.e., Macaulay duration as maturity of liabilities) is thus in accordance with theoretical predictions on the riskiness of firm debt and maturity, and is consistent with Stohs and Mauer (1996), who report that bonds with very high and very low ratings tend to have shorter maturity. It is also consistent with findings in Xie et al. (2009) who report that default risk for risky bonds decreases bond durations.

TABLE 3
TEST OF DIFFERENCES-IN-MEANS ACROSS RATING CATEGORIES FOR PDS AND TIME TO MATURITY

Panel A: PD by rating category

Rating	N Obs	Mean	Median	Std Dev	Minimum	Maximum	t- stat for diff in means	Z- test for diff in medians
AAA	723	0.0234 (5.17***)	0.0000	0.1219	0.0000	0.9513		
AA	3725	0.0231 (13.91**)	0.0000	0.1012	0.0000	0.9751	0.08	-10.64***
A	11771	0.0314 (32.54***)	0.0005	0.1047	0.0000	1.0000	-4.27***	-27.95***
BBB	13003	0.0638 (54.35***)	0.0080	0.1338	0.0000	1.0000	-21.04***	-46.07***
BB	10759	0.1487 (81.54***)	0.0677	0.1892	0.0000	0.9915	-40.41***	-51.04***
B	6618	0.2520 (97.84***)	0.1820	0.2334	0.0000	0.9881	-31.93***	-34.04***
CCC & below	405	0.3759 (27.29***)	0.3309	0.2772	0.0000	1.0000	-10.25***	-8.78***

Panel B: Macaulay duration for time to maturity of liabilities by rating category								
AAA	723	3.8110 (76.87***)	3.5053	1.3331	1.3855	8.3768		
AA	3725	4.3861 (145.17***)	4.0523	1.8402	0.5606	8.8073	-8.09***	-7.54***
A	11771	4.8603 (299.08***)	4.8091	1.7631	0.5000	9.3591	-14.16***	-14.49***
BBB	13003	5.3340 (353.53***)	5.3301	1.7205	0.5000	9.5532	-21.39***	-21.03***
BB	10759	5.5495 (322.22***)	5.7074	1.7864	0.5000	9.5900	-9.44***	-10.35***
B	6618	5.7543 (255.58***)	5.8811	1.8316	0.5674	9.6496	-7.27***	-6.91***
CCC & below	405	5.0243 (46.55***)	5.3152	2.1722	0.5000	9.2376	7.70***	6.16***
Total	47004							

***,** and*: Significant at 1%, 5% and 10% levels of significance, respectively.

This table provides tests of differences in the mean and median for risk-adjusted probability of default (PD) and time to maturity variables by rating categories, based on the t-stat and Mann-Whitney-Wilcoxon Z-stat, respectively. In all test comparisons, one letter rating category is compared to the one preceding it. Numbers in brackets below mean values test for significant differences from zero. t and Z-test results are based on two sample test statistics. All variable values are rounded to four decimal places. Panel A reports results for PD, while Panel B reports results for the Macaulay duration (that is, the time to maturity) of liabilities. Computation details on the Macaulay duration and PD are provided in the Appendix A.

Taken together, the results in Tables 2 and 3 suggest that the constructed variable Merton PD provides a good reflection of the default risk of the universe of firms on average, individual cardinal values notwithstanding. Time to maturity follows the distribution predicted by theory, while PDs themselves track the credit spreads observed in the market during this period. Consequently, we can utilize the parameter values from the PD distribution for analyzing rating standards.

Results on Rating Agency Model

A typical rating agency model contains accounting variables that measure leverage, profitability, and interest coverage while controlling for size. Although Blume et al. (1998) introduce market betas and idiosyncratic volatility as proxies for off-balance sheet risks of the firm, Amato and Furfine (2004) contend that such measures do not adequately capture the risks of the firm. In a later study, Bharath and Shumway (2008) find that PD adds information that enhances predictability of bankruptcy prediction models. We thus append the accounting variables with PD, a variable more pertinent to credit ratings, and control for industry effects. The rating agency model we employ is ordered probit and the model is specified as:

$$rating_{it} = f(Lev, ROA, OpInctoSales, OpInctoDebt, ICR, PD, controls\ for\ size\ and\ industry)$$

where rating of firm i at time t varies from 1 (AAA) to 7(CCC or below). We correct for heteroskedastic errors using square root of total assets as given in equation (5) earlier.

The variables used for the determination of ratings can be categorized as interest coverage ratios, profitability ratios, and leverage ratios. We use averages of prior twelve quarters for independent variables. Lev (long-term debt to total assets)[11] denotes leverage which plays an important role in ratings as higher leverage ratios imply a higher risk of bankruptcy for a firm. Therefore, a high-leverage firm should receive low letter ratings. We include return on assets[12] (ROA) as a measure of profitability. The higher the ROA is, the higher is the letter rating that a company can expect to obtain. A firm's ability to cover its operational expenses is captured through the operating income to sales ($OpInctoSales$)[13] and operating income to debt ($OpInctoDebt$) ratios.[14] These ratios are expected to be positively related to the letter ratings; i.e., the higher these ratios are, the higher is the rating. A firm's ability to meet its debt obligations is reflected in the interest coverage ratio (ICR).[15] Consistent with Blume et al. (1998), we model the relationship between ratings and interest coverage as piece-wise-linear (see Blume et al., 1998 for details). The higher this ratio is, the higher are the letter ratings assigned. The $size$ effect on ratings is modeled using the natural log of market capitalization, and to control for $industry$ effects we employ one-digit SIC code dummies. Bharath and Shumway (2008) find that PD adds predictability to accounting-based bankruptcy models, and Guttler and Wahrenburg (2007) report that Moody's and S&P adjust their ratings based on increase in firm default risk, hence we include PD from the Merton model in the regressions. This variable is directly linked to rating levels and also accounts for Amato and Furfine's (2004) critique of Blume et al. (1998). We expect this variable to be positive and significant.

Table 4 provides the summary statistics of the variables utilized in the ordered probit rating model. Sixty-nine percent of the observations of the sample belong to the manufacturing, transportation and communications industries (SICs 2, 3, and 4), as defined in Appendix B. There are nine industries and therefore eight industrial control dummies. All reported results are with respect to the agriculture and forestry industry, which is coded 0. The mean rating over the entire period is BBB. The four interest coverage ratios (ICR) are for the ranges $(-\infty, 5)$, $[5, 10)$, $[10, 20)$ and $[20, \infty)$. In the Table, interest coverage ratios 2, 3, and 4 are the average values above the baseline values of 5, 10 and 20, respectively. Thus, ICR2 has a mean of 7.2579 (= 5+2.2579). The ICR is binned into intervals to account for the piece-wise-linear marginal effect of ICR on ratings (Blume et al., 1998). The average PD over a time period of five years (i.e., the average time to maturity in Merton model) is 10%. A year-by-year analysis fails to show any appreciable monotonic trend in the financial statement variables.[16] The quarterly correlation over the sample period between PD and Lev is 42%, implying that PD captures other risk effects beyond leverage.

TABLE 4
SUMMARY STATISTICS OF VARIABLES USED IN THE AGENCY RATING MODEL

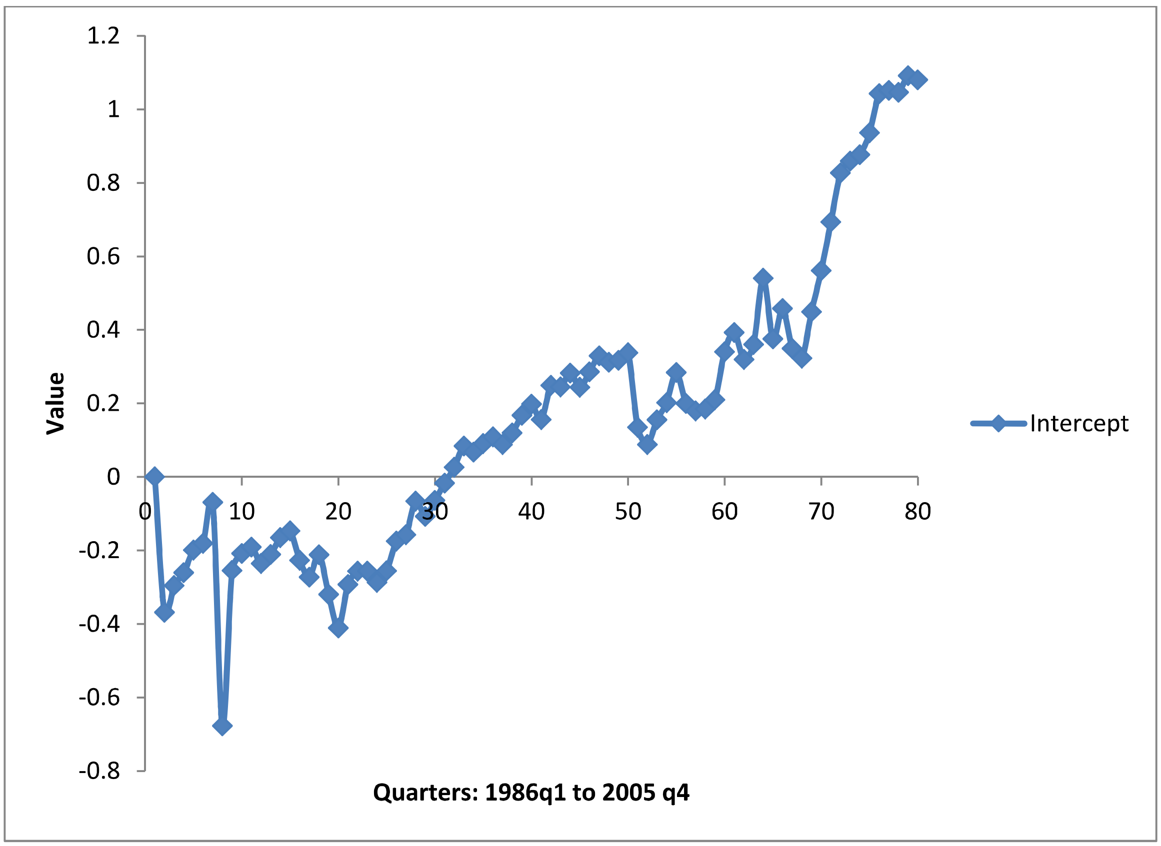
Variable	Mean	Std Dev	Minimum	Maximum
Dependent Variable				
Numerical Code S & P Rating	4.0813	1.2526	1	7
Balance Sheet Variables				
Operating Income to Debt Ratio	0.2131	0.325	-0.1042	4.3617
Operating Income to Sales Ratio	0.1745	0.1248	-0.3853	0.6695
Return on Assets	0.0180	0.0205	-0.0892	1.0048
Long-Term Debt to Assets Ratio	0.2887	0.1513	0	0.8971
Interest Coverage Ratio 1 (-∞,5)	4.4109	1.1379	-2.7903	5
Interest Coverage Ratio 2 [5,10)	2.2579	2.1536	0	5
Interest Coverage Ratio 3 [10,20)	1.6468	3.253	0	10
Interest Coverage Ratio 4 [20,∞)	1.4605	7.158	0	91.5075
Control Variables (Cross - Section Indicator)				
Industry 1 Dummy	0.0608	0.2389	0	1
Industry 2 Dummy	0.2236	0.4167	0	1
Industry 3 Dummy	0.2589	0.4380	0	1
Industry 4 Dummy	0.2146	0.4105	0	1
Industry 5 Dummy	0.1328	0.3393	0	1
Industry 6 Dummy	0.0137	0.1162	0	1
Industry 7 Dummy	0.0643	0.2453	0	1
Industry 8 Dummy	0.0270	0.1622	0	1
Control Variable (Size)				
Log (Market Capitalization)	14.0994	1.6582	4.115	19.7496
Market Determined Probability of Default				
Firm Level PD	0.1004	0.1784	0	1
Variance Parameter Proxy				
Square Root Total Assets	55.7561	44.9268	0.2074	558.6242
Total Observations: 47004				

This table presents the summary statistics for the panel of variables used in modeling credit ratings. The Numeric Ratings are AAA = 1, AA =2, A = 3, BBB = 4, BB = 5, B = 6 and CCC & below = 7. The Operating Income to Debt Ratio is the ratio of operating income before depreciation to the total long-term debt. Operating Income to Sales ratio equals operating income before depreciation by net sales. Return on Assets equals Net Income plus Interest Expense divided by total assets. Long-term debt to assets is the ratio of long-term debt on the balance sheet to total assets. The interest coverage ratio (ICR) is the ratio of the sum of operating income before depreciation and interest expense to interest expense. We t is divided into four ranges to account for possible piece-wise-linear marginal effects. The four interest coverage ratio ranges are (-∞, 5), [5, 10), [10, 20) and [20, ∞). The means given for each range are the means of the ratio above the baseline value of the range. For example, an ICR of 2 in the second range will generate a mean of 7.589 (that is, 5+2.2589). Similarly, the min and max in each ICR range is obtained by subtracting the limits of that range from the lower limit of the range. This is equivalent to equating the lower limit of each range to zero. Cross-section indicator variables equal 1 if the firm belongs to a particular one-digit SIC code or industry. Because there are nine one-digit SIC codes, there are eight dummies. Log (Market Capitalization) is the natural log of the market value of all shares as at the end of that quarter. Firm-level PD is the market-determined probability of default, as calculated according to the Merton model and adjusted for risk. Total Assets indicate the assets of a firm, and its square root is used to model heteroskedasticity. The calculation of PD is explained in the main body and Appendix A of this paper, and industry classification is explained in the Appendix B.

In the first step of the model of ratings, we run a pooled ordered probit with time dummies for each of the 80 quarters with fixed slope coefficients, as carried out in Blume et al. (1998) and Amato and Furfine (2004). The plot of the quarterly dummies (intercept term) from this regression is shown in Figure 2. The results of the ordered probit regression for the full sample over the entire period 1986-2005 are shown in Table 5.

If the intercept captures the rating standards, a rising intercept implies that for a given level of firm risk, the firm would be assigned a lower letter grade in later years.[17] The intercepts in the figure exhibit an upward time trend, consistent with findings in earlier papers, (e.g., Blume et al. 1998; Amato and Furfine, 2004; Baghai et al. 2014). However, as noted earlier, the intercepts encapsulate effects of all omitted variables. Moreover, forcing a fixed slope does not account for variations in the weights of different variables across time. Therefore, our analysis of rating standards is based on thresholds obtained from quarter-by-quarter regressions.

FIGURE 2
PLOT OF THE TIME INTERCEPTS OBTAINED FROM POOLED ORDERED PROBIT MODEL OF AGENCY RATINGS



This figure plots the time intercepts for the agency rating model, $y_{it} = \alpha_t + \beta X_{it} + \varepsilon_{it}$, where $E(\varepsilon_{it}) = 0; and\ E(\varepsilon_{it}^2) = [\exp(\gamma W_{it})]^2$, and y_{it} is the issuer rating of firm i at time t. y_{it} is 1 if rating is AAA, y_{it} equals 2 if rating is AA, y_{it} is coded 3 if rating equals A, y_{it} is 4 if rating is BBB, y_{it} is assigned a value 5 if rating equals BB, y_{it} equals 6 if rating equals B, and y_{it} is 7 if rating is CCC and below. The figure mirrors the intercept trend observed in previous studies by Blume et al (1998), and Amato and Furfine (2004). Unlike previous studies, however, the intercepts are not equated to rating standards in this paper. X_{it} is the vector of determinant variables provided in Table 4, while W_{it} equals square root of total assets of the firm and is used to model heteroskedastic errors. The period is from Yr1986q1 – Yr2005q4 totaling eighty quarters, and the model is implemented using ordered probit.

Table 5 shows that all variables in the ordered probit regression are significant and have the expected signs. For example, Lev is positive and significant. Similarly, other balance sheet variables have the expected sign. Very high interest coverage ratios do not contribute significant information to the ratings, and thus, the magnitude of very high ICR (>20) on ratings is small (0.0019) but has the expected negative sign. Also, because PD predicts the future default risk *a priori*, it should be highly positively correlated with lower letter ratings, which are a subjective measure of the likelihood of default by a firm on its debt obligations in the future. As expected, the sign of firm-level PD is both positive and significant at the 1% level, confirming this intuitive correlation. The $size$ variable shows that larger firms are more likely to have higher letter ratings.

TABLE 5
RESULT OF THE PANEL ORDERED PROBIT WITH TIME-VARYING INTERCEPTS

Dependent variable: Numerical rating class (1 to 7)			
Variable	**Coefficient**	**t -stat**	
Intercept	9.6510	81.14	***
Time Dummies by Quarter (Yr1986q2 to Yr2005q4)			
Balance Sheet Variables			
Operating Income to Debt Ratio	-0.0633	-3.74	***
Operating Income to Sales Ratio	-0.4837	-9.86	***
Return on Assets	-1.2789	5.31	***
Long-Term Debt to Assets Ratio	1.9367	40.65	***
Interest Coverage Ratio 1 (-∞,5)	-0.1357	-24.55	***
Interest Coverage Ratio 2 [5,10)	-0.0569	-15.35	***
Interest Coverage Ratio 3 [10,20)	-0.0079	-3.49	***
Interest Coverage Ratio 4 [20,∞)	-0.0019	-2.71	***
Control Variables(Cross - Section Indicator)	**8 Industry**	**Dummies**	
Control Variable (Size)			
Log (Market Capitalization)	-0.4181	-103.27	***
Market-Determined Probability of Default			
Firm Level PD	1.1812	39.74	***
Variance Parameter Proxy			
Square root of Total Assets	-0.0008	-12.73	***
Threshold Parameter Values			
Upper Limit for AA	1.2249	54.30	***
Upper Limit for A	2.5485	93.59	***
Upper Limit for BBB	3.6836	119.43	***
Upper Limit for BB	4.8896	143.02	***
Upper Limit for B	6.9496	166.55	***
N Obs	47004		
Log-Likelihood	-76232		

***,** and *: Significant at 1%, 5% and 10% level of significance, respectively.

This table reports the results for the pooled ordered probit regressions on the entire sample for the period from 1986 to 2005. There are 80 quarters and, hence, 79 quarterly dummies. The plot of the intercept dummies is given in Figure 2. All quarterly constants are measured with respect to Yr1986q1, which is set at 0. Similarly, all industry coefficients are reported with respect to the agriculture and forestry industry. The numeric ratings are AAA = 1, AA =2, A = 3, BBB = 4, BB = 5, B = 6 and CCC & below = 7. The Operating Income to Debt Ratio is the ratio of operating income before depreciation to the total long-term debt. The Operating Income to Sales ratio equals operating income before depreciation by net sales. Return on Assets equals Net Income plus Interest Expense divided by total assets. Long-term debt to assets is the ratio of long-term debt on the B/S to the total assets. The interest coverage ratio (ICR) is the ratio of the sum of operating income before depreciation and interest expense to interest expense. We t is divided into four ranges to account for possible piece-wise-linear marginal effects. The four interest coverage ratio ranges are (-∞, 5), [5, 10), [10, 20) and [20, ∞). The means given for each range are the means of the ratio above the baseline value of the range. For example, an ICR of 2 will have a mean of 7.2450 (that is, 5+2.2450). Similarly, the min and max in each ICR range is obtained by subtracting the limits of the range from the lower limit of the range. This is equivalent to equating the lower limit of each range to zero. Cross-section indicator variables equal 1 if the firm belongs to a particular one-digit SIC code or industry. Because there are nine one-digit SIC codes, there are eight dummies. The industries are tabulated in the Appendix B of this paper. Log (Market Capitalization) is the natural log of market value of all shares as at the end of that quarter. Firm-level PD is the market-determined probability of default as calculated according to the Merton model and adjusted for risk. Total Assets include assets of a firm, and its square root is used to model heteroskedasticity. PDs are computed from the Merton model of default, including equity as a call option and the optimal hedge equation. The implementation of the Merton model is described in the Appendix A.

Results on Rating Standards and Economy Wide Default Risk Distribution

To analyze the relation between rating standards and credit risk distribution, we run the ordered probit model of agency ratings for 80 quarters separately. Each quarterly regression yields a six-tuple vector of thresholds for the seven rating categories, thus leading to a time series of thresholds. As stated earlier, the longitudinal series of thresholds represents the rating standards over time.

Next, we test if the rating standards are pegged to the distribution of the underlying quality of firms. Because previous papers have reported that standards exhibit a secular tightening of standards over time, the basic model takes the following form.

$$\mu_{jt} = f(t, \bar{\theta}, \sigma_\theta)_t, \tag{6}$$

where j is the threshold between rating categories $j - 1$ and j, and varies from 1 to 6. t = 1..80 quarters starting from 1986q1 and ending with 2005q4. $\bar{\theta}$ is the mean credit quality, and σ_θ is its dispersion measure at time t. As given in equation (1), the standards are positively linked to $\bar{\theta}$ and negatively to σ_θ .

Although equation (6) is sufficient to test the theoretical link between rating standards and distribution quality, it does not incorporate the statistical finding that ratings are sticky. Ratings are sticky for the reason that agencies attempt to provide rating stability and do not alter them as frequently as the market expects (for a review, see Loffler, 2004). To capture this empirical fact, we modify equation (6) to include the stickiness of the rating standards. Consequently, the following regression is adopted.

$$\mu_{jt} = f\left(t, \bar{\theta}, \sigma_\theta, \mu_{jt-1}\right)_t, \tag{7}$$

where lagged values of the thresholds are included to replicate stickiness in ratings. Equations (6) and (7) are estimated using cross-sectional fixed effects for j.

The period-by-period regressions yield a panel of thresholds and intercepts. If intercepts mirror the rating standards, then visual analysis reveals that the changes in standards over time have not been linear. There have been periods of both tightening and relaxation in standards. Hence, in the algebraic analysis of equations (6) and (7), we include both linear and squared time terms.

The results of regression equations (6) and (7) are shown in Table 6 in Specifications 1 and 2, respectively. There are six thresholds (and hence, six cross-sectional units), and to identify the model, the intercept term is suppressed.

Specification 1 reveals that there is a non-linear relationship between rating standards and time. The linear trend is negative, while the coefficient on the square term is positive but small in magnitude. The marginal effect of time on rating standards is $-0.0337 + 2 \times 0.004t$, which remains negative for 42 quarters, after which it becomes positive. Thus, under this specification, the tightening of rating standards can be observed until the first half of 1996, i.e., over the period also covered in Blume et al's (1998) study. After that, however, the marginal effect becomes positive, implying a secular relaxing of standards thereafter.

In line with expectations, the marginal effect of the mean quality is positive (slope is 6.60), while that of dispersion is negative (slope equals -13.62), both significant at the 1% level.

Specification 2 modifies Specification 1 and includes the lagged value of the standards and is consistent with modeling of credit rating transitions (Lando and SkØdeberg, 2002). As ratings are known to react with a lag, this term captures the stickiness in ratings. The term is positive (0.53) and significant at the 1% level. Because the previous standard takes into account historical behavior, the time variable becomes insignificant. The marginal effect of mean (dispersion) of PD decreases (increases) in magnitude from 6.6 (-13.62) to 4.86 (-8.01) but remains significant at the 1% level.

TABLE 6
RATING STANDARDS AND THE UNDERLYING QUALITY OF THE UNIVERSE OF FIRMS

	Spec 1		Spec 2	
CS Fixed Effects	YES		YES	
Time	-0.0337	***	-0.0089	
	(-3.75)		(-1.16)	
$Time^2$	0.0004	***	0.0001	
	(3.73)		(1.24)	
Mean PD	6.6021	***	4.8639	**
	(2.68)		(2.33)	
Std PD	-13.6157	***	-8.0052	**
	(-3.25)		(-2.23)	
lagthresh			0.5239	***
			(13.62)	
R^2	0.9016		0.9011	
N Obs	480		474	

***,** and *: Significant at 1%, 5% and 10% levels of significance, respectively.

This panel shows the results of the fixed effect (cross-sectional) regressions of thresholds. The regression model is $\mu_{jt} = f(t, \bar{\theta}, \sigma_\theta,)_t$ in Specification 1 and $\mu_{jt} = f\left(t, \bar{\theta}, \sigma_\theta, \mu_{jt-1}\right)_t$ in Specification 2. Here μ_{jt} denotes the rating thresholds between categories j and $j+1$ at time t. The six cross-sectional units correspond to the rating thresholds between the seven rating categories AAA and AA, AA and A, A and BBB, BBB and BB, BB and B, B and CCC & below. $\bar{\theta}$ and σ_θ denote the mean and standard deviations of the probability of default (PD) distribution, respectively. Because there are six cross-sectional units (or thresholds), the intercept term is suppressed for identification purposes. Specification 2 captures the stickiness in ratings. Lagthresh refers to the threshold value in the previous time period. CS Fixed Effects are six intercept terms for the cross-section of six thresholds. The panel of thresholds is obtained by running the ordered probit regressions of the agency rating model for 80 quarters (that is, from Yr1986q1 to Yr2005q4).

The findings reveal that rating standards have varied over time, and previous standards are a dominant determinant of current standards, consistent with the observation that ratings are sticky. Moreover, the standards are linked to the underlying default risk distribution. Therefore, a broad conclusion that standards have been made stricter (or more lenient) across the board cannot be supported.

Robustness Checks

We subject the findings in Tables 6 to different measures of location and dispersion parameters with respect to the credit quality distribution. This is necessary because results in Tables 2 and 3 point to a highly skewed distribution of PDs. In particular, we use the median and 60th percentile for location, and inter-quartile range and mean absolute deviation for the dispersion in the credit quality distribution. The results are qualitatively similar to those obtained with mean and standard deviation.

Furthermore, verification of the relationship between standards and distribution of PDs reveals it is not an artifact of the Merton model. We also utilize a down-and-out (DOC) call option model of default (Reisz and Perlich, 2007) in place of the Merton model for a sub-sample (Years 1986 to 2000) and find that the results are qualitatively similar.[18] The mean and median, and, dispersion and range obtained from the DOC option model display a similar relationship to standards. The findings of Tables 6 are thus robust to alternative models of the underlying distribution of firm quality.[19]

A potential critique of the findings is we do not control for the effects of the economic cycle and through the cycle rating methodology. Consequently, we include an economic expansion dummy to account for business cycle effects. The dummy takes value 1 for a quarter if all months in a calendar quarter are documented as expansionary by the National Bureau of Economic Research (NBER), 0 otherwise.

Loffler (2004) notes that rating agencies rate through the cycle. That is ratings change only to long-term (or permanent) changes in credit quality, not short-term (or transitory) fluctuations. We uses a 37 period centered moving average (CMA) to compute the long-term components of credit quality of the mean and the variance of the economy wide credit quality distribution.

The results of incorporating the business cycle effects and taking into account only the permanent components of the parameters of credit risk distribution are shown in Table 7. The mean and standard deviation continue to be statistically significant in Spec 1, and have the same sign as hypothesized. The business cycle dummy is statistically insignificant. The marginal effects of long term components are comparatively higher in Specification 2, and all variables are statistically significant.[20]

TABLE 7
ROBUSTNESS CHECKS FOR RATING STANDARDS AND THE UNDERLYING CREDIT QUALITY DISTRIBUTION

	Spec 1		Spec 2	
CS Fixed Effects	YES		YES	
Time	-0.0095		0.3272	***
	(0.00)		(0.10)	
Time^2	0.0001		-0.004	***
	(0.00)		(0.00)	
Mean PD	4.6853	**		
	(2.10)			
Std PD	-7.4365	**		
	(3.66)			
Perm Comp Mean PD			239.0389	***
			(64.73)	
Perm Comp Std PD			-386.631	***
			(98.52)	
lagthresh	0.5189	***	0.2271	***
	(0.04)		(0.06)	
Bus Cycle Dummy	0.1161		0.4378	**
	(0.14)		(0.20)	
R^2	0.9812		0.9859	
N Obs	474		258	

***,** and *: Significant at 1%, 5% and 10% levels of significance, respectively.

This panel shows the results of the robustness checks for results in Table 6. Specification 1 captures the business cycle effects with the Bus cycle dummy = 1 if all three months in a calendar quarter are expansionary, else 0. All other variables of Specification 1 are explained in Table 6. In Specification 2, Perm comp Mean PD and Perm comp Std PD are the centered moving average of Mean PD and square root of variance PD of 37 quarters, centered at the nineteenth quarter. Correspondingly, for each threshold the length of time series is reduced to 43. The models are explained in Table 6.

CONCLUSION

In this paper, we study rating standards (not ratings). The profit maximization behavior of rating agencies dictates that they are anchored to the underlying distribution of the quality of the universe of firms that they can only discover with some imprecision. Tests using S & P issuer ratings confirm this relationship. While previous studies have sought to answer whether standards have become more stringent or lenient over time, the notion of anchored yet changing rating standards leads to the conclusion that the answer to the question is not so straightforward. To the extent that there is no trend, positive or negative, or cyclicality in the underlying distribution of credit quality, we cannot expect to find any in the rating standards.

We find that the effect of decline in mean of credit quality on standards is the opposite of the increase in dispersion in credit quality. With an increase in dispersion, standards tighten. These findings are robust to alternative specifications of the relationship between rating standards and underlying credit quality, the measures of location and dispersion of the credit distribution, controls employed for business cycle effects, and the model used to quantify the underlying credit distribution. The findings are consistent with the argument that rating agencies face a trade-off between market share (short term gains) and reputational value (long term costs).

The findings add to the debate on the usefulness of ratings, since the paper assumes that rating agencies cannot detect true quality of the firm with 100% precision. By no stretch of imagination does our study suggest rating agencies do not reduce information asymmetry. But our findings are in line with previous studies that question information content of ratings. Additionally, the examination shows when markets do not know due to increase in uncertainty, the rating agencies also do not know and they take steps to protect their reputational capital. During such times, the belief of investors in ratings becomes critical as argued in Boot, Milbourn and Schmeits (2006), even though the information content suffers.

Finally, even though we have analyzed the behavior of credit rating agencies, the paper has relevance for studies of other certifiers in financial markets facing problems of information asymmetry. A theoretical literature in finance examines the certification role of investment banks, venture capitalists, and newsletter producers (see Jin et al., 2004 for a review). This paper sheds light on rating agency standards by undertaking an empirical analysis. Similar research can be undertaken for certification intermediaries in other markets in the future, and the results compared and contrasted with findings in this one.

ACKNOWLEDGEMENT

We thank Neil Doherty, Otgo Erhemjamts, Manu Gupta, Jayant Kale, Sven Kepes, Spyros Pagratis, Srikanth Paruchuri, Richard Phillips, Sanjay Srivastava, Ajay Subramanian, and participants at various conferences for helpful comments. Usual disclaimer applies.

ENDNOTES

1. Typically, dispersion is measured for a random variable. This measure implies probability (of default) itself is random, so we label it as uncertainty for lack of a better term.
2. Given we employ Standard and Poors (S & P) ratings in our study, we prefer to use the nomenclature followed by S & P. AAA represents the highest quality amongst investment-grade firms, while BB represents the same in non-investment-grade firms. We ignore +/- qualifiers in the study, and group AAA+, AAA and AAA- firms together as AAA. Similarly, we group firms rated BB+, BB and BB- into the BB class.
3. We also employ the Bharath and Shumway (2008) method when we carry out robustness checks.
4. We quote from the abstract of the published study. Again, Eom, Helwege and Huang (2004) study different structural models (including Merton model) to examine which model explains the observed market credit spreads best. We use Merton model as it serves our purpose in this paper. However, to rule out the critique that our findings are an artifact of the model, we also carry out robustness checks using the down-and-out call option model of Reisz and Perlich (2007).
5. Note that there are several ways of determining the Merton probability of default (see Bharath and Shumway, 2008). The paper is not about determining the exact probability of default, but the trends in mean and standard deviation. Our calculations track the credit spreads quite closely. Comparisons are available upon request from the authors. In the robustness tests we extend the data and employ the Bharath and Shumway (2008) approach.
6. Professor French's data URL is: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
7. Blume et al. (1998) allowed the slope coefficients to vary as part of their robustness checks.
8. Christie, L., May 16, 2006. Real estate cools down. *CNNmoney.com*
9. A plot of computed PDs and credit spreads is available from the authors.

10. While these numbers represent defaults on bond issues, we reiterate that the present study is about issuer ratings.
11. Long Term Debt/Total Assets
12. Net Income + Interest Expense /Total Assets
13. Operating Income Before Depreciation /Net Sales
14. Operating Income Before Depreciation/Long-Term Debt
15. Operating Income Before Depreciation + Interest Expense / Interest Expense
16. Yearly summary statistics of the variables provided in Table 4 are available from the authors.
17. In this study, higher letter grades are assigned lower integers, similar to Amato and Furfine (2004) but in contrast to Blume et al. (1998).
18. The DOC option model is a four equation-four unknown model comprising transcendental equations. Reisz and Perlich, 2007 utilize IBM computers with 10 processors running for 72 hours to obtain convergence for 53% observations. Owing to the computational complexity involved and only desktop computing available to us, we restrict the sample to 1986-2000 for this part of robustness check.
19. The results of robustness checks are not included but are available from the authors.
20. We also sample data for the period 2006-2013 using same data sources and employing the Bharath and Shumway (2008) methodology to compute PDs. However, data on which firms undertake risk management (S&P criterion after 2005) is unavailable. We find similar results as reported in this paper and are available from the authors.

REFERENCES

Altman, E.I. (1989). Measuring corporate bind mortality and performance. *Journal of Finance*, 44, 909-922.

Altman, E.I., & Rijken, H. A. (2004). How Rating Agencies Achieve Rating Stability? *Journal of Banking & Finance*, 28, 2679-2714.

Altman, E.I., & Rijken, H. A. (2005). The Impact of the Rating Agencies' Through-the-cycle Methodology on Rating Dynamics. *Economic Notes*, 34(2), 127-54.

Amato, J.D., & Furfine, C. H. (2004). Are Credit Ratings Pro-cyclical*? Journal of Banking and Finance*, 28, 2641-2677.

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

Bar-Isaac H., & Shapiro, J. (2011). Credit ratings accuracy and analyst incentives. *The American Economic Review*, 101, 120-124.

Bar-Isaac H., & Shapiro, J. (2013). Ratings quality over the business cycle. *Journal of Financial Economics*, 108, 62-78.

Becker, B., & Milbourn, T. (2011). How did increased competition affect credit ratings? *Journal of Financial Economic*, 101, 493-514.

Berger, A., & Udell, G. (2004). The institutional memory hypothesis and the procyclicality of bank lending behavior. *Journal of Financial Intermediation*, 13, 458–495.

Bernoth, K., & Pick, A. (2010). Forecasting the fragility of the banking and insurance sectors. *Journal of Banking & Finance*. doi:10.1016/j.jbankfin.2010.10.024

Bharath, S.T., & Shumway, T. (2008). Forecasting Default with the Merton Distance to Default Model. *Review of Financial Studies*, 21, 1339-1369.

Blume M.E., Lim F., & MacKinlay C. A. (1998). The declining Credit Quality of U.S. Corporate Debt: Myth or Reality? *Journal of Finance*, 53, 1389 – 1413.

Bolton, P., Freixas, X., & Shapiro, J. (2012). The credit ratings game. *Journal of Finance*, 67, 85-111.

Boot, A. W., Milbourn, T. T., & Schmeits, A. (2006). Credit ratings as coordination mechanisms. *Review of Financial Studies*, 19, 81–118.

Brockman, P., & Turtle, H. J. (2003). A Barrier Option Framework for Corporate Security Valuation. *Journal of Financial Economics*, 67, 511-529.

Delianedis, G., & Geske, R. (1999). *Credit Risk and Risk Neutral Default Probabilities: Information About Rating Migrations And Defaults*. Working Paper, The Anderson School, UCLA.
Diamond, D. W. (1991). Debt maturity structure and liquidity risk. *Quarterly Journal of Economics*, 106, 709-737.
Dimson, E. (1979). Risk Measurement When Shares Are Subject to Infrequent Trading. *Journal of Financial Economics*, 7, 197-226.
Doherty, N. A., Kartasheva, A.V., & Phillips, R.D. (2012). Information effect of entry into credit ratings market: The case of insurers' ratings. *Journal of Financial Economics*, 106, 308-330.
Ebnother S., & Vanini, P. (2007). Credit portfolios: What defines risk horizons and risk measurement? *Journal of Banking & Finance*, 31, 3663–3679.
Eichengreen, B., Hale, G., & Mody, A. (2001). Flight to quality: Investor risk tolerance and the spread of emerging market crises. In S. Claessens & K. J. Forbes (Eds.), *International Financial Contagion* (pp. 129–55). Kluwer Academic Publishers, MA.
Eom, Y.H., Helwege, J., & Huang, J-Z. (2004). Structural models of corporate bond pricing: An empirical analysis. *Review of Financial Studies*, 17, 499-544.
Fama, E., & French, K. (1997). Industry Costs of Equity. *Journal of Financial Economics*, 43, 153-193.
Ferri, G., Liu, L., & Stiglitz, J.E. (1999). The procyclical role of rating agencies: Evidence from the east Asian crisis. *Economic Notes*, 28, 335–355.
Goel, A. M., & Thakor, A.V. (2015). Information reliability and welfare: A theory of coarse credit ratings. *Journal of Financial Economics*, 115, 541-557.
Greene, W.H. (2003). *Econometric Analysis* (5th ed.). International Edition. NJ- Prentice Hall.
Guttler, A., & Wahrenburg, M. (2007). The adjustment of credit ratings in advance of defaults. *Journal of Banking & Finance*, 31, 751–767.
Hillegeist S.A., Keating E.K., Cram, D.P., & Lundstedt, G. (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, 9, 3-54.
Holden, S., Natvik, G. J., & Vigier, A. (2014). *An Equilibrium Theory of Credit Rating*. Working paper, University of Oslo.
Jeon, D-S., & Lovo, S. (2013). Credit rating industry: A helicopter tour of stylized facts and recent theories. *International Journal of Industrial Organization*, 31, 643-651.
Jin, G., Kato, A., & List, J.A. (2004). *Evolution of professional certification markets: Evidence from field experiments*. Working Paper, University of Maryland and NBER.
Lando, D., & Skødeberg, T. (2002). Analyzing Rating Transitions and Rating Drift with Continuous Observations. *Journal of Banking and Finance*, 26, 423-444.
Leland, H.E. (2004). *Predictions of default probabilities in structural models of debt*. Working paper University of California, Berkeley.
Lizzeri, A. (1999). Information Revelation and Certification Intermediaries. *RAND Journal of Economics*, 30, 214-231.
Loffler, G. (2004). An Anatomy of Rating through the Cycle. *Journal of Banking and Finance*, 28, 695-720.
Loviscek, A.L., & Crowley, F.D. (1990). What is in a Municipal Bond Rating? *The Financial Review*, 25, 25-53.
Manso, G. (2013). Feedback effects of credit ratings. *Journal of Financial Economics*, 109, 535–548.
Mathis, J., McAndrews, J., & Rochet, J-C. (2009). Rating the raters: are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics*, 56, 657–674.
Merton R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29, 449-470.
Opp, C.C., Opp, M. M., & Harris, M. 2013. Rating agencies in the face of regulation. *Journal of Financial Economics*, 108, 46–61.
Reisz, A.S., & Perlich, C. (2007). A Market-based Framework for Bankruptcy Prediction. *Journal of Financial Stability*, 3, 85-131.

Ronn, E., & Verma, A. (1986). Pricing Risk Adjusted Deposit Insurance: An Option Based Model. *The Journal of Finance*, 41, 871-895.
Ruckes, M. (2004). Bank competition and credit standards. *The Review of Financial Studies*, 17, 1073–1102.
Skreta, V., & Veldkamp, L. (2009). Ratings shopping and asset complexity: A theory of ratings inflation. *Journal of Monetary Economics*, 56, 678–695.
Stohs, M.H., & Mauer, D.C. (1996). The Determinants of Corporate Debt Maturity Structure. *The Journal of Business*, 69, 279-312.
Strausz, R. (2005). Honest Certification and the Threat of Capture. *International Journal of Industrial Organization*, 23, 45-62.
Thakor, A.V. (2015a). *The highs and the lows: a theory of credit risk assessment and pricing through the business cycle*. Working Paper, Washington University in St. Louis.
Thakor, A.V. (2015b). The financial crisis of 2007-2009: Why did it happen and what did we learn? *Review of Corporate Finance Studies*, 4, 155-205.
White, L. J. (2010). Markets: The credit rating agencies. *Journal of Economic Perspectives*, 24, 211–226.
Xie Y.A., Liu, S., Wuc, C., & Bing, A. (2009). The effects of default and call risk on bond duration. *Journal of Banking & Finance*, 33,1700–1708.

APPENDIX A

Implementation of the Merton Model

The strike price is M = total liabilities – 0.5 current liabilities. This is similar to the strike price considered by Delianades and Geske (1999).

The time to maturity of liabilities is calculated using the Macaulay duration. While this equals one year for the Moody's KMV model, we prefer to use the Macaulay duration for two reasons. First, we focus on long-term ratings. As per S & P, this reflects the ability of the obligor to meet liabilities beyond one year, even though the horizon is not exactly defined. Also, we know from Altman and Rijken (2004) that the RTTC focuses on long-term default horizon. Second, the choice of the Macaulay duration as time to maturity accounts for the gap between the ratings and the market Probabilities of Default (PDs) arising out of a point-in-time versus a long-term rating-through-the-cycle perspective (again, see Altman and Rijken, 2005 who use six years for long term default horizon in an attempt to bridge this gap). Ebnother and Vanini (2007) have argued that risk of credit portfolios be analyzed in multi-period set-up because the effect of shocks lasts for more than one period. The current liabilities are assumed to have a maturity of 6 months, and long-term liabilities are assumed to have a 10-year maturity.

Computation of Asset Betas

We run a regression of daily excess returns of the following form for all firms in the sample for the year prior to the current quarter to account for the effects of non-synchronous trading (Dimson, 1979).

$$R_{it} - R_{ft} = \alpha_0 + \alpha_1\left(R_{it+2} - R_{ft+2}\right) + \alpha_2\left(R_{it+1} - R_{ft+1}\right) + \\ \alpha_3\left(R_{it} - R_{ft}\right) + \alpha_4\left(R_{it-1} - R_{ft-1}\right) + \alpha_5\left(R_{it-2} - R_{ft-2}\right) + \varepsilon_{it} \qquad \text{(A.1)}$$

Equity beta for firm i is given by $\beta_i = \sum_{k=1}^{5} \alpha_k$. Then we recombine individual firm betas into Fama and French (1997) industry portfolio betas β_p, where for each of the 48 industries, $\beta_p = \frac{1}{n}\sum_{i=1}^{n} \beta_i$, and n is the number of firms in the industry portfolio in that quarter. We then form an asset beta (that is, an unlevered beta) for firm i using the following equation.

$$\beta_i{}^A = \frac{\beta_p}{1+(D/_E)_i} \qquad \text{(A.2)}$$

$D/_E$ is the financial leverage for firm i in the quarter. D/E is equal to long-term debt / (BV assets – BV liabilities) from COMPUSTAT.

Market Risk Premium

The expected daily market risk premium at the end of the month is the average of daily actual excess returns of last three months. It is then multiplied by 252 to yield an annualized estimate of the expected market risk premium.

The market risk premium and asset betas are utilized to compute the expected return $\mu = r + \beta_A[E(R_m) - r]$, which is input into equation (4') to calculate the risk-adjusted probability of default (PD).

Computation of Volatility

Total volatility of equity returns for firm i is the standard deviation of daily firm returns (σ_i) computed on a rolling basis using 60 days of historical data, thereby approximating the number of trading days in the last quarter. The return on day t equals $\ln\left(\frac{P_t}{P_{t-1}}\right)$, where ln equals the natural log. To annualize the computed volatility, it is multiplied by the square root of 252.

Initial Values

The price at the close of the quarter is multiplied by the number of shares outstanding (that is, the variable SHROUT) to obtain an estimate of the market value of stock. To this, the book value of debt is added. The final value is used to initialize the market value of assets to solve equations (3) and (4') simultaneously. The annual volatility of equity returns at the end of the quarter is used as the initial value for asset volatility.

APPENDIX B

DESCRIPTION OF INDUSTRIES IN THE SAMPLE BY ONE-DIGIT SC CODE

1 – digit SIC code (DNUM from Compustat)	Industry description	Industry code dummies
0	Agriculture and Forestry	-
1	Minerals, Metal Mining, Bldg and Construction	Dummy 1
2	Manufacturing – Wood, Paper, Food and Chemicals	Dummy 2
3	Manufacturing – Rubber, Clay and Miscellaneous	Dummy 3
4	Transportation and Communications	Dummy 4
5	Durable goods	Dummy 5
6	Financial Services	Dummy 6
7	Hotels, Motion Pictures	Dummy 7
8	Legal, Educational and Social Services	Dummy 8
9	Government and Public Administration	Deleted from sample