Comprehending China's Domestic Agency Ratings: A Perspective from Default Probability-Implied S&P Ratings

Shida Liu* Hao Wang†

Abstract

We establish a mapping between the Chinese domestic agency ratings and S&P global ratings by matching firms' expected default probabilities (PDs) estimated using a dynamic logit model with the actual default rates of S&P ratings. The AAA, AA, and A ratings assigned by the domestic agencies correspond to S&P BB+, BB, and BB- by median default probability, suggesting that the agency ratings are inflated by 11 notches on average in the light of actual default probability and the S&P rating standard. The PD-implied ratings outperform the agency ratings in predicting default and complement the latter explaining credit spreads. The superior default predictive power originates from their use of more dynamic fundamental and stock information. In contrast, the agency ratings give more weights to static firm characteristics.

Key Words:
Dynamic Logit Model, Actual Default Rate, Receiver Operating Characteristic Curve, Accuracy Ratio, Rating Inflation, Machine Learning.

JEL: G21, G24, G28

* Tsinghua University, School of Economics and Management, Beijing, China, 100084; Email: liushd.17@sem.tsinghua.edu.cn;
† Tsinghua University, School of Economics and Management, Beijing, China, 100084; Email: wanghao@sem.tsinghua.edu.cn.

Yongcheng Coal & Electricity Holding Group Co., whose surprise payment failure in November (2020) triggered a slump in bonds of some other state-owned enterprises, held a AAA rating when it failed to repay at bond maturity by the deadline. It was cut to BB the following day.

---www.bloomberg.com

1. Introduction

In the wake of recent pickup in the surprising defaults of highly-rated Chinese firms, people have become growingly skeptical about the informativeness and effectiveness of Chinese domestic agency ratings. They ask the following questions: how do we make sense of the Chinese domestic agency ratings in the light of commonly accepted credit rating standards? How effective are the agency ratings in terms of monitoring default? Answers to these compelling questions are important to regulators, issuers, and investors that are increasingly drawn to the world's second-largest bond market (Schipke, Markus, and Zhang, 2019). However, it is not easy to answer these questions. The domestic agency ratings do exhibit some unusual features, e.g., rating clusters at the high-investment end between AAA and AA (Amstad and He, 2020; Jiang and Packer, 2019); there are disproportionately more upgrades than downgrades (Liu and Wang, 2020). Yet, one may still argue that the ratings are high and clustered because of different market conventions. The high upgrade-downgrade ratio can be explained by China's continued economic growth that alleviates firms' insolvency risk. Hence, these unique features may not necessarily undermine the ratings' informativeness and effectiveness.[1]

This work aims to comprehend China's domestic agency ratings by assessing their effectiveness in monitoring corporate defaults. To achieve the goal, we collect a comprehensive sample of listed firm defaults in China during 1998-2020 and establish a mapping between the domestic agency and S&P global ratings, using the latter for the benchmark rating standard. The mapping is established by matching firms' one-year expected default probabilities to S&P ratings' one-year actual default rates (ADR) with the following steps: we first create ADR bins of S&P ratings. We then follow Shumway (2001), Chava and Jarrow (2004) and Campbell, Hilscher, and Szilagyi (2008) to use a dynamic logit model to estimate firms' expected default probabilities. Lastly, we allocate firms into the ADR bins by expected default probability to pair up the domestic agency and S&P ratings.

We find that the domestic agency ratings are remarkably higher than S&P ratings for the same

[1] See Section 6 for further discussion.

level of default risk. The median expected default probabilities of firms with the domestic AAA, AA+, AA, AA-, A+, A, and Below-A ratings are 0.57%, 0.79%, 1.05%, 1.37%, 1.61%, 1.70% and 1.88%. In comparison, the actual default rates of S&P BB+, BB, BB- and B+ are 0.49%, 0.70%, 1.19% and 2.08%. The results suggest that the domestic AAA corresponds to S&P BB+; AA+ corresponds to S&P BB; AA corresponds to S&P BB; AA- corresponds to S&P BB-; A+ corresponds to S&P BB-; A corresponds to S&P BB-; Below-A corresponds to S&P BB-, respectively. On average, the agency ratings are higher than S&P ratings by 11 notches for the same level of default risk. The results give the first default-based evidence of the Chinese agency ratings being inflated in the light of a widely accepted rating standard.

Moreover, one domestic rating's default probabilities match the default rates of a broad range of S&P ratings. For example, the 10-to-90-percentile expected default probabilities of the domestic AAA, AA, and A ratings match the actual default rates of S&P BBB+ to B+ (seven notches), BBB- to B+ (five notches), and BBB- to B- (seven notches), respectively. The default probabilities of firms with identical domestic ratings vary significantly in a wide range, revealing that firms exposed to very different insolvency risks have clustered into the same rating categories. Such massive overlap would undermine the rating's primary purpose of information discovery (Ramakrishnan and Thakor, 1984; Millon and Thakor, 1985). It would also reduce credit ratings' capability to detect default because a financially distressed firm could have a high credit rating but suddenly default. The implication explains the recent pickup in the surprising defaults of highly-rated Chinese firms.

Since the ultimate goal of credit ratings is to discover insolvency risk, we compare the default predictability of the agency ratings to that of our PD-implied ratings. First, we follow Vassolu and Xing (2004) and Duan, Sun, and Wang (2012), and S&P (2020) to apply the Receiver Operating Characteristic (ROC) curve and Accuracy Ratio (AR) to assess their differentiability of default.[2] The ROC curve is a graphical plot illustrating the ratings' ability to differentiate default from no-default. The PD-implied ratings' ROC curve is significantly above the domestic agency rating counterpart, implying that the PD-implied ratings have superior default differentiability. The Accuracy Ratio is a numerical measure of a rating model's discriminatory power of default. A rating model can more accurately predict default if its Accuracy Ratio is closer to one. We find that the PD-implied ratings' Accuracy Ratio is 63.45%, which is approximately two times the domestic agency ratings' Accuracy

[2] We describe the ROC curve and Accuracy Ratio in Section 4 and Appendix B.

Ratio, 32.23%. Taking together, the PD-implied ratings can more accurately predict default than the domestic agency ratings can.

We then focus on defaulted firms to generate the before-default migration paths of the agency and PD-implied ratings. They exhibit strikingly different patterns. The defaulted firms' median domestic rating starts to fall slowly from AA eight months before default, reaching A two months and then BB one month before default. The pattern echoes the cliff-falling downgrade phenomenon observed in the Chinese bond market. In contrast, the median PD-implied rating of defaulted firms starts to fall gradually from BB- eight months to CCC one month before default. The PD-implied ratings' default signals are much more accurate and consistent than those given by the agency ratings. We also find that the PD-implied ratings complement the agency ratings explaining credit spreads in the subsequent month. The implied rating's explanatory power is economically significant in all agency rating categories, suggesting that the implied ratings constitute a valuable complement to the agency ratings, which are found coarse in Livingston, Poon, and Zhou (2018).

To understand where the PD-implied ratings' superior default predicting power originates from and what information they capture while the agency ratings do not, we carry out reverse engineering to regress the domestic agency ratings on our default probability predictors. The results show that the agency ratings give more weights to static firm characteristics, such as size and leverage ratio, in credit risk assessment. In contrast, the PD-implied ratings use more dynamic fundamental information, such as change in the pr ofitability and cash holding. Stock information is also processed differently--- the agency ratings emphasize stock idiosyncratic volatility, while the PD-implied ratings focus on the first-moment returns. Our findings help identify potential routes by which to improve agency rating strategy and methodology.

This paper appeals to a broad readership on several fronts. It first assesses the performance of the Chinese domestic agency ratings from the perspective of actual default and global rating standard.[3] Our findings will help those interested in international asset pricing and investing in the world's largest emerging market to more accurately assess the credit risk of Chinese entities. Our work will also help corporate finance analysis, as default risk has a fundamental role in firm operations, governance, and

[3] Jiang and Packer (2019) apply the comparable rank-ordering method to compare the ratings assigned by both domestic agencies (in the domestic market) and global agencies (in the overseas market) for bonds issued by the same Chinese entities. They document an average difference of 6–7 notches between the domestic and global credit ratings. Our works share the same spirit in establishing a match between the domestic and global ratings. We, however, take very different approaches. Our method is default-probability-oriented and applicable to a larger sample of general firms.

financial decisions. More specifically, our work adds to the burgeoning Chinese bond literature (Liu, Lyu, and Yu, 2017; Amstad and He, 2019; Ang, Bai, and Zhou, 2019; Geng and Pan, 2019; Chen, Chen, He, Liu, and Xie, 2020; Ding, Xiong, and Zhang, 2020). Although we focus on the Chinese agency ratings, our approach is general and can be used to research credit ratings in different markets and different rating schemes.

For broader literature, our findings advocate default-probability-based models that hold great potential to improve current rating practices, especially in an environment where traditional ratings are more likely to be compromised. We find that the machine learning models remarkably outperform the agency rating models and the Merton model in detecting default, adding support to Bharath and Shumway (2008), Campbell, Hilscher, and Szilagyi (2008), and Duan, Sun, and Wang (2012). More importantly, we identify the sources of superior default predictive power of machine learning models. Like the structural models, they use dynamic fundamental and stock information, but use the information more efficiently, in stark contrast to the traditional ratings that rely more on static characteristic information.

The remainder of our paper is organized as follows. Section 2 describes our empirical strategy. Section 3 analyzes the empirical results. Section 4 assesses the PD-implied rating in predicting default and explaining corporate bond yield spread. Section 5 implements a battery of robustness checks. Section 6 discusses the explanatory factors of rating inflation. Section 7 concludes the paper.

2. Empirical Strategy

Our primary purpose is to assess the domestic agency ratings' capability of measuring corporate default risk in the light of commonly accepted rating standards. To achieve the goal, we establish a mapping between the agency ratings and S&P global ratings by matching the firm's one-year expected default probability (PD) to the S&P rating's one-year actual default rate (ADR). We implement the mapping in the following steps: first, creating ADR bins of S&P ratings. Second, applying a dynamic logit model to estimate firms' expected default probabilities. Third, allocating firms into the ADR bins of S&P ratings by expected default probability to pair up domestic and S&P ratings.

2.1 ADRs of S&P Ratings

S&P ratings' one-year ADRs during 1981-2019 are reported in Table 9 of the *2019 Annual Global Corporate Default and Rating Transition Study*. Columns (2) to (4) of Table 1 duplicate S&P ratings and the means and standard deviations of their ADRs. The ADRs provide a quantitative measure of default risk underneath S&P ratings straightforwardly. For example, the ADR of S&P BBB is 0.21%, indicating that, on average, 21 out of 10,000 firms with S&P BBB ratings have defaulted in one year. The ADR increases monotonically as the S&P rating deteriorates.

[Insert Table 1 here.]

We note that the reported ADRs of S&P AAA and AA+ are zero because high investment-grade firms were typically downgraded one year before default. The ADRs of the low junk grades, CCC+ to C, are bundled together and reported in a single number of 24.34%. We need to obtain the ADR of every single S&P rating, to which we can later relate firms' expected default probabilities. We follow Duan and Li (2020) to interpolate the ADRs by first transforming them into the logarithm format, that is, $\mathrm{Log}(\frac{ADR}{100-ADR})$ to avoid generating negative ADR estimates and then apply Cubic Splines to create the term structure of $\mathrm{Log}\left(\frac{ADR}{100-ADR}\right)$ according to the numerical ratings reported in Column (1).[4] We transform the fitted $\mathrm{Log}(\frac{ADR}{100-ADR})$ back into ADR, as reported in Column (6). The fitted ADRs of AAA and AA+ are 0.0073% and 0.0096%, respectively, slightly lower than 0.0133% of AA. The ADRs of CCC+, CCC, CCC- are 9.84%, 14.50%, and 20.33%, respectively, and C has the highest ADR of 34.02%.

Next, we construct the ADR bins of S&P ratings into which we will later allocate firms according to their expected default probabilities. We use the middle point between the ADRs of two adjacent S&P ratings as their ADR bins boundary. Section 3 shows that each S&P ADR bin's real default rate closely resembles its ADR, suggesting that this simple mechanism works.[5] Columns (7) and (8) of Table 3 report the lower and upper ADR boundaries of each S&P rating, respectively. We allocate a firm into the higher rating's bin if its expected default probability happens to fall on the boundary. For example, the firm-month observation with an expected default probability in the range of (0.0085%, 0.0115%] will be assigned with an implied S&P rating of AA+.

2.2 Estimating Expected Default Probability

This section describes how to estimate a firm's expected default probability. It starts with the methodology, followed by defining default events and selecting default predictors.

2.2.1 Methodology

[4] We follow the literature to transform the letter ratings into numerical ratings, that is, AAA equals 1, AA+ equals 2, …, and C equals 21. We assume that CC, the middle grade of the CCC-C bundle, has default rate of 24.34% in the interpolation.

[5] We purposefully use the real default rate to distinguish from the actual default rate (ADR) of S&P global rating. The real default rate is computed as the percentage of firm-month observations in the ADR bin of one S&P rating experienced actual default in 12 months. In this sense, it is essentially the actual default rate of (PD-implied) S&P rating for the Chinese firms. Robustness checks in Section 5 show that more complex bin boundary determination mechanisms, such as simulation and backward induction, yield similar results.

We follow Shumway (2001), Chava and Jarrow (2004) and Campbell, Hilscher and Szilagyi (2008) to estimate firms' one-year expected default probabilities using a dynamic logit model. As noted earlier, this reduced form model accommodates diversified information from firm fundamentals and equity without imposing a structure on information application.

We estimate the logit model using a recursive scheme. That is, we use all the information available at the end of each estimation time. In particular, we assume for each estimation time t, the marginal default probability of firm i in a month s before t $(s < t)$ follows a logistic distribution and is expressed as

$$P\left(Default_{i,s}\right) = \frac{1}{1 + exp\left(-\alpha_t - \beta_t x_{i,s}\right)}, \tag{1}$$

where $P\left(Default_{i,s}\right)$ represents the one-year expected default probability of firm i in month s; $Default_{i,s}$ is a default indicator specified below; $x_{i,s}$ is a vector of default predictors observed in month s. The specification ensures the marginal default probability to be within the (0, 1] range.

For firm i in any month $s < t$, the default indicator is determined as

$$Default_{i,s} = \begin{cases} 1 & if\ D_i \in [s+1, s+12] \text{ and } D_i \le t; \\ Null & if\ D_i \le \text{s}; \\ 0 & otherwise. \end{cases} \tag{2}$$

$Default_{i,s}$ tells whether firm i will default in 12 months at s. D_i denotes firm i's default date. If firm i defaulted before or in month t, the default indicator of month s is equal to one if the default occurs in 12 months. The default indicator is null if firm i has defaulted before or in month s, implying that the defaulted firm has been removed from the sample. Otherwise, the default indicator is equal to zero, indicating that the default date D_i is later than t or time to default is longer than 12 months as of s. For example, Shanghai ChaoRi Solar defaulted on March 7, 2014. The firm's default indicator in any estimation time t before March 2014 is zero. For t of March 2014 and after, the market has learned about the default. The default indicator is equal to one for month s from March 2013 to February 2014, indicating that the firm will default within one year. The default indicator is equal to zero for month s before March 2013 and null after March 2014. This backward-looking approach is important as it determines the value of the default indicator based on default information available at the estimation time t, and relates the information to default predictors observed in a previous month s. It effectively trains the model to identify relevant information to predict defaults.

For each estimation time t, we fit the dynamic logit model using firm-month panel data to obtain the estimated values $\widehat{\alpha_t}$ and $\widehat{\beta_t}$ by maximizing the following likelihood function:

$$\widehat{\alpha_t}, \widehat{\beta_t} = argmax \prod_{i=1}^{N} \prod_{s=0}^{t-1} P(Default_{i,s}), \tag{3}$$

where *N* represents the total number of firms, and *s* ranges from 0 to $t-1$. We set the month of January 1998 to be 0 for data availability reasons discussed in the next section and estimate the model for the period between January 2005 and December 2019. We use the recursive scheme, that is, for each month, we use all the information available since January 1998.

Then we use $\widehat{\alpha_t}$, $\widehat{\beta_t}$, and $x_{i,t}$ to predict firm $i's$ expected default probability in one year at time t as

$$P(Default_{i,t}) = \frac{1}{1+exp(-\widehat{\alpha_t}-\widehat{\beta_t}x_{i,t})}. \tag{4}$$

The prediction begins in 2005 as we have a decent number of observations. The Chinese corporate bond market has begun to grow rapidly since 2005. The State Council of China issued a work report (No. 3 [2004]) to guide the development of corporate bond market in 2004 and ignited the bond market boom. Many bond products, such as commercial papers and asset-backed securities, have been launched since 2005.

2.2.2 Default Events

Our goal is to collect and use the most comprehensive default data in China. This paper follows the literature to investigate the listed firms as they have routinely audited financial reports and stock market information (Shumway, 2001; Chava and Jarrow, 2004; Campbell, Hilscher, and Szilagyi, 2008; Duan, Sun, and Wang, 2012). The listed firms are representative. According to WIND, the annual default rates of all (the listed) corporate bond issuers are 0.14% (0.17%), 0.59% (0.30%), 0.62% (0.00%), 0.20% (0.14%), 0.87% (1.88%) and 0.87% (1.96%) from 2014 to 2019, respectively. Unreported results show that both median ratings of listed and non-listed defaulters are AA six months before default. The median rating is BB for the listed defaulters and A+ for the non-listed defaulters one month before default. The listed firms' credit ratings appear to be more sensitive to default information and up-to-date than the unlisted firms'. Investigating the listed firms' credit ratings may slightly overestimate the effectiveness of domestic ratings.

Another advantage of using the listed firms is being able to consider a broader definition of corporate failure by including Special Treatment (ST) firms. In 1998, China Securities Regulatory

Commission introduced the ST rules to protect investors from firms that are imminent to financially driven delisting or bankruptcy (Zhang, Altman, and Yen, 2010; Altman, Iwanicz-Drozdowska, Laitinen, and Suvas, 2017). A firm is labeled with ST for the following conditions: (1) net profits were negative in the last two fiscal years; (2) net worth of equity was negative in the previous fiscal year; (3) operating income was less than 10 million yuan in the last fiscal year; (4) auditor issued adverse or disclaimer opinion on the financial statement of the previous fiscal year. More than 90% of ST firms have negative earnings for two consecutive years or negative net worth of equity (Cheng, Xia, and Wang, 2014). They were in a de facto default state according to asset value- and cash flow-based credit risk models (Black and Scholes, 1973; Merton,1974; Leland, 1994; Goldstein, Ju, and Leland, 2001).

The practice of including ST firms in the default sample is in line with the literature. Chen and Chen (2000), Zhang, Altman, and Yen (2010), and Altman, Iwanicz-Drozdowska, Laitinen, and Suvas (2017), among others, include the ST firms in their default samples. In broader literature, Beaver (1966), Altaman (1968), Chava and Jarrow (2004) and Campbell, Hilscher, and Szilagyi (2008) consider financially driven delisting, negative cumulative earnings, nonpayment of preferred stock dividends as financial failures besides default and bankruptcy. Including the ST firms expands our sample, extending the sample period by 16 years, as shown in Table 2. A larger sample will help estimate the model more accurately, while ignoring the ST firms could lead to a severe underestimation of default probabilities. We estimate Equation (3) with information starting from January 1998 when the ST rules were introduced. Compared to the developed economies, The Chinese bond market has a relatively short history and fewer observations of corporate defaults. Our data present a comprehensive description of listed corporate defaults and distresses during 1998-2020.

[Insert Table 2 here.]

2.2.3 Default Predictors

We follow existing work to account for a variety of default predictors: firm fundamental variables (Altman, 1968; Ohlson, 1980; Zmijewski, 1984; Shumway, 2001; Bharath and Shumway 2008; Campbell, Hilscher, and Szilagyi 2008); the stock market variables (Campbell, Hilscher, and Szilagyi, 2008; Bharath and Shumway; 2008; Duan, Sun, and Wang, 2012; Brogaard, Li, and Xia, 2017); and distance-to-default (DTD) as in Crosbie and Bohn (2001).

We consider the following fundamental variables: (1) RSize is a firm's relative size estimated using firm market equity value divided by the average market equity value of the Shanghai Stock

Exchange Composite Index's component stocks. (2) MB is the ratio of market equity value to book equity value. (3) WC_MTA is the ratio of working capital to the market asset value computed as the sum of market equity value and total liabilities. According to Campbell, Hilscher and Szilagyi (2008), the market asset value incorporates stock market information relative to the book asset value. (4) RE_MTA is the ratio of retained earnings to the market asset value. (5) EBIT_MTA is the ratio of earnings before interest and taxes (EBIT) to the market asset value. (6) SALE_MTA is the ratio of sales to the market asset value; (7) NI_MTA is the ratio of net income to the market asset value. (8) TL_MTA is the ratio of total liabilities to the market asset value. (9) CASH_MTA is the ratio of cash plus cash equivalents to the market asset value. These variables capture key firm characteristics from various aspects: size, profitability, growth potential, leverage, tangibility, and short-term liquidity. To mitigate the forward-looking bias, we use lagged fundamental variables by one quarter in the estimation.

The stock market variables include: (10) ExRet is the trailing 12-month excess return over the Shanghai Stock Exchange Composite Index return. (11) Dimson β is a proxy for systematic risk (Dimson, 1979). (12) Vol^{Idio} is a proxy for idiosyncratic risk. (13) ILLIQ is a proxy for illiquidity. (14) distance-to-default (DTD) is estimated using the method of Bharath and Shumway (2008). Appendix A presents detailed information on variable construction.

Campbell, Hilscher, and Szilagyi (2008) and Duan, Sun, and Wang (2012) find that comprehensive historical information helps predict default more accurately. Thus, we use the trailing 12-month moving average values of the fundamental variables and DTD in the estimation. We also compute and include the difference between the current value and the moving average value to capture the dynamic change in a default predictor.

[Insert Table 3 here.]

Table 3 reports the summary statistics of the default predictors. There are 462,647 firm-month observations with default indicators of zero and 9,831 observations with default indicators of one. Firms associated with the default indicator of one are less profitable, hold less cash and liquid assets, and display lower stock returns and shorter distance-to-default. They also exhibit higher systematic risk, idiosyncratic risk and stock illiquidity, and more volatile market/book ratios. However, they show a higher average market/book ratio than the firms with zero default indicators. When a firm is imminent to distress, its book value of equity tends to be extremely low or even negative, leading to a mechanically high market/book ratio. On the other hand, a distressed listed firm's market value could be inflated, as many potential buyers are willing to pay a premium to take over the firm as a shortcut to go public. It is not easy to become a listed firm in China due to high qualification requirements and stringent approval processes (Lee, Qu, and Shen, 2019).

Our default predictors do not explicitly include the rating agencies' private information that is difficult to identify, disentangle, and measure. However, to some degree, private information reflected in financial markets can be captured by credit valuation models (Duffie and Singleton, 2003; Saunders and Allen, 2002, Vassalou and Xing, 2004). The rating agencies use private information to adjust ratings both upward and downward, so the private information's effects can be even out or substantially mitigated in cross-section. Consider the agencies' private information is ultimately about assessing default risk, we compare our model-implied default probabilities to actual default rates in Section 3. The result shows that they quantitatively resemble each other, suggesting that the agencies' private information will not substantially affect our conclusions.

3. Mapping the Domestic Agency Ratings to S&P Global Ratings

We establish a mapping between the domestic agency and S&P ratings by allocating the firm-months into the ADR bins of S&P ratings according to their expected default probabilities. The matched sample contains 69,635 firm-month observations from 1,044 firms during September 2005 and December 2019.[6] Table 4 describes the default probabilities for the agency and S&P ratings.

[Insert Table 4 here.]

It shows that the domestic agency ratings are remarkably higher than S&P ratings for the same level of default risk. The mean (median) expected default probabilities of firms with the domestic AAA, AA+, AA, AA-, A+ and Below-A ratings are 1.13% (0.57%), 1.28% (0.79%), 1.61% (1.05%), 2.27% (1.37%), 2.86% (1.61%), 3.86% (1.70%) and 18.83% (1.88%), implying that the domestic AAA corresponds to S&P BB (BB+); AA+ corresponds to S&P BB- (BB); AA corresponds to S&P BB- (BB); AA- corresponds to S&P B+ (BB-); A+ corresponds to S&P B+ (BB-); A corresponds to S&P B (BB-); Below-A corresponds to S&P CCC- (BB-), respectively.[7] The domestic agency ratings are on

[6] For each firm-month, we use the firm's latest and lowest long-term issuer rating obtained from WIND. According to the Chinese regulation (China Securities Regulatory Commission, CSRC No. [2015]113), rating agencies should surveil outstanding credit ratings and perform reviews at least once a year. Thus, we mark the rating missing if the latest rating record is more than one year old to avoid using out-of-date unreliable information.

[7] The domestic below-A ratings are combined into one category due to their small number of observations. The statistics associated with each individual rating are reported for reference only.

average higher than S&P ratings by 11 notches, suggesting that Chinese domestic agency ratings are inflated according to commonly accepted S&P credit rating standards.

One domestic rating's expected default probabilities match the ADRs of a wide range of S&P ratings. For example, the 10-90 percentile expected default probabilities of firms rated AAA match the ADRs of S&P BBB+ to B+ (seven notches); AA+ matches S&P BBB to B+ (six notches); AA matches S&P BBB- to B+ (five notches); AA- matches S&P BBB- to B (six notches); A+ matches S&P BB+ to B (five notches); A matches S&P BBB- to B- (seven notches), respectively. The pattern reveals that firms of significantly different default risks have been clustered into one domestic rating category. The cluster will undermine credit rating's essential role for information discovery (Ramakrishnan and Thakor, 1984; Sangiorgi and Spatt, 2017). It also prevents credit rating from raising default alarm on time, as a distressed firm can have an inflated rating but suddenly defaults. The results explain the recent pickup in surprising defaults of highly-rated firms in China.

The AAA rating has critical regulatory implications in China. For example, only issuers and issues rated AAA are eligible for public issuance (China Securities Regulatory Commission (CSRC) No. [2015]113). Financial institutions' capital reserve requirement for holding AAA bonds is 10% compared to 30% for holding below-AAA bonds (China Banking Regulatory Commission (CBRC) No. [2009]116; CSRC No. [2016]30). Only 0.6% of the observations with domestic AAA ratings are qualified for the S&P AAA rating. 40% of the observations are qualified for S&P investment grades with BBB- as the threshold. Besides, 0.07%, 0.54%, and 0.16% of the observations with domestic AA+, AA, and AA- ratings are qualified for S&P AA+, AA, and AA- ratings, respectively. The evidence confirms that Chinese domestic agency ratings are seriously inflated under the S&P rating standards.

[Insert Figure 1 here.]

The firm-month observations' distributions confirm that the domestic and the PD-implied S&P ratings are subject to very different standards. Graph A of Figure 1 shows that under the PD-implied S&P rating, the observations are distributed in a smooth hump shape with a BB peak. AAA and C contain more observations than their adjacent grades, as these polar ratings contain outlier default probabilities. Nonetheless, the overall pattern is sensible. In contrast, the domestic rating distribution concentrates at the high-investment grade end between AAA and A+, peaked at AA. The lowest C grade has a disproportionately high number of junk grade observations, leaving few observations between BBB- and CC. Amstad and He (2020) and Jiang and Parker (2019) also find that Chinese domestic agency ratings are polarized and abnormally skewed toward the high end. The rating agencies appear to be too lenient in assigning high corporate ratings but conservative in assigning junk ratings.

To address the concern that our default predictors do not explicitly involve the rating agencies' private information, we compute the PD-implied S&P ratings' real default rates and compare them to S&P ratings' ADRs. The real default rate is computed as the percentage of firm-month observations experienced defaults in 12 months for each rating. By examining how the expected default probabilities resemble the real default rates, we assess whether our model sensibly describes default risk.

Graph B of Figure 1 depicts the domestic and PD-implied S&P ratings' real default rates, respectively. The real default rate of each PD-implied S&P rating closely resembles its ADR. Their discrepancy is within one standard deviation of the ADR, suggesting that our model sensibly captures default risk. On the other hand, the domestic agency ratings' real default rates are significantly higher than their S&P counterparts' ADRs. The domestic AAA's real default rate is 0.50%, equal to the ADR of S&P BB+. The real default rates of BBB and BBB- are 17.91% and 57.14%, respectively. It is shocking to observe that the domestic below-BBB ratings' s real default rates soar to 100%. Indeed, it reflects that the domestic rating agencies rarely assign junk ratings until defaults are imminent. Section 4 shows that the defaulted firms' median domestic rating starts to fall slowly from AA eight months before default, reaching BB one month before default.

To assess the overall default predictability of the domestic agency and S&P ratings, we develop the following R-squared measure:

$$R\text{-}squared = 1 - \frac{\sum_{r=1}^{17}\left(S\&P\ ADR_r - Real\ Default\ Rate_r\right)^2}{\sum_{r=1}^{17}\left(S\&P\ ADR_r - \overline{S\&P\ ADR_r}\right)^2}. \tag{5}$$

R-squared is typically used to measure how close data are to a fitted regression line. Our R-squared represents how close the real default rates under different rating schemes are to the ADRs of S&P ratings. In a normal situation, the R-squared takes a value between zero and one. A model is more capable of fitting S&P ratings' ADRs if its R-squared is closer to one. The R-squared of PD-implied S&P ratings is 0.64, while the R-squared of the domestic agency ratings is -46.51. The abnormally big second term in Equation (5) 's right-hand side function gives rise to the negative R-squared, implying that the domestic agencies' rating standards are less stringent than S&P's. Investors should exercise caution to infer default risk from the domestic agency ratings under commonly accepted credit rating standards.

The findings suggest that the rating agencies' private information does not substantially affect our conclusions. As noted earlier, our model can capture the rating agencies' private information embedded in the financial markets. Besides, the rating agencies use private information to adjust ratings both upward and downward, so the private information's effects are mitigated in cross-section.

Graphs A and B of Figure 2 depict the time series of the composition of domestic and PD-implied S&P ratings. In both graphs, the vertical axis represents the cumulative percentages of ratings, with AAA at the bottom and C on the top. To facilitate illustration, we combine ratings by letter category. For example, AA+, AA, and AA- are combined into AA; CCC, CC, and C are combined into CCC/C.

[Insert Figure 2 here]

Graph A shows that the composition of the domestic agency ratings has changed dramatically over time. In 2007, there were only three rating categories: AAA (10.00%), AA (47.50%), and A (42.50%). The percentages of AAA and AA firms continued to increase as time passed. In 2019, 25.69% of firms were rated AAA, and 72.03% were rated AA. In sharp contrast, the percentage of firms rated A has fallen to 0.98%. We barely observe low-investment grades (BBB) and junk grades (BB-C) until 2017; however, these firms accounted for only about 1% of the population in 2019. As noted earlier, the issuers and investors of high-investment grade issues enjoy tremendous regulatory advantages. Heavy regulatory reliance on credit ratings promotes domestic rating inflation (Liu and Wang, 2020).

Graph B shows that the composition of the PD-implied S&P ratings is relatively stable over time. In 2007, the percentages of AAA, AA, A, BBB, BB, B and CCC/C ratings were 0%, 0%, 0.56%, 11.43%, 58.41%, 29.60% and 0%, respectively.[8] In 2019, their percentages were 0.95%, 1.13%, 4.09%, 20.33%, 54.07%, 15.97% and 3.47% respectively. The 2019 composition appears to be more reasonable, as the percentages of AAA, AA, and CCC/C ratings are no longer zero. The rating composition remains relatively stable after accounting for seasonal and economic cycles, e.g., a higher percentage of investment-grade ratings in 2009. Firms' financial performance was improved due to the expansionary monetary policy and the "four-trillion-yuan" stimulus package in combating the Global Financial Crisis (Chen, He, Liu, 2019; Wang, Wang, Wang, Zhou, 2020).

4. PD-Implied Ratings

[8] The reported percentages are the averages of monthly percentages in a year.

This section studies the PD-implied ratings constructed in Section 3. It examines the ratings' default predictability and explanatory power of corporate bond yield spread. The last part sheds light on why the agency ratings and PD-implied ratings perform differently in predicting default.

4.1 Predicting Default

We study the default predictability of PD-implied ratings from two angles. First, we follow Vassolu and Xing (2004), Duan, Sun, and Wang (2012), and S&P (2020) to apply the Receiver Operating Characteristic (ROC) curve and accuracy ratio (AR) to assess their default differentiability (Appendix B provides an introduction of ROC curve and AR). Second, we depict the median PD-implied ratings and domestic agency ratings of defaulted firms 24 months before default to illustrate their default predictability.

[Insert Figure 3 here.]

The ROC curve is a graphical plot illustrating the diagnostic ability of a binary classification model (in this case, a rating model predicts default or no-default) at various discrimination thresholds (in this case, at different ratings).[9] In particular, the ROC curve establishes a comparison of the True Positive Rate (TPR) to the False Positive Rate (FPR) for every rating (as a discrimination threshold) on a TPR-FPR coordinate. For example, let C be the default threshold. The TPR is computed as the ratio of the number of firms rated C defaulted to the total number of firms defaulted in 12 months. The FPR is the ratio of the number of firms rated C that did not default to the total number of firms that did not default in 12 months.

Graph A of Figure 3 depicts the ROC curves of the PD-implied and domestic agency ratings. Both are benchmarked to a 45-degree linear line representing the ROC curve of a hypothetical random model with no power of predicting default. The ROC curves of the PD-implied and domestic agency ratings are above the random model's ROC curve, suggesting that both models possess some ability to detect default. The PD-implied ratings' ROC curve is significantly above the domestic agency ratings' ROC curve, implying that the PD-implied ratings have superior default differentiability.

Accuracy Ratio provides a numerical measure of the discriminatory power of a binary classification model. In this case, the Accuracy Ratio is the ratio of the space between a rating model's ROC curve and the ROC curve of the hypothetical random model (the 45-degree line) to the space

[9] The U.S. military first used the ROC curve to detect radio signals from noises, after the Japanese attack on Pearl Harbor. The initial research was motivated to determine to what extent the U.S. radar receiver operators had missed detecting the Japanese aircraft, after which ROC is named. The ROC curve was then introduced to psychology to account for perceptual detection of stimuli and has been widely used in medicine, radiology, and machine learning.

between the ROC curves of a perfect theoretical model and the hypothetical random model, which is the total space above the 45-degree line. A model can more accurately predict default if its Accuracy Ratio is closer to one. We find that the Accuracy Ratio of the PD-implied ratings is 63.45%, approximately two times the Accuracy Ratio of domestic agency ratings, 32.23%, confirming that the PD-implied ratings outperform the conventional ratings in predicting default.

Graph B depicts the median PD-implied and domestic agency ratings of 36 defaulters in a 24-month window before default. The domestic agency ratings' dash line shows that these defaulters' median ratings remain AA until eight months before default. The median domestic rating falls to AA- six months and then further to BB one month before default. This pattern echoes cliff-falling downgrades observed in the market. In contrast, the solid line of PD-implied ratings indicates that the defaulters are risky with non-investment grades in the first place. The median implied rating starts to fall steadily from BB 12 months before default, reaching CCC one month before default. The PD-implied ratings give stronger alarming signals and avoid imminent-to-default cliff downgrading.

The rating migration paths of the default and ST firms tell a consistent story. The domestic rating's dotted line shows that the median rating remains AA until seven months before distress. It falls to AA- six months and further to A+ one month before distress. In contrast, the dash-dotted line of the PD-implied rating shows that the median rating starts to fall steadily from BB- 18 months before distress, reaching CCC one month before distress. Deterioration in the PD-implied ratings accelerates about 11 months before distress, giving earlier and more precise warnings.

On March 4, 2014, Shanghai ChaoRi Solar defaulted on the interest payments on its one-billion-yuan bond. It was the first case of publicly issued bond default in China. We illustrate how the domestic and PD-implied ratings had evolved before its insolvency in Graph C. The domestic rating's dashed line starts to fall from AA to AA- in December 2012. The rating drops dramatically from AA- to CCC within three months after the firm being labeled as ST in February 2013. The solid line of the PD-implied rating falls from BB in December 2011 to C in April 2013. The persistent downgrading would have generated alarming signals to creditors. Although the implied ratings use only public information, they accommodate diverse information and have superior default predictivity.

4.2 Explaining Yield Spread

This section investigates whether the PD-implied rating has incremental explanatory power of corporate bond yield spread in the presence of conventional rating and primary credit risk factors.

There is ample evidence that Chinese domestic agency ratings significantly explain corporate yield spreads (Livingston, Poon, and Zhou, 2018; Liu and Wang, 2020). Our corporate bond transaction data are from RESSET, a well-known and reliable database, covering a period from January 2005 to December 2019. For each month, we use a bond's latest price to compute its yield to maturity (YTM) and then subtract the Treasury yield of matched maturity (interpolated with Cubic Splines) to obtain yield spread. If a firm has multiple bonds traded in a month, we select the bond with the largest issuance amount. Issues of larger sizes are generally more liquid, so their prices can better reveal intrinsic values.

The regression is formulated as

$$Spread_{i,t+1} = \theta_1 * Implied\ Rating_{i,t} + \sum_{i=2}^{N} \theta_i * Control_{i,t} + \varepsilon_{i,t}, \tag{6}$$

where $Spread_{i,t+1}$ denotes yield spread of firm i at time $t+1$; $Implied\ Rating_{i,t}$ denotes PD-implied rating and $Control_{i,t}$ denotes control variables observed at time t, respectively. We further control for the firm and year-month fixed effects. Ratings may reflect changes in aggregate fundamentals so that standard errors may be correlated across firms. Thus, we compute standard errors clustered by firm.

We follow Collin-Dufrene, Goldstein, and Martin (2001) and Chen, Lesmond, and Wei (2007) to account for a variety of bond characteristics: Log (Issue Size) is the logarithm of bond issue amount in a million yuan; Log (Time to Mat) is the logarithm of time to maturity; Coupon denotes coupon rate; Log (TrdVol) is the logarithm of total trading volume in the month in a million yuan. Firm-level control variables include Stock Vol, which is annualized 12-month stock return volatility; Leverage is the ratio of total liabilities to total assets. According to Black and Sholes (1973) and Merton (1974), these variables are primary credit risk determinants. We do not involve other financial variables to avoid the multicollinearity problem because the PD-implied rating is constructed using financial information. Our final sample contains 21,531 firm-month observations involving 559 different firms.

[Insert Table 5 here.]

Panel A of Table 5 reports that the average (median) yield spread is 251 (213) basis points (bps) with a standard deviation of 189 bps, suggesting that the sample is diversified. Approximately 1% of the observations have negative yield spreads. They mainly come from large state-owned firms that finance at low costs. We follow the literature to transform the letter ratings into numerical ratings, that is, AAA equals 1, AA+ equals 2, …, and C equals 21. The average domestic numerical rating is 2.62, corresponding to AA. The domestic agency ratings, however, cluster around five notches between AAA and A+.

In contrast, the average PD-implied numerical rating is 12.03, corresponding to BB. The PD-implied ratings distribute in a wide range between A+ and CCC. The average (median) coupon rate is 5.93% (5.88%) with a standard deviation of 1.16%. The average book leverage and stock return volatility are 57% and 40%, respectively. The statistics of the variables are similar to those reported in previous work.

Panel B reports that the PD-implied and domestic agency ratings have a correlation coefficient of 0.31, suggesting that the PD-implied rating shares some information with the conventional rating yet possesses some unique knowledge. The correlation coefficient of yield spread and the PD-implied rating (domestic rating) is 0.22 (0.45). The PD-implied rating (domestic rating) and leverage and stock return volatility have correlation coefficients of 0.26 (0.00) and 0.19 (0.23), respectively. The correlation coefficients of PD-implied rating (domestic rating) and Log (Issue Size) and Coupon are -0.13 (-0.58) and 0.18 (0.56), respectively. The PD-implied rating appears to capture more fundamental information, while the conventional rating reflects more bond characteristics.

[Insert Table 6 here.]

Table 6 reports the regression results. Column (1) shows that yield spread is positively correlated with the PD-implied rating, statistically significant at the 1% level. The coefficient estimate is 13.88, implying that a one-notch decrease in the PD-implied rating corresponds to an average increase in yield spread by 13.88 bps. The PD-implied rating's coefficient estimate is 11.52 after controlling for the domestic rating. The coefficient estimate is lower by 2.36 bps but remains economically significant. The coefficient estimate of the domestic rating is 45.13, statistically significant at the 1% level, consistent with the findings of Dhawan and Yu (2015) and Livingston, Poon, and Zhou (2018). The adjusted R-squareds' reported in Columns (1) and (2) are 0.54 and 0.55, respectively, implying that adding the conventional rating slightly increases the model's explanatory power.

The PD-implied rating's explanatory power is further confirmed after including bond- and firm-level control variables in the regression. Column (4) reports that the PD-implied rating's coefficient estimate is 9.53, statistically significant at the 1% level. The adjusted R-squared's of Columns (3) and (4) are 0.56 and 0.57, confirming that the PD-implied rating complements the conventional rating and other primary credit risk determinants in explaining yield spread in the subsequent month.

The average values (standard deviations) of credit spreads for firms with domestic AAA, AA+, AA, and AA- ratings are 127 bps (100 bps), 194 bps (142 bps), 278 bps (185 bps), and 385 bps (190 bps), respectively. On the one hand, the credit prices appear to reflect issuers' credit quality. The result is not surprising as the Chinese bond markets are less efficient than the developed markets, and bond pricing heavily relies on explicit ratings. On the other hand, the high standard deviations imply substantial overlapping between the credit spreads of issuers with adjacent ratings, suggesting that the ratings may not be accurate.

Therefore, we examine the explanatory power of the PD-implied rating within each of the domestic agency ratings and report the results in Columns (5) to (8). The PD-implied rating's coefficient estimates are 6.82, 7.94, 14.02, and 18.06, respectively, all statistically significant. The significant explanatory power persists for bonds of varying credit quality. Intuitively, the coefficient estimate increases monotonically as the bond issuer's agency rating deteriorates. The yield spreads of lower-grade bonds decrease more in magnitude for a one-notch increase in the PD-implied rating. The PD-implied rating can further differentiate issuers' credit quality from the same domestic rating.

In summary, the PD-implied rating can better differentiate default. The accuracy ratio of the PD-implied rating is twice that of the domestic rating. The PD-implied rating provides earlier and more precise distress warnings than the traditional rating and complements the latter in explaining yield spread for different credit quality bonds.

4.3 Why Do Agency Ratings and PD-Implied Ratings Perform Differently?

One may be curious about where the PD-implied ratings' promiment default predicting power originates and what information they capture while the agency ratings do not. To address these questions, we carry out reverse engineering to regress the agency ratings on our default predictors with a monthly sample from January 2007 to December 2019.

We consider three approaches: (1) the Ordered Probit model following Kaplan and Urwitz (1979), Blume, Lim, and Mackinlay (1998) and Alp (2013); (2) the OLS model with firm and year-month fixed effects to control the influence of potential latent factors beyond our predictors;[10] and (3) the Fama and Macbeth (1973) method to generate cross-sectional estimations each month and then average the coefficients, considering that the rating agencies might alter their models over time. We report the

[10] See Baghai, Servaes, and Tamayo (2014) for a discussion about that controlling for firm fixed effects in the Ordered Probit Model could result in biased and inconsistent estimates.

results in Columns (1) to (3) of Table 7.

[Insert Table 7 Here.]

For comparison, we report in Column (4) the predictors deemed important in predicting default in our model. In the spirit of Section 2, the default probability is estimated using penalized logistic regressions (Elastic Net) under a monthly rolling scheme.[11] We know how many times a predictor is selected and used to predict default during the 156 sample months. The more times a predictor is selected, a more important role it has in predicting default. Comparing the important default predictors for the two rating schemes discover the reasons behind the PD-implied ratings' good performance and identify potential routes to improve the agency ratings.

We classify the predictors into four groups: Group 1 contains those significant/important to both types of ratings; Group 2 reports the predictors important to the PD-implied ratings only; Group 3 includes the predictors significant to the agency ratings only; Group 4 involves the predictors that are insignificant or inconsistently significant to both types of ratings. Our analysis focuses on the predictors in Groups 2 and 3.

In Group 2, the predictors most frequently selected by our default prediction model but insignificant to the agency ratings are the ratio of net income to total assets and its change, the ratio of cash to total assets, and excess stock return. The agency ratings tend to overlook dynamic change in profitability, liquidity, and stock return that strongly predict defaults. The Group 3 results indicate that the domestic agencies pay greater attention to MB ratio, relative size, leverage ratio, and idiosyncratic volatility, which are not equally valuable for default prediction in our model. Size and capital structure are static, and idiosyncratic risk is time-varying. The agency ratings are negatively correlated with MB ratios, which seems counterintuitive. This result is, however, consistent with those reported in Table 3. As explained earlier, a distressed Chinese listed firm usually has an extremely low or even negative book value of equity. Yet, its market value tends to be inflated, as many potential buyers are willing to pay a premium to take over the firm as a shortcut to go public.

Overall, the PD-implied ratings use more dynamic information, such as (change in) profitability and cash holding, to predict default, while the domestic agency ratings rely more on static

11 Gu, Kelly, and Xiu (2020) show that Elastic Net is the general case of commonly used penalized regressions, such as Lasso or Ridge regressions. According to Avramov, Li, and Wang (2021), as a generalized method, Elastic Net not only allows shrinkage (Ridge regressions) that helps mitigate multicollinearity among predictors but also imposes variable selection (Lasso regressions) that is well suited for assessing the relative importance of predictors in generating our PD implied ratings.

characteristics, such as size and leverage ratio, in credit risk assessment. Stock information is also processed differently---the PD-implied rating models focus on returns, but the agency ratings emphasize return volatility. Their uses of different information with perhaps different focuses lead to default signals of different accuracy.

5. Robustness Check

This section implements a battery of robustness checks. First, we examine whether the mechanism we have applied to estimate S&P ratings' ADR bin boundaries affects our findings. Second, we investigate whether different default probability estimation models generate consistent empirical results.

Previously, we arbitrarily use the middle point between the ADRs of two adjacent S&P ratings as their ADR bin boundaries. Although it has shown that these bins' real default rates are reasonably close to their ADRs, we apply two alternative schemes to determine the optimal ADR bin boundaries by maximizing the R-Squared defined in Equation (5). The first scheme is backward induction, in which we search the optimal ADR bin boundaries grade by grade, starting with the lowest rating C. We restrict the searched value within the range of the two adjacent ratings' ADRs. For example, the optimal ADR bin boundary between CC and C is within the scope of CC's ADR of 27.0179% and C's ADR of 34.0191%. The backward induction scheme's R-Squared is 0.808, higher than 0.641 of the “middle point” approach. The result implies that our simple approach is reasonable, though not optimal. The second estimation scheme is a simulation by which we generate 100,000 times a random series of ADR bin boundaries and keep track of their R-Squareds'. The series of ADR bin boundaries that yields the highest R-Squared is considered optimal. The baseline approach’s R-squared of 0.641 is slightly higher than the median R-squared of the 100,000 simulations, confirming that the simple “middle point” approach is not the best but sensible.

We also consider the Merton (1974) model and the other 11 machine learning models to estimate expected default probabilities.[12] The untabulated results confirm that the Chinese domestic agency ratings are significantly inflated by the commonly accepted rating standard. One domestic rating’s

---

[12] We consider the following models: cost-sensitive logistic model (being cost-sensitive means that balanced weights are assigned to the default and non-default samples), penalized logistic model (Elastic Net), random forest model, cost-sensitive random forest models, XGBoost, cost-sensitive XGBoost, Neural Network with hidden layers from one to five. Guo et al., (2017) provide an excellent review of the methods and applications of machine learning models on data with class imbalance (e.g., a sample with rare default events, like ours). Gu, Kelly and Xiu (2020) present a detailed description of the empirical application of machine learning models in asset pricing.

expected default probabilities correspond to the actual default rates of a wide range of S&P ratings, revealing that the domestic agency ratings are not accurate in describing credit risk. A more significant percentage of firms have received lower PD-implied S&P ratings in the recent years, adding evidence to rating standard relaxation (Liu and Wang, 2020).

Figure 4 reports the actual default rates and R-Squareds of these alternative models. Some machine learning models marginally outperform the dynamic logit model. Among them, the XGBoost model and the Neural Network model have R-Squareds' of 0.97 and 0.94. The cost-sensitive models that assign balanced sample weights to the default and non-default samples underperform their non-cost-sensitive counterparts. The R-Squared of the Neural Network model monotonically decreases as the number of hidden layers exceeds two. The Random Forest models estimate integer default probabilities and cannot generate implied ratings of AA+ to BB- as these ratings' default probabilities are below one. Thus, we do not report their R-Squareds.

[Insert Figure 4 here.]

Most machine learning models' R-Squareds are significantly higher than that of the Merton (1974) model, suggesting that the machine learning models have superior default predictability. Our findings support the notion that reduced-form models are useful in credit risk management (Bharath and Shumway, 2008; Campbell, Hilscher, and Szilagyi, 2008; Duan, Sun, and Wang, 2012). As noted earlier, the machine learning models accommodate dynamic and diversified information from multiple sources and are flexible in terms of not imposing a structure on information application (Gu, Kelly, and Xiu, 2020).

6. Explaining Rating Inflation in China

This section discusses why the domestic agency ratings are significantly higher than S&P ratings for the same level of default risk. Previous research has identified the following potential explanatory factors: implicit guarantee, pro-cyclical pattern of credit ratings, rating industry competition, interest conflict under the issuer-pays business model, and regulatory reliance on credit ratings.

Implicit guarantee for corporate issuers with links to the government may partly justify higher ratings (Amstad and He, 2020). According to WIND, about 80% of the corporate debenture issuers are state-owned enterprises as of 2019. Walker, Zhang, and Zhang (2019) show that the yield spreads of quasi-municipal corporate bonds (“Chengtou” bonds) that carry an implicit government guarantee are significantly lower than those of corporate bonds issued by privately-owned enterprises. Provincial fiscal capacity and political risk are important determinants of "Chengtou" bond yield spreads (Liu, Lyu, and Yu, 2017; Ang, Bai, and Zhou, 2019). Although recognized by the market as a kind of soft credit enhancement, implicit guarantees may give rise to inflated ratings because they are not legally-bound commitments.

Bar-Issac and Shapiro (2013) and Dilly and Mahlmann (2016) show that upward biased ratings are more likely to occur during economic booms. Entangling with implicit guarantees, booming economy or expansionary monetary policy during financial crisis may play a role in rating inflation (Chen, He, and Liu, 2019 Chen, Chen, He, Liu, and Xie, 2020; Cong, Gao, Ponticelli, and Yang, 2019; Ding, Xiong, and Zhang, 2020).

Rating industry competition is related to inflated ratings (Becker and Milbourn, 2011). Jiang and Packer (2019) document that the Chinese agency ratings are on average 6-7 notches higher than the global ratings from Moody's, S&P, or Fitch, according to comparable rank ordering. They find that industry competition helps explain the gap between the domestic agency and global ratings. Liu and Wang (2020) find the domestic rating standard relaxation in 2006-2019 is partially explained by rising competition in the rating industry.

Fee-paying mechanism also causes inflated ratings (Jiang, Stanford, and Xie, 2012), especially when tangled with regulatory reliance on ratings (Opp, Opp, and Harris, 2013; Behr, Kisgen, and Taillard, 2018). All Chinese domestic rating agencies collect fees from issuers except one. Hu, Huang, Pan, and Shi (2019) find that the issuer-paid rating agencies lowered their ratings and increased rating informativeness after observing investor-paid rating agency coverage. Liu and Wang (2020) show that issuers tend to receive better ratings if they have longer relationships with an issuer-paid agency.

Credit ratings are extensively used for the regulation of bond issuance and risk control in China. Liu and Wang (2020) examine firms' upgrading patterns before and after new regulations enacted in 2015 to give issuers of AAA rating more significant regulatory advantages. They find that the upgrading rate of AA+ to AAA nearly doubled from 5.5% in 2014 to 10.9% in 2015. A *difference-in-difference* study tells that the post-regulation upgrades are explained by relaxed rating standards rather than improved firm fundamentals. Laxer rating standards have resulted in an average higher rating by 0.7 notches. In comparison, a one standard deviation increase in competition and the issuer-agency relationship (a proxy for conflict of interest) led to higher ratings by 0.02 and 0.18 notches on average.

7. Conclusion

This paper establishes the first mapping between the Chinese domestic and S&P global ratings by matching firms' expected default probabilities with S&P ratings' actual default rates. At least half of the Chinese domestic agency ratings are higher than S&P ratings by ten notches for the same level of default risk. The results constitute the first default-based evidence of the Chinese agency ratings being inflated in the light of a widely accepted rating standard.

The PD-implied ratings can better differentiate defaults and provide earlier and more precise default warnings than the conventional ratings. The superior default predictive power comes from their employment of more dynamic information, such as changes in firm profitability and stock returns. In contrast, the agency ratings rely more on static firm characteristics, such as size and leverage ratio. The findings advocate default-probability-based models that have great potential to improve current rating practices, especially in an environment where agency ratings are more likely to be compromised.

Our findings render several policy implications. First, adopting more rigorous rating standards is essential to facilitate the healthy development of credit markets in China. Second, making the domestic rating agencies more responsible for erroneous and inflated credit ratings helps reinforce rating effectiveness and restore market confidence. Third, default probability-based rating models involving artificial intelligence, machine learning, and big data could innovatively complement the traditional ratings in credit risk assessment and supervision.

Reference

1) Alp, Aysun, 2013, Structural shifts in credit rating standards, *The Journal of Finance* 68, 2435-2470.
2) Altman, Edward, 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance*, 23, 589-609.
3) Altman, Edward I., Iwanicz-Drozdowska, Małgorzata., Laitinen, Erkki K., Suvas, Arto. 2017. Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. Journal of International Financial Management & Accounting, 28: 131-171.
4) Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets*, 5, 31-56.
5) Amstad, Marlene, and Zhiguo He, 2020, Chinese Bond Market and Interbank Market, in Amstad, Marlene, Guofeng Sun, and Wei Xiong, eds. The Handbook of China's Financial System. Princeton University Press, 2020.
6) Ang, Andrew, Jennie Bai, and Hao Zhou, 2018, The great wall of debt: real estate, political risk, and Chinese local government financing cost. Working Paper.
7) Avramov, Doron, Minwen Li, and Hao Wang, 2021, Predicting corporate policies using downside risk: A machine learning approach. *Journal of Empirical Finance*, in press.
8) Baghai, Ramin P., Henri Servaes, and Ane Tamayo. 2014. Have rating agencies become more conservative? Implications for capital structure and debt pricing. *The Journal of Finance*, Vol (69), No.5: 1961-2005.
9) Bar-Isaac, Heski, and Joel Shapiro, 2013, Ratings quality over the business cycle, *Journal of Financial Economics*, 108(1): 62-78.
10) Beaver, William H, 1966, Financial ratios as predictors of failure, *Journal of Accounting Research*, 71-111.
11) Becker, Bo, and Todd Milbourn, 2011, How did increased competition affect credit ratings?, *Journal of Financial Economics*, 101, 493-514.
12) Behr, Patrick, Darren J. Kisgen, and Jérôme P. Taillard, 2018, Did government regulations lead to inflated credit ratings? *Management Science*, 64(3): 1034-1054.
13) Bharath, Sreedhar T, and Tyler Shumway, 2008, Forecasting default with the Merton distance to default model, *The Review of Financial Studies*, 21, 1339-1369.
14) Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-654.
15) Blume, Marshall E., Felix Lim, and A. Craig MacKinlay. 1998. The declining credit quality of US corporate debt: Myth or reality? *The Journal of Finance*, Vol(53), No.4: 1389-1413.
16) Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2012, The credit ratings game, *The Journal of Finance*, 67, 85-111.
17) Brogaard, Jonathan, Dan Li, and Ying Xia, 2017, Stock liquidity and default risk, *Journal of Financial Economics*, 124, 486-502.
18) Campbell, John Y, Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *The Journal of Finance*, 63, 2899-2939.
19) Chava, Sudheer, and Robert A Jarrow, 2004, Bankruptcy prediction with industry effects, *Review of Finance*, 8, 537-569.
20) Chen, Hui, Zhuo Chen, Zhiguo He, Jinyu Liu, and Rengming Xie, 2019, Pledgeability and asset

prices: Evidence from the Chinese corporate bond markets. NBER Working Paper.
21) Chen, Long, David A Lesmond, and Jason Wei, 2007, Corporate yield spreads and bond liquidity, *The Journal of Finance* 62, 119-149.
22) Chen, Tianqi, and Carlos Guestrin, 2016, Xgboost: A scalable tree boosting system, in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785-794.
23) Chen, Zhuo, Zhiguo He, and Chun Liu, 2020, The financing of local government in China: Stimulus loan wanes and shadow banking waxes, *Journal of Financial Economics* 137, 42-71.
24) Chen, Xiao, and Zhihong Chen, 2000, Predicting Financial Distress in Chinese Listed Firms, *China Accounting and Finance Review*, Vol (2), No.3, 55-92.
25) Cheng, Enoch, Fan Xia, and George Yungchih Wang, 2014, The effect of special treatment designation on information transmission in the Chinese stock market, *International Journal of Economics and Statistics* 1, 294-305.
26) Cong, Lin William, Haoyu Gao, Jacopo Ponticelli, and Xiaoguang Yang, 2019, Credit allocation under economic stimulus: evidence from China, *Review of Financial Studies* 32 (9), 3412–3460.
27) Collin-Dufresne, Pierre, Robert S Goldstein, and J Spencer Martin, 2001, The determinants of credit spread changes, *The Journal of Finance* 56, 2177-2207.
28) Crosbie, PJ, and JR Bohn, 2001, Modeling portfolio risk, Technical report, Working Paper, Moody's KMV.
29) Dhawan Raghav and Fan Yu, 2015, Are Credit Ratings Relevant in China’s Corporate Bond Market?, *The Chinese Economy*, 48:3, 235-250.
30) Dilly, Mark, and Thomas Mahlmann, 2016, Is there a “boom bias” in agency ratings?, Review of Finance 20, 979-1011.
31) Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197-226.
32) Ding, Yi, Wei Xiong, and Jinfan Zhang, 2020, Overpricing in china's corporate bond market, NBER Working Paper.
33) Duan, Jin-Chuan, and Shuping Li, 2020, PD-implied ratings via referencing a credit rating/scoring pool's default experience, *Data Science for Financial Econometrics (in press)*.
34) Duan, Jin-Chuan, Jie Sun, and Tao Wang, 2012, Multiperiod corporate default prediction—a forward intensity approach, *Journal of Econometrics,* 170, 191-209.
35) Duffie, Darrell, and Kenneth J. Singleton, 2003, Credit Risk: Pricing, Measurement, and Management, *Princeton University Press*, Princeton, NJ.
36) Fama, Eugene F., and James D. Macbeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy*, Vol (81), No.3: 607-636.
37) Flannery, Mark J, Joel F Houston, and Frank Partnoy, 2010, Credit default swap spreads as viable substitutes for credit ratings, *University of Pennsylvania Law Review*, 2085-2123.
38) Geng, Zhe, and Jun Pan, 2019, Price discovery and market segmentation in china's credit market. NBER Working Paper.
39) Goldstein, Robert, Nengjiu Ju, and Hayne Leland, 2001, An EBIT-based model of dynamic capital structure, *The Journal of Business* 74, 483-512.
40) Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *The Review of Financial Studies* 33, 2223-2273.
41) Guo, Haixiang, Yijing Li, Shang Jennifer, Mingyun Gu, Yuanyue Huang, and Bing Gong, 2017,

Learning from class-imbalanced data: Review of methods and applications, *Expert Systems with Applications* 73, 220-239.
42) Hillegeist, Stephen A, Elizabeth K Keating, Donald P Cram, and Kyle G Lundstedt, 2004, Assessing the probability of bankruptcy, *Review of Accounting Studies* 9, 5-34.
43) Hu, Xiaolu, Haozhi Huang, Zheyao Pan, and Jing Shi, 2019, Information asymmetry and credit rating: A quasi-natural experiment from China, *Journal of Banking & Finance*, 106, 132-152.
44) Jiang, John Xuefeng, Mary Harris Stanford, and Yuan Xie, 2012, Does it matter who pays for bond ratings? Historical evidence, *Journal of Financial Economics* 105(3): 607-621.
45) Jiang, Xianfeng, and Frank Packer, 2019, Credit ratings of Chinese firms by domestic and global agencies: Assessing the determinants and impact, *Journal of Banking & Finance*, 105, 178-193.
46) Kaplan, Robert S., and Gabriel Urwitz. 1979. Statistical models of bond ratings: A methodological inquiry. *Journal of business*, 231-261.
47) Lee, Charles MC, Yuanyu Qu, and Tao Shen, 2019, Going public in china: Reverse mergers versus IPOs, *Journal of Corporate Finance* 58, 92-111.
48) Leland, Hayne E, 1994, Corporate debt value, bond covenants, and optimal capital structure, *The Journal of Finance* 49, 1213-1252.
49) Liu, Laura Xiaolei, Yuanzhen Lyu, and Fan Yu, 2017, Implicit government guarantee and the pricing of Chinese LGFV debt, Working Paper.
50) Liu, Shida, and Hao Wang, 2020, The real effects of rating inflation: Evidence from the Chinese corporate credit ratings. SSRN Working Paper.
51) Livingston, Miles, Winnie PH Poon, and Lei Zhou, 2018, Are Chinese credit ratings relevant? A study of the Chinese bond market and credit rating industry, *Journal of Banking & Finance* 87, 216-232.
52) Merton, Robert C, 1974, On the pricing of corporate debt: The risk structure of interest rates, *The Journal of Finance* 29, 449-470.
53) Millon, Marcia H, and Anjan V Thakor, 1985, Moral hazard and information sharing: A model of financial information gathering agencies, *The Journal of Finance* 40, 1403-1422.
54) Ohlson, James A, 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 109-131.
55) Opp, Christian C., Marcus M. Opp, and Milton Harris, 2013, Rating agencies in the face of regulation, *Journal of Financial Economics*, 108(1): 46-61.
56) Poon, Winnie PH, Iftekhar Hasan, Gaiyan Zhang, and Jianfu Shen, 2016, Are market implied ratings viable alternatives to credit ratings? SSRN Working Paper.
57) Saunders, Anthony and Linda Allen, 2002, Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms, *John Wiley and Sons*, New York.
58) Schipke, Alfred, Markus Rodlauer, and Longmei Zhang, 2019, *The Future of China's Bond Market* (International Monetary Fund).
59) Ramakrishnan, Ram TS, and Anjan V Thakor, 1984, Information reliability and a theory of financial intermediation, *The Review of Economic Studies* 51, 415-432.
60) S&P Global Ratings, 2020, Default, Transition, and Recovery: 2019 Annual Global Corporate Default and Rating Transition Study.
61) Shumway, Tyler, 2001, Forecasting bankruptcy more accurately: A simple hazard model, *The Journal of Business,* 74, 101-124.
62) Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *The Journal of Finance*,

59, 831-868.

63) Wang, Hao, Honglin Wang, Lisheng Wang, and Hao Zhou, 2019, Shadow banking: China's dual track interest rate liberalization, *Working Paper*.
64) Walker, Thomas John and Zhang, Xueying and Zhang, Aoran, 2019, The Great Wall and Beyond: The Value of Implicit Government Guarantees for Corporate Bonds in Mainland China and its Northeast Regions, *Working Paper.*
65) Zhang, Ling, Edward I Altman, and Jerome Yen, 2010, Corporate financial distress diagnosis model and application in credit rating for listing firms in China, *Frontiers of Computer Science in China* 4, 220-236.
66) Zmijewski, Mark E, 1984, Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research,* 59-82.

Appendix A. Definitions and Construction of Default Predictors

Variable	Definition and Construction
RSize	The firm's relative size, constructed as the logarithm of market equity value divided by the Shanghai Stock Exchange Composite Index component stocks' average market value.
MB Ratio	The ratio of market equity value to book equity value. The market equity value is equal to the common share price times number of common shares outstanding.
WC_MTA	The ratio of working capital to the market value of total assets. Working capital is equal to the difference between current assets and current liabilities. The market value of total assets is the sum of market equity value and total liabilities.
RE_MTA	The ratio of retained earnings to the market value of total assets. Retained earnings is equal to the sum of earned surplus and undistributed profit.
EBIT_MTA	The ratio of earnings before interest and taxes (EBIT) to the market value of total assets.
SALE_MTA	The ratio of sales to the market value of total assets.
NI_MTA	The ratio of net income to the market value of total assets.
TL_MTA	The ratio of total liabilities to the market value of total assets.
CASH_MTA	The ratio of cash plus cash equivalents to the market value of total assets.
ExRet	Trailing 12-month excess return over the Shanghai Stock Exchange Composite Index return.
ILLIQ	Stock illiquidity, constructed using the method of Amihud (2002): ILLIQ=$(\sum \lvert r_{i,t}\rvert / TrdVol_{i,t})/N$, where $r_{i,t}$ denotes daily stock return; $TrdVol_{i,t}$ is daily trading volume, N is the number of trading days in the past one year with a minimum of 50 trading days required.
β	Dimson beta as a proxy of systematic risk. As in Dimson (1978), the regression model is formulated as $r_{i,t} = \alpha_{i,t} + \beta_1 r_{m,t-2} + \beta_2 r_{m,t-1} + \beta_3 r_{m,t} + +\beta_4 r_{m,t+1} + \beta_5 r_{m,t+2} + \varepsilon_{i,t}$, and then calculate $\beta = \sum \beta_i$. The regression utilizes the past one-year daily stock returns and Shanghai Stock Exchange Composite Index returns as a proxy for the market returns. A minimum of 50 daily observations is required.
$\mathrm{Vol}^{\mathrm{Idio}}$	The standard deviation of the regression residual in the Dimson model as a proxy for idiosyncratic risk.
DTD	According to the Merton (1974) model, $V_E = V_A N(d_1) - e^{-rT} D N(d_2)$; $\sigma_E = \frac{V_A}{V_E} N(d_1)\sigma_A$, where $d_1 = \frac{\ln\left(\frac{V_A}{D}\right) + (r + \frac{\sigma_A{}^2}{2})T}{\sigma_A\sqrt{T}}$, $d_2 = d_1 - \sigma_A\sqrt{T}$; $DTD = \frac{\ln\left(\frac{V_A}{D}\right) + (\mu - \frac{\sigma_A{}^2}{2})T}{\sigma_A\sqrt{T}}$. D is equal to the sum of short-term debt and half of the long-term debt; V_E is market equity value; σ_E is stock return volatility estimated using past 12 monthly returns; μ is stock return in past 12 months. A minimum of three observations is required.

Appendix B. A Brief Introduction of Receiver Operating Characteristic Curve and Accuracy Ratio

The Receiver Operating Characteristic (ROC) curve is a graphical plot that illustrates a binary classification model's diagnostic ability at various discrimination thresholds. The U.S. military first used the ROC curve to detect radio signals from noises after the Japanese attack on Pearl Harbor. The initial research was motivated to determine how and to what extent the U.S. radar receiver operators had missed detecting the Japanese airplanes, after which ROC is named. The ROC curve was then introduced to psychology to account for perceptual detection of stimuli and has been widely used in medicine, radiology, and machine learning.

The ROC curve depicts the relationship between True Positive Rate (TPR) and False Positive Rate (FPR) at various thresholds. The ROC curve is formed by plotting TPR against FPR at different thresholds. Panel A of Figure A1 illustrates the ROC curve of our rating model. We sort the ratings from C to AAA and compute the TPR and FPR for each rating. For example, when C is the default threshold, the TPR is computed as the ratio of the number of firms rated C defaulted to the total number of firms defaulted in 12 months. The FPR is computed as the ratio of the number of firms rated C that did not default to the total number of firms that did not default in 12 months. We then proceed to estimate the TPR and FRP for CC and obtain the second dot, and so on so forth.

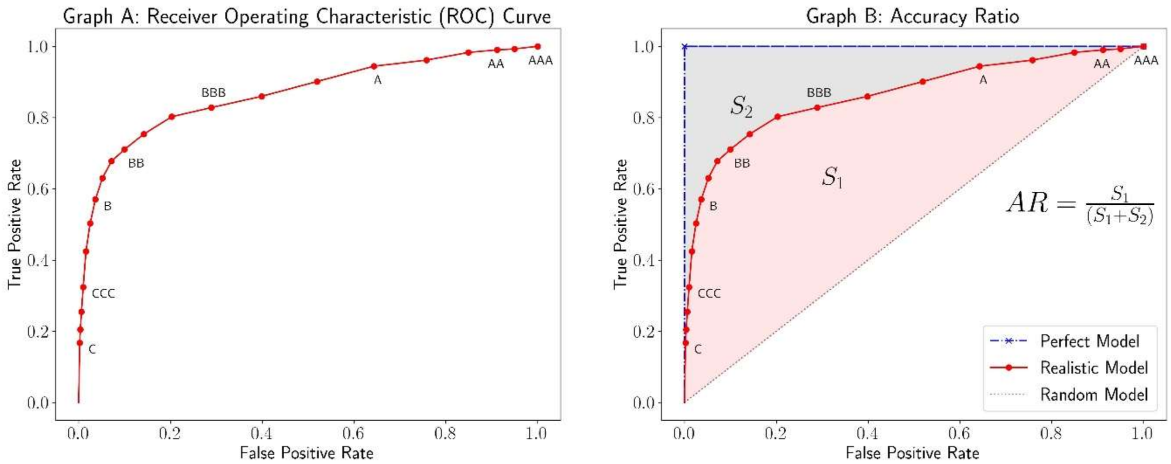

Figure A1. Graphical Illustration of ROC Curve and Accuracy Ratio

Accuracy Ratio (AR) is a numerical measure of the discriminatory power of a classification model. Panel B of Figure A1 depicts the ROC curves of three models. The dash-dot line is the ROC curve of a hypothetical perfect model, for which the TPR is one, and the FPR is zero at an ideal threshold, where default and no-default are perfectly separated. The dotted line represents the ROC curve of a theoretical random model with no discriminative power. Thus, the TPR equals the FPR at every threshold, forming a 45-degree line on the coordinate. The solid line is the ROC curve of a rating model. Denote the space between the ROC curves of the rating model and the random model S1 and the space between the perfect model's ROC curves and the random model S1+S2=0.5. The Accuracy Ratio of the rating model is the ratio of S1 to S1+S2 that equals 2S1. The Accuracy Ratio, taking a value between zero and one, tells how close a model is to a perfect model. The model possesses greater discriminative power as its Accuracy Ratio approaches one. Being a standardized measure, the Accuracy Ratio lets us compare the discriminatory power of multiple models.

Table 1. S&P Ratings and Actual Default Rates

This table reports S&P ratings and published and interpolated one-year actual default rates (ADR). Columns (1) and (2) report numeric ratings corresponding to S&P ratings, respectively. Columns (3) and (4) report the means and standard deviations of S&P ratings' one-year ADRs from 1981 to 2019 (Table 9 of 2019 Annual Global Corporate Default and Rating Transition Study). Column (5) reports the logarithm format of ADRs before interpolation with Cubic Splines. The purpose of interpolating $\text{Log}(\frac{ADR}{100-ADR})$ is to avoid negative ADR estimates. Column (6) reports the interpolated ADRs, and Columns (7) and (8) report the lower and upper boundaries of the ADR bins of S&P ratings, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Numeric Rating	Letter Grades	Reported S&P ADRs Mean	Reported S&P ADRs Stdev	$\text{Log}(\frac{ADR}{100-ADR})$	Fitted ADRs	Lower Boundary of ADR Bins	Upper Boundary of ADR Bins
1	AAA	0.00	0.00	-	0.0073	0.0000	0.0085
2	AA+	0.00	0.00	-	0.0096	0.0085	0.0115
3	AA	0.01	0.07	-9.21	0.0133	0.0115	0.0163
4	AA-	0.02	0.09	-8.52	0.0192	0.0163	0.0240
5	A+	0.04	0.13	-7.82	0.0287	0.0240	0.0365
6	A	0.05	0.11	-7.60	0.0443	0.0365	0.0573
7	A-	0.07	0.20	-7.26	0.0702	0.0573	0.0921
8	BBB+	0.12	0.29	-6.72	0.1140	0.0921	0.1511
9	BBB	0.21	0.34	-6.16	0.1881	0.1511	0.2512
10	BBB-	0.25	0.42	-5.99	0.3143	0.2512	0.4215
11	BB+	0.49	0.89	-5.31	0.5286	0.4215	0.7096
12	BB	0.70	0.82	-4.95	0.8905	0.7096	1.1920
13	BB-	1.19	1.64	-4.42	1.4934	1.1920	1.9854
14	B+	2.08	2.04	-3.85	2.4774	1.9854	3.2573
15	B	5.85	4.91	-2.78	4.0371	3.2573	5.2248
16	B-	8.77	7.44	-2.34	6.4125	5.2248	8.1295
17	CCC+	24.34	11.36	-	9.8465	8.1295	12.1720
18	CCC	24.34	11.36	-	14.4974	12.1720	17.4117
19	CCC-	24.34	11.36	-	20.3259	17.4117	23.6719
20	CC	24.34	11.36	-1.13	27.0179	23.6719	30.5185
21	C	24.34	11.36	-	34.0191	30.5185	100.0000

Table 2. Defaults and ST Firms by Year

This table reports the numbers of listed firms, Special Treatment (ST) firms, and defaulted firms in each year in our sample. The percentage is computed as the number of distressed (ST, Defaulted, ST and Defaulted) firms divided by the year-end number of listed firms.

Year	Number of Listed Firms	ST Firms		Defaulted Firms		ST & Defaulted Firms	
		Number	Percentage	Number	Percentage	Number	Percentage
1998	889	24	2.70%	0	0.00%	24	2.70%
1999	1016	34	3.35%	0	0.00%	34	3.35%
2000	1154	28	2.43%	0	0.00%	28	2.43%
2001	1257	23	1.83%	0	0.00%	23	1.83%
2002	1323	41	3.10%	0	0.00%	41	3.10%
2003	1358	67	4.93%	0	0.00%	67	4.93%
2004	1469	41	2.79%	0	0.00%	41	2.79%
2005	1482	35	2.36%	0	0.00%	35	2.36%
2006	1482	62	4.18%	0	0.00%	62	4.18%
2007	1577	59	3.74%	0	0.00%	59	3.74%
2008	1703	31	1.82%	0	0.00%	31	1.82%
2009	1721	31	1.80%	0	0.00%	31	1.80%
2010	2004	40	2.00%	0	0.00%	40	2.00%
2011	2336	19	0.81%	0	0.00%	19	0.81%
2012	2547	26	1.02%	0	0.00%	26	1.02%
2013	2583	21	0.81%	0	0.00%	21	0.81%
2014	2634	31	1.18%	1	0.04%	32	1.21%
2015	2892	31	1.07%	2	0.07%	33	1.14%
2016	2990	43	1.44%	0	0.00%	43	1.44%
2017	3424	41	1.20%	2	0.06%	43	1.26%
2018	3642	47	1.29%	14	0.38%	61	1.67%
2019	3756	75	2.00%	17	0.45%	92	2.45%

Table 3. Descriptive Statistics of Default Predictors

This table reports the summary statistics of default predictors. DTD denotes Merton (1974) model distance-to-default; RSize is the logarithm of market equity value normalized by the average market value of the component stocks of Shanghai Stock Exchange Composite Index; MB denotes ratio of market equity value to book equity value; WC_MTA denotes ratio of working capital to market asset value, which is equal to the sum of market equity value and total liabilities; RE_MTA denotes ratio of retained earnings to market asset value; EBIT_MTA denotes ratio of earnings before interest and taxes to market asset value; (7) SALE_MTA denotes ratio of sales to market asset value; NI_MTA denotes ratio of net income to market asset value; TL_MTA denotes ratio of total liabilities to market asset value; CASH_MTA denotes ratio of cash plus cash equivalents to market asset value; ExRet denotes trailing 12-month excess return over the Shanghai Stock Exchange Composite Index return; β represents systematic risk estimated using the Dimson (1979) model; Vol^{Idio} represents idiosyncratic risk; ILLIQ represents illiquidity estimated using the method of Amihud (2002). Superscript MA denotes trailing 12-month moving average value. Superscript Diff denotes the difference between the current value and the moving average value. See Appendix A for more detailed variable information. N denotes the number of firm-month observations. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	(1) Default Indicator = 0 (N=462647)			(2) Default Indicator = 1 (N=9831)			(2)-(1)
Variables	Mean	Stdev.	Median	Mean	Stdev.	Median	Mean
DTD^{MA}	7.45	4.73	6.59	5.08	4.27	4.53	-2.37***
DTD^{Diff}	-0.05	3.41	-0.12	-0.40	3.19	-0.37	-0.34***
$RSize^{MA}$	-1.12	1.12	-1.16	-1.46	1.07	-1.37	-0.34***
$RSize^{Diff}$	-0.01	0.19	-0.03	-0.09	0.21	-0.10	-0.08***
MB^{MA}	4.04	4.41	3.02	4.80	7.37	2.97	0.76***
MB^{Diff}	-0.16	2.28	-0.11	0.06	4.19	-0.01	0.22***
WC_MTA^{MA} (%)	9.46	14.29	8.13	3.49	16.34	3.23	-5.97***
WC_MTA^{Diff} (%)	0.02	4.38	0.00	-1.70	5.16	-1.18	-1.71***
RE_MTA^{MA} (%)	7.12	9.72	5.89	-0.18	12.19	1.22	-7.29***
RE_MTA^{Diff} (%)	0.33	2.45	0.28	-2.16	4.01	-1.48	-2.49***
$EBIT\_MTA^{MA}$ (%)	0.92	1.12	0.76	-0.60	1.57	-0.37	-1.52***
$EBIT\_MTA^{Diff}$ (%)	-0.04	0.97	-0.03	-0.24	2.20	-0.01	-0.20***
$SALE\_MTA^{MA}$ (%)	10.29	11.25	6.89	8.60	10.10	5.61	-1.69***
$SALE\_MTA^{Diff}$ (%)	-0.11	3.75	-0.06	-0.80	4.07	-0.39	-0.69***
NI_MTA^{MA} (%)	0.62	0.95	0.52	-1.01	1.58	-0.65	-1.62***
NI_MTA^{Diff} (%)	-0.03	0.87	-0.03	-0.24	2.15	0.00	-0.20***
TL_MTA^{MA} (%)	27.09	22.91	20.09	36.28	23.01	32.36	9.19***
TL_MTA^{Diff} (%)	0.77	4.94	0.49	1.35	6.16	1.15	0.58***
$CASH\_MTA^{MA}$ (%)	9.24	8.50	6.82	6.30	6.41	4.36	-2.94***
$CASH\_MTA^{Diff}$ (%)	0.04	3.24	-0.09	-0.40	2.92	-0.24	-0.44***
ExRet (%)	7.76	43.46	-1.98	-6.62	42.64	-16.09	-14.37***
ILLIQ	0.04	0.25	0.01	0.06	0.30	0.01	0.02***
β	1.15	0.46	1.12	1.18	0.47	1.17	0.03***
Vol^{Idio}	0.37	0.13	0.35	0.41	0.13	0.40	0.04***

Table 4. Chinese Domestic ratings and S&P Ratings

This table reports the statistics of expected one-year default probabilities (PD, in percentage) of the firm-month observations and their domestic agency ratings and ADR-implied S&P ratings. Columns (1) and (2) present the domestic CRA ratings and numbers of observations, respectively. Columns (3)-(11) report the statistics of expected default probabilities. Ratings in parentheses are S&P ratings, whose ADR bins contain PDs reported in front of them. P10, P25, P75, and P90 denote the 10, 25, 75, and 90 percentiles. Note that the statistics of ratings below A may not be robust due to small observations, and they are reported for reference only.

(1) Domestic Ratings	(2) Obs	(3) Mean	(4) Std	(5) Min	(6) p10	(7) p25	(8) Median	(9) P75	(10) p90	(11) Max
AAA	11112	1.13(*BB*)	0.02	0.00(*AAA*)	0.09(*BBB*+)	0.24(*BBB*)	0.57(*BB*+)	1.25(*BB*-)	2.42(*B*+)	58.66(*C*)
AA+	12788	1.28(*BB*-)	0.02	0.00(*AAA*)	0.22(*BBB*)	0.41(*BBB*-)	0.79(*BB*)	1.47(*BB*-)	2.58(*B*+)	99.68(*C*)
AA	28049	1.61(*BB*-)	0.03	0.00(*AAA*)	0.26(*BBB*-)	0.58(*BB*+)	1.05(*BB*)	1.82(*BB*-)	3.00(*B*+)	99.71(*C*)
AA-	11960	2.27(*B*+)	0.05	0.00(*AAA*)	0.40(*BBB*-)	0.76(*BB*)	1.37(*BB*-)	2.34(*B*+)	3.87(*B*)	99.96(*C*)
A+	4370	2.86(*B*+)	0.06	0.01(*AA*-)	0.50(*BB*+)	0.86(*BB*)	1.61(*BB*-)	3.02(*B*+)	5.02(*B*)	98.88(*C*)
A	919	3.86(*B*)	0.11	0.01(*AA*-)	0.36(*BBB*-)	0.92(*BB*)	1.70(*BB*-)	3.46(*B*)	5.36(*B*-)	99.95(*C*)
Below A	437	18.83(*CCC*-)	0.32	0.00(*AAA*)	0.23(*BBB*)	0.60(*BB*+)	1.88(*BB*-)	21.75(*CCC*-)	84.81(*C*)	99.98(*C*)
A-	128	2.98(*B*+)	0.12	0.07(*A*-)	0.18(*BBB*)	0.28(*BBB*-)	0.69(*BB*+)	1.41(*BB*-)	2.79(*B*+)	90.99(*C*)
BBB+	117	4.25(*B*)	0.11	0.00(*AAA*)	0.08(*A*-)	0.49(*BB*+)	1.69(*BB*-)	2.70(*B*+)	6.04(*B*-)	98.70(*C*)
BBB	87	28.92(*CC*)	0.37	0.02(*A*+)	0.45(*BB*+)	1.04(*BB*)	5.06(*B*)	57.06(*C*)	97.61(*C*)	99.98(*C*)
BBB-	7	7.09(*B*-)	0.16	0.21(*BBB*)	0.25(*BBB*-)	0.37(*BBB*-)	1.71(*BB*-)	2.20(*B*+)	18.45(*CCC*-)	42.55(*C*)
BB+	13	65.28(*C*)	0.35	0.73(*BB*)	19.41(*CCC*-)	30.63(*C*)	78.79(*C*)	94.06(*C*)	96.80(*C*)	98.36(*C*)
BB	26	52.09(*C*)	0.38	0.54(*BB*+)	2.81(*B*+)	21.82(*CCC*-)	44.97(*C*)	94.23(*C*)	98.90(*C*)	99.92(*C*)
BB-	10	1.31(*BB*-)	0.01	0.63(*BB*+)	0.69(*BB*+)	0.71(*BB*)	1.06(*BB*)	1.75(*BB*-)	2.08(*B*+)	2.91(*B*+)
B	10	48.74(*C*)	0.22	15.73(*CCC*)	18.97(*CCC*-)	30.89(*C*)	53.90(*C*)	60.28(*C*)	70.91(*C*)	84.66(*C*)
B-	1	7.96(*B*-)		7.96(*B*-)	7.96(*B*-)	7.96(*B*-)	7.96(*B*-)	7.96(*B*-)	7.96(*B*-)	7.96(*B*-)
CCC	18	73.72(*C*)	0.33	0.64(*BB*+)	33.51(*C*)	40.72(*C*)	95.82(*C*)	98.44(*C*)	98.97(*C*)	99.88(*C*)
CC	12	9.51(*CCC*+)	0.09	0.68(*BB*+)	3.49(*B*)	4.14(*B*)	5.13(*B*)	10.80(*CCC*+)	25.40(*CC*)	26.50(*CC*)
C	8	78.99(*C*)	0.22	30.47(*CC*)	58.02(*C*)	75.83(*C*)	83.85(*C*)	89.89(*C*)	99.66(*C*)	99.82(*C*)

Table 5. Statistics of Corporate Yield Spread Regression Variables

This table reports the summary statistics and univariate correlation coefficients of the key regression variables. Spread denotes the difference between month-end yield to maturity (YTM) and Treasury yield of matched maturity. Implied Rating means PD-implied rating; Rating means rating assigned by one of the domestic rating agencies. Log (Issue Size) is the logarithm of bond issue amount in a million yuan. Log (Time to Mat) is the logarithm of time to maturity in years. Coupon denotes coupon rate. Log (TrdVol) is the logarithm of total trading volume in the month in a million yuan. Stock Vol is the annualized 12-month stock return volatility; Leverage denotes the ratio of total liabilities to total assets. Except for implied rating and rating, all variables are winsorized at the 1st and 99th percentiles level, respectively.

Panel A: Summary Statistics

Variable	Unit	Obs.	Mean	St. Dev.	P5	P10	P50	P90	P95
Spread	bp	21354	254.19	189.02	57.08	147.11	216.47	309.54	560.92
Implied Rating		21354	12.03	2.29	8.00	11.00	12.00	13.00	15.00
Rating		21354	2.62	1.08	1.00	2.00	3.00	3.00	4.00
Log (Issue Size)		21354	6.95	0.79	5.71	6.40	6.91	7.38	8.29
Log (Time to Mat)		21354	1.28	0.51	0.29	0.95	1.36	1.66	2.02
Coupon	%	21354	5.93	1.16	3.85	5.20	5.88	6.80	7.80
Log (TrdVol)		21354	2.82	1.68	0.04	1.44	3.02	4.11	5.40
Stock Vol		21354	0.41	0.19	0.17	0.27	0.37	0.51	0.79
Leverage	%	21354	57.53	15.44	30.60	46.52	58.34	69.26	81.35

Panel B: Univariate Correlation Coefficients

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Spread	(1)	1.00								
Implied Rating	(2)	0.22	1.00							
Rating	(3)	0.45	0.31	1.00						
Log (Issue Size)	(4)	-0.26	-0.13	-0.58	1.00					
Log (Time to Mat)	(5)	-0.04	0.07	-0.02	0.07	1.00				
Coupon	(6)	0.46	0.18	0.56	-0.39	-0.02	1.00			
Log (TrdVol)	(7)	0.11	0.09	0.11	0.15	-0.05	0.18	1.00		
Stock Vol	(8)	0.08	0.19	0.23	-0.17	0.02	0.17	0.08	1.00	
Leverage	(9)	0.08	0.26	0.00	0.31	0.05	0.02	0.10	0.02	1.00

Table 6. Corporate Yield Spreads and PD-Implied Ratings

This table reports the results of the following regression:

$$Spread_{i,t+1} = \theta_1 * Implied\ Rating_{i,t} + \sum_{i=2}^{N} \theta_i * Control_{i,t} + \varepsilon_{i,t},$$

where $Spread_{i,t+1}$ denotes yield spread; Implied rating denotes PD-implied rating; Rating denotes rating assigned by a domestic rating agency. Log (Issue Size) is the logarithm of bond issue amount in a million yuan. Log (Time to Mat) is the logarithm of time to maturity in years. Coupon denotes coupon rate. Log (TrdVol) is the logarithm of total trading volume in the month in a million yuan. Stock Vol is the annualized 12-month stock return volatility; Leverage denotes the ratio of total liabilities to total assets. We further control for the firm and year-month fixed effects and compute standard errors clustered by firm. *, ** and *** denote the statistical significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample				AAA	AA+	AA	AA-
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Implied Rating	13.88***	11.52***	11.69***	9.53***	6.82**	7.94**	14.02***	18.06***
	(6.40)	(5.63)	(5.86)	(4.83)	(2.16)	(2.27)	(4.69)	(4.32)
Rating		45.13***	38.99***	37.80***				
		(5.75)	(5.16)	(5.07)				
Log (Issue Size)			-32.59**	-35.50**	-27.57**	-62.11*	3.23	-28.04
			(-2.21)	(-2.49)	(-2.22)	(-1.79)	(0.11)	(-0.22)
Log(Time to Mat)			3.25	1.84	10.26	6.23	4.90	71.00
			(0.25)	(0.14)	(0.94)	(0.25)	(0.18)	(1.19)
Coupon			43.15***	43.34***	29.00***	35.24**	49.28***	50.30*
			(6.02)	(6.20)	(4.13)	(2.37)	(4.45)	(1.97)
Log (Volume)			2.81**	2.64*	2.68	1.97	2.48*	2.99
			(2.03)	(1.91)	(1.20)	(0.74)	(1.66)	(0.91)
Stock Vol				-7.16	40.81**	70.15	-39.95	2.23
				(-0.38)	(2.21)	(1.31)	(-1.38)	(0.06)
Leverage				1.66***	-0.67	1.02	1.33**	0.93
				(3.65)	(-0.82)	(0.99)	(2.18)	(0.97)
Firm FE	√	√	√	√	√	√	√	√
Year-Month FE	√	√	√	√	√	√	√	√
Observations	21354	21354	21354	21354	3918	4263	9969	2761
Adj. R^2	0.540	0.553	0.564	0.566	0.513	0.448	0.550	0.678

Table 7. Domestic Agency Ratings and Default Predictors

This table reports the regression results of domestic agency ratings on the default predictors in our estimation model. We use the Ordered Probit model, the OLS model, and the Fama-Macbeth approach in the agency rating regressions. Selected # and Selected (%) represent the importance of the predictors in our logistic model through Elastic Net Procedures. Selected # is the number of times that the predictors are selected to predict defaults in the 156 months from January 2007 and December 2019. Selected (%) presents the percentage of times a predictor is selected (Selected #/156). *, ** and *** denote the statistical significance at the 10%, 5%, and 1% levels, respectively. Pseudo R-squared, Adjusted R-squared, and the average R-squared values are reported in Columns (1) to (3).

	(1) Ordered Probit		(2) OLS		(3) Fama-Macbeth		(4) Elastic Net	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Selected #	Selected (%)
Significant/Important predictors of agency ratings and PD-implied ratings								
$RSize^{Diff}$	1.28***	(10.94)	0.58***	(5.33)	0.27***	(2.68)	151	96.79%
RE_MTA^{Diff}	2.22***	(3.17)	1.89**	(2.56)	-2.07***	(-2.84)	152	97.44%
RE_MTA^{MA}	5.78***	(11.36)	2.40***	(5.69)	1.53***	(6.20)	137	87.82%
Important predictors of PD-implied ratings only								
NI_MTA^{MA}	12.98	(0.81)	44.34***	(3.43)	66.65***	(12.82)	156	100.00%
NI_MTA^{Diff}	5.21	(1.39)	-0.30	(-0.09)	-6.21	(-1.30)	156	100.00%
$CASH\_MTA^{MA}$	0.28	(0.55)	-0.27	(-0.72)	-0.18	(-0.77)	156	100.00%
ExRet	-0.33***	(-6.40)	0.06*	(1.77)	0.05	(1.23)	71	45.51%
Significant predictors of agency ratings only								
DTD^{MA}	0.07***	(9.13)	-0.01***	(-2.89)	0.03***	(6.06)	23	14.74%
MB^{MA}	-0.14***	(-9.32)	-0.06***	(-4.03)	-0.11***	(-15.88)	3	1.92%
MB^{Diff}	-0.07***	(-6.41)	-0.05***	(-4.33)	-0.07***	(-4.34)	0	0.00%
$RSize^{MA}$	1.35***	(26.47)	0.79***	(12.47)	0.75***	(64.56)	0	0.00%
TL_MTA^{MA}	1.91***	(9.87)	0.93***	(4.47)	0.82***	(12.89)	1	0.64%
TL_MTA^{Diff}	1.00***	(3.19)	0.82***	(2.95)	0.91***	(4.65)	11	7.05%
Vol^{Idio}	-0.61***	(-2.83)	-1.12***	(-4.64)	-1.44***	(-13.99)	3	1.92%
Insignificant or inconsistently significant predictors								
DTD^{Diff}	0.00	(0.84)	-0.01***	(-3.31)	0.00	(0.36)	0	0.00%
WC_MTA^{MA}	-0.91***	(-4.18)	0.23	(1.03)	-1.05***	(-11.26)	58	37.18%
WC_MTA^{Diff}	0.23	(0.87)	0.30	(1.38)	0.82***	(3.35)	0	0.00%
$EBIT\_MTA^{MA}$	-10.74	(-0.77)	-27.70***	(-2.76)	-38.34***	(-9.42)	8	5.13%
$EBIT\_MTA^{Diff}$	-4.87	(-1.41)	-0.44	(-0.16)	11.39**	(2.61)	1	0.64%
$SALE\_MTA^{MA}$	0.16	(0.53)	0.77**	(2.30)	-0.25**	(-2.50)	14	8.97%
$SALE\_MTA^{Diff}$	0.08	(0.33)	0.04	(0.28)	-0.69***	(-2.72)	3	1.92%
$CASH\_MTA^{Diff}$	0.04	(0.11)	-0.31	(-1.16)	-1.60***	(-5.14)	0	0.00%
ILLIQ	1.48	(0.73)	-1.25	(-0.91)	-7.17***	(-2.76)	0	0.00%
β	0.12***	(3.26)	0.03	(1.56)	-0.05*	(-1.93)	19	12.18%
Obs	69,494		69,494		69,494			
Firm & YM FE	-		√		-			
Cluster	Firm		Firm		-			
R-squared	0.33		0.82		0.65			

Figure 1. Rating Distributions and Real Default Rates

This figure depicts in Graph A the firm-month observations' distributions according to PD-implied S&P ratings (red) and domestic agency ratings (blue). The purple color represents their overlapping; in Graph B, the real default rates of PD-implied S&P ratings (dotted line with X) and domestic agency ratings (dotted line with O) in comparison to the ADRs of S&P global ratings (dotted line with diamond). R-squared is expressed as

$$\text{R-squared} = 1 - \frac{\sum_{r=1}^{17}(S\&P\ ADR_r - Real\ Default\ Rate_r)^2}{\sum_{r=1}^{17}(S\&P\ ADR_r - \overline{S\&P\ ADR_r})^2}.$$

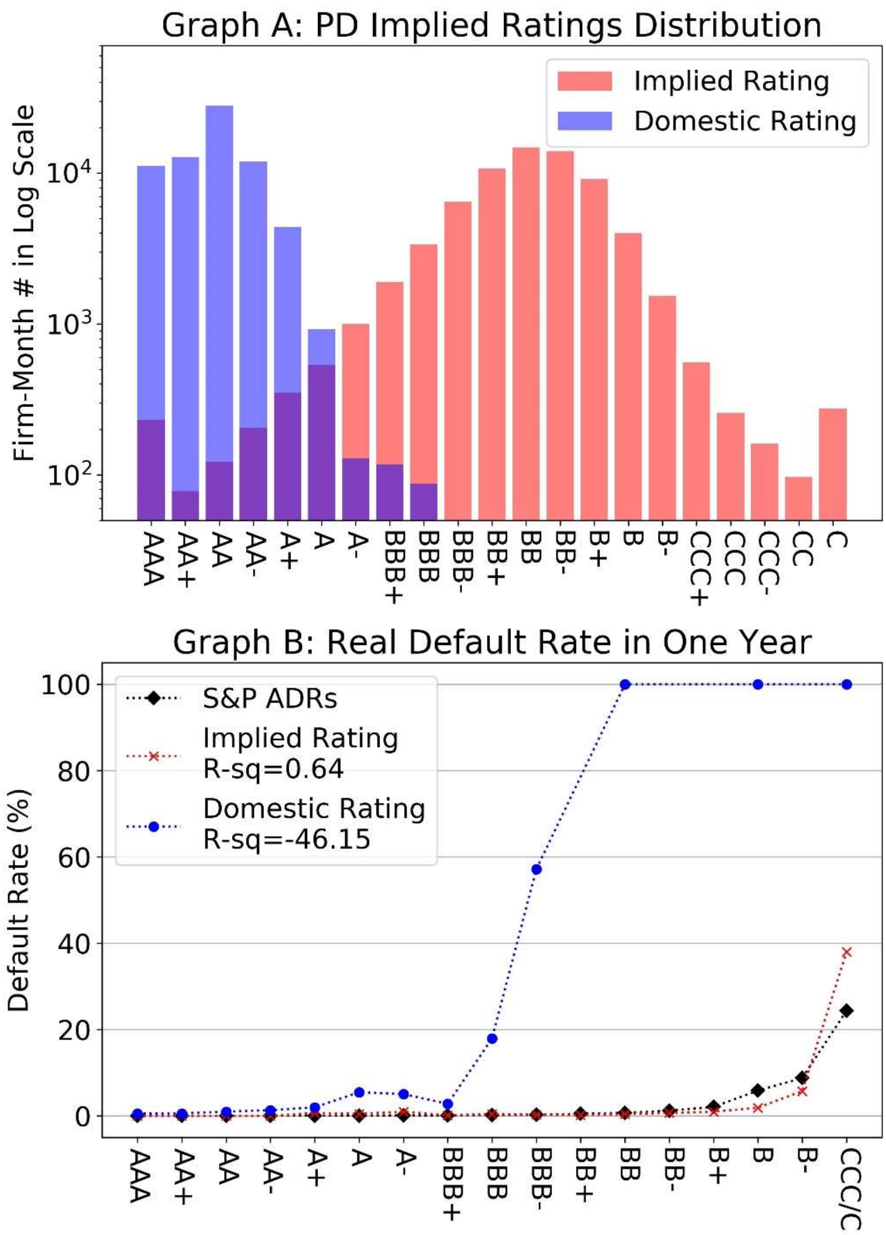

Figure 2. Time Series Patterns of Rating Composition

This figure depicts the time series patterns of rating composition of domestic agency ratings in Graph A and PD-implied S&P ratings in Graph B. In both graphs, the vertical axis represents the cumulative percentage of ratings, with the highest rating AAA at the bottom and the lowest rating C on the top. To facilitate illustration, we combine ratings by letter category. For example, AA+, AA, and AA- are reported as AA. Moreover, CCC, CC, and C are combined into one single CCC/C category.

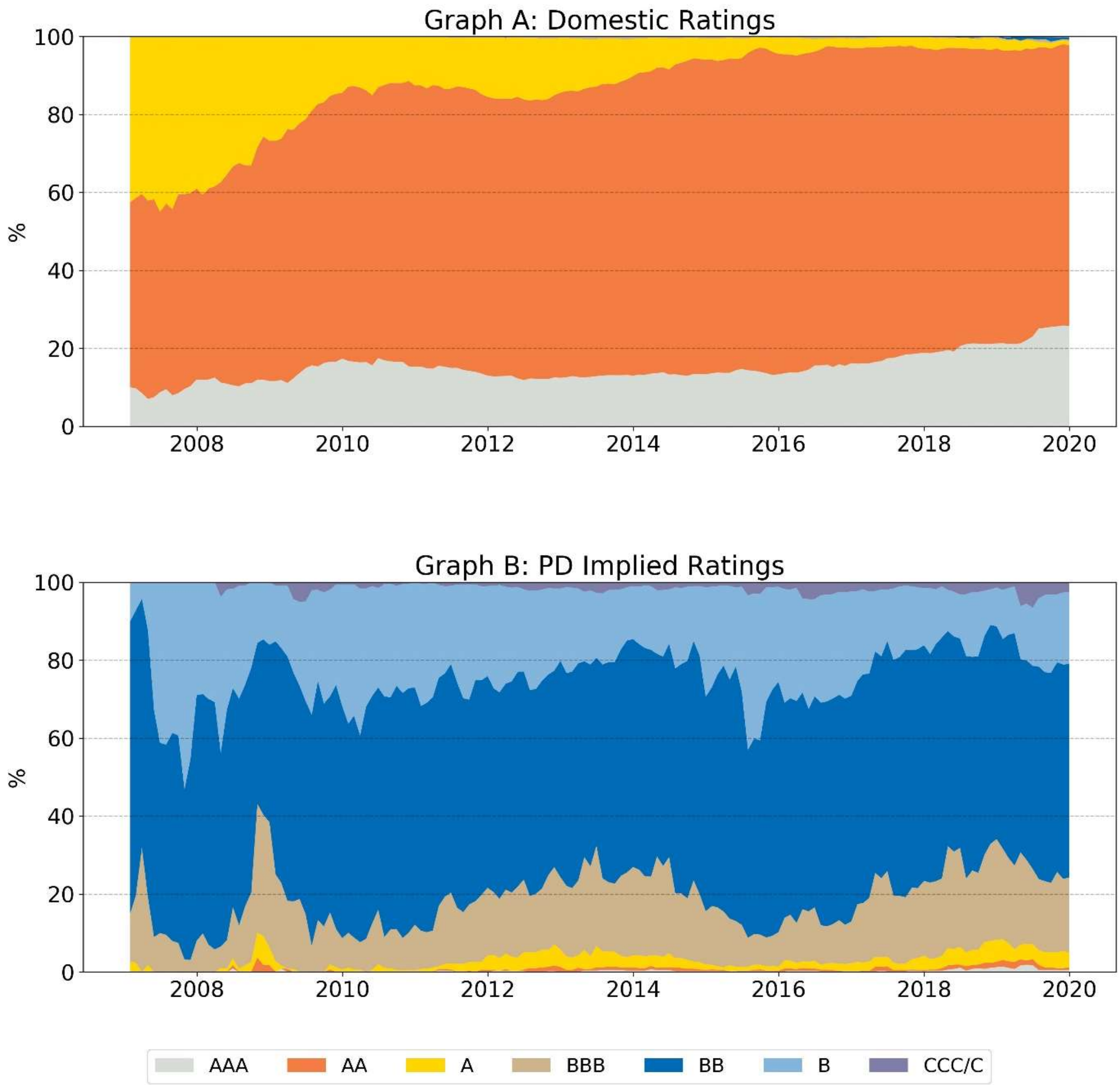

Figure 3. Default Differentiation: ROC Curve and Accuracy Ratio

The figure depicts in Graph A the receiver operating characteristic (ROC) curves and Accuracy Ratio (AR) of PD-implied S&P ratings and domestic agency ratings. The 45-degree line represents the ROC curve of a hypothetical random model with no detective power of default. Graph B depicts the PD-implied S&P ratings and domestic agency ratings 24 months before distress. Graph C illustrates the PD-implied S&P rating and the domestic agency rating of Shanghai ChaoRi Solar that experienced the first publicly issued bond default in China. See Appendix B for a brief introduction of the ROC curve and AR.

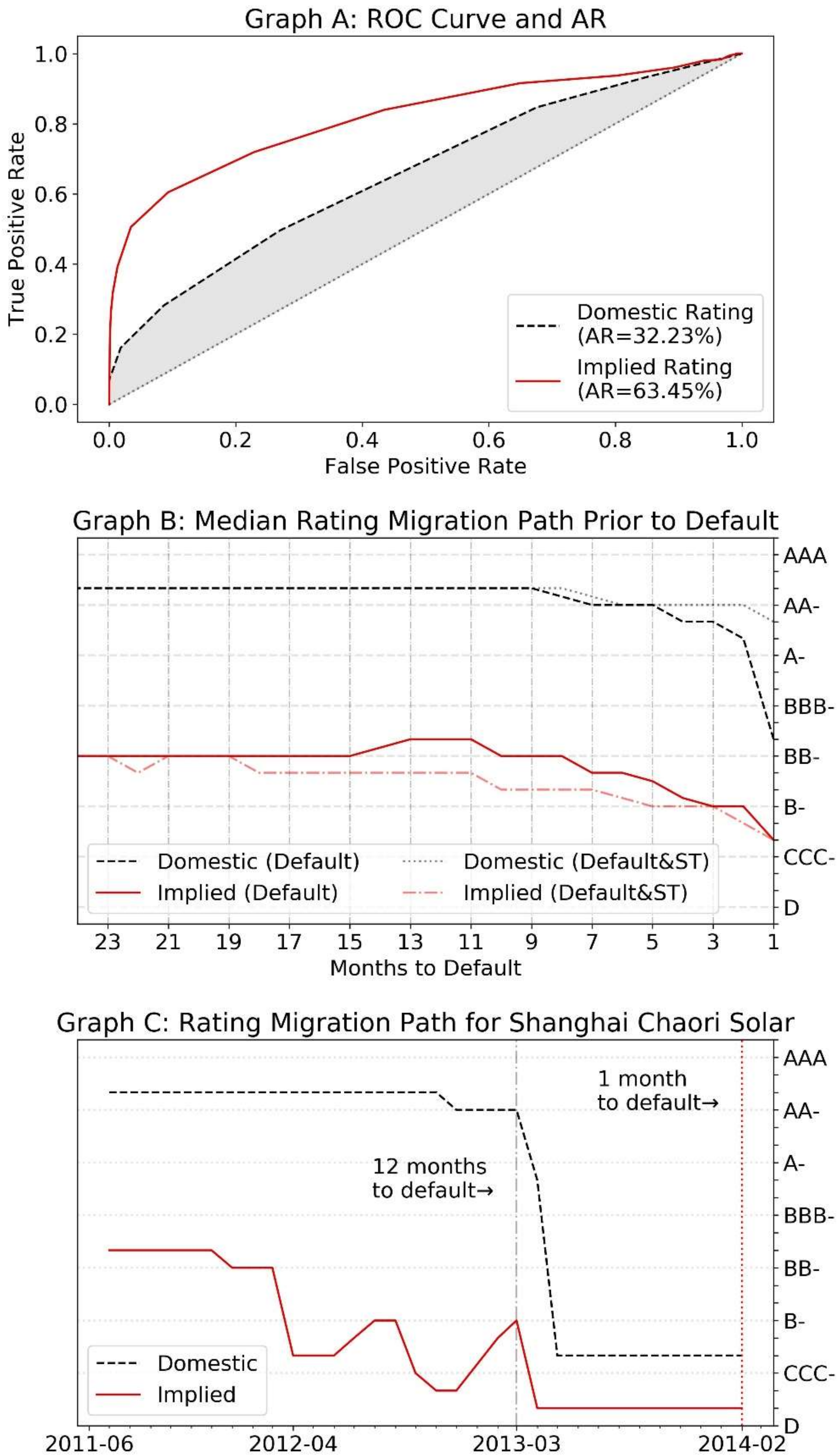

Figure 4. Default Prediction Accuracy of Alternative Models

This figure depicts the real default rates of PD-implied ratings and R-Squareds' of the baseline model and the following models: Merton (1974) model, cost-sensitive logistic model (being cost-sensitive means that balanced weights are assigned to the default and non-default samples), penalized logistic model (Elastic Net), random forest model, cost-sensitive random forest models, XGBoost, cost-sensitive XGBoost, Neural Network with hidden layers from one to five. Gu, Kelly and Xiu (2020) provide an excellent description of the machine learning models. The R-squared measure is expressed as

$$\text{R-squared} = 1 - \frac{\sum_{r=1}^{17}(S\&P\ ADR_r - Real\ Default\ Rate_r)^2}{\sum_{r=1}^{17}(S\&P\ ADR_r - \overline{S\&P\ ADR_r})^2}.$$

The dotted line presents S&P ADRs; the dotted line with diamond represents the real default rates of the baseline model-implied S&P ratings; the dotted line with X means the real default rates of alternative model-implied S&P ratings. The R-Squared of the baseline model is 0.641, and the R-Squareds of alternative models are reported in their graphs, respectively.