How to compare two relative rankings?

Hatef Dastour (U. Calgary), Qiwei Feng (U. Manitoba),
Neal Madras (York U.), Weifei Ouyang (Shanghai Jiao Tong U.), Negin Pasban Roozbahani (U. Manitoba), Alessandro Selvitella (McMaster U.)
Tuan Tran (McMaster U.)

Abstract. This work proposes two methods to compare the creditworthiness of two arbitrary borrowers coming from two different groups: Financial Firms and Oil & Gas Companies. The first method relies on a static measure of Distance to Default, and the second one is based on modelling the credit score as a Markovian process.

Table of Contents

Introduction

Credit ranking plays an important role in banking industry. Banks rely on credit scores of their borrowers rated by rating agencies to decide whether or not they lend money to these borrowers. When borrowers come from the same industry, it is straightforward to compare their creditworthiness based on their credit scores. However, if two borrowers come from two different groups, it is no longer straightforward just to compare their scores because each group has its own industry-specific characteristics. The question is how can we come up with a global measure in order to normalize the different relative rankings.

In this project, we work on a dataset with two groups of firms: the first one consists of firms coming from financial sector (non banks), and the second one consists of firms coming form Oil& Gas industry, each group has a large number of firms (around 2000 and 9000 respectively). Rank of firms is any integer number between 1 to 17, where 1 corresponds to the best credit quality and 17 corresponds to default firms. Firms are rated based on experts' opinion. We suppose that ranks are relatively accurate and try to convert them into some sort of global measure.

We propose in this work two different measures for creditworthiness: the first one relies on a static measure of Distance to Default, and the second one is based on modelling the credit igration as a markovian process.

1 First Approach

We need some notations: Suppose that there are two groups of firms: Group A_i has N_i firms, $i = 1, 2$. The ranks of borrower a at time t is $R_t^a, t = 0, \dots T, R_t^a \in \{1, \dots 17\}$.

The idea is the following:

- Fix a borrower a in one group, say $a \in A_1$. We come up with an static score for a. The easiest way to do this is to take arithmetic mean or weighted average with power weights.
$$\overline{R^a} = \sum_{t=1}^{T} w_t R_t^a.$$
Here, we can chose $w_t = \frac{1}{T}$ or $w_t = \frac{\rho^{t-1}}{1-\rho^T}$ for some $\rho \in (0, 1)$. This measure is still a local one as it does not take into account any industry-related information.
- Now, in each group, we suppose that firms bear common information about the group. By this, it is reasonable to suppose that $(\overline{R^a})_{a \in A_1}$ are drawn from the same distribution function F_1 of some random variable M_1. We can build an empirical distribution for F_1 based on $(\overline{R^a})_{a \in A_1}$.
- Now, if a given borrower comes from group A_1 with local score $\overline{R^a} = x$, then we define the Relative Distance to Default as
$$RDD^1(x) = P(R^1 > 16 | R_1 \geq x) = \frac{P(R^1 > 16)}{P(R^1 \geq x)}.$$
- We repeat the previous procedure for borrowers from group 2. Now if a borrower from Group 1 with local score x and another borrower from Group 2 with local score y, the bank has to compare $RDD^1(x)$ and $RDD^2(y)$ to decide which borrower is riskier to lend money.
- We can go a step further by converting credit scores from Group 1 into those in Group 2. More precisely, we find an creasing mapping $\phi : \{1, \dots, 17\} \to [1, 17]$ that matches the conditional probability of default, i.e. we find ϕ in such a way that
$$RDD^1(x) = RDD^2(\phi(x)).$$

2 Second Approach

In this approach, we suppose that the credit profiles of all borrowers in the same group (say Group 1) are realizations of the same homogeneous Markov process M_1 of constant transition matrix Q and constant generator matrix Λ. By definition Q_{ij} denotes the probability that M_1 is migrated from rank i to rank j after one year. It is a well-known fact that $Q = e^{\Lambda}$.

The idea of this method is the following

- First, we estimate the matrices Q (or equivalently Λ) form the data. There are two ways of doing this. The first one is to estimate the one-month step transition matrix directly from the data. The formula we use is the following
$$Q_{ij}^{1/12} = \frac{N_{ij}[0,T]}{\sum_1^T N_i[t]},$$
where $N_{ij}[0,T]$ denotes the number of migration from state i to state j during the period $[0,T]$ (here $T = 127$ months), and $N_i[t]$ denotes how many firms whose rank is i at time t. Finally, the one-year transition matrix is given by $Q = [Q^{1/12}]^{12}$.
The second way to estimate Q is to estimate the generator matrix Λ first and then Q is given by e^{Λ}. To estimate Λ, we follow Lando and Skodeberg (2012). We have $\Lambda_{ii} = -\sum_{j\neq i} \Lambda_{ij}$. and
$$\Lambda_{ij} := 12\frac{N_{ij}[0,T]}{\sum_1^T N_i[t]}.$$
The only advantage of the second estimation is that it can help deal with the case when there are many zeros in the estimated matrix Q.
- Once we have estimate of Λ, The $\tau-$ year transition matrix is given by $Q^{tau} = e^{\Lambda\tau}$. τ is determined by the bank, typically one year.
- Now, if a borrower from Group 1 has current rank of x, its probability of default in $\tau-$year will be $Q_{x,17}^{tau}$.
- Finally, in order to compare the creditworthiness of two borrowers, one from Group 1 with current score x and one from Group 2 with score y, the bank only needs to compare $Q_{x,17}^{\tau}$ and $P_{y,17}^{\tau}$, where P denotes the transition matrix of Group 2 estimated in the same way for Q.

Another possible way to compare the creditworthiness of two borrowers with current ranks x, y (comming from Group 1 and Group 2, respectively) is to compare the two first passage times when the transition probabilities from these states to default excess certain threshold p_0 (for instance $p_0 = 0.5$). This "half-life" method can be considered as a dual one to the previous method, it gives more information about necessary time for the default event to be likely to happen. More precisely, we define
$$N_x = \inf\{n : Q_{x,17}^{tau} \geq p_0\}$$
and
$$N_y = \inf\{n : P_{y,17}^{tau} \geq p_0\}$$
If $N_x > N_y$ then we can conclude that rank x in Group 1 is less risky than rank y in Group 2 with default level p_0.

Results of the Markov chain analysis were somewhat mixed. The one-year default probabilities were comparable between the two groups for ranks 8 through 15; the probabilities for rank k in Oil and Gas roughly corresponded to rank $k+1$ in Financial Services. The better ranks had very low default probabilities, which were harder to compare. The "half-life" method gave a different picture, suggesting that all but the weakest ranks of Financial companies took significantly longer to default than most of the Oil and Gas companies.