USING CREDIT HISTORY DATA TO MONITOR FINANCIAL STABILITY OF THE BELARUSIAN ECONOMY

A.Yu. Novopoltsev
Belarusian State University
Minsk, BELARUS
e-mail: novopsacha@gmail.com

Abstract

Data on credit histories may be very helpful for analysis of vulnerability of private sector to monitor financial stability in central banks. In National Bank of Belarus, there are credit register including vast database of credit histories of most companies and individuals on daily basis. These data may be used for estimating some useful risk measures such as Probabilities of Default (PD), Loss Given Default (LGD), etc. To construct aggregated measures from daily data special algorithms are developed. The event of default is defined according to Basel methodology. Some issues related to estimation of the risk measures are solved through analysis of the data distribution. The constructed measures can be further used for analytical reporting and statistical models estimation using supervised learning techniques.
Keywords: data science, credit history, financial stability

1 The issues and the data

This study is devoted to a problem of assessment financial vulnerability and credit risk of nonfinancial companies on microdata. Some research projects based on analysis of companies' balance sheets data have been already conducted in National Bank of Belarus [1, 2, 3]. In these projects some analytical and program tools to estimate companies financial risks were proposed. Statistical credit ratings were estimated on a set of financial ratios with cluster analysis algorithm. Unsupervised learning technique was applied because of the lack of real data on defaults. Estimated credit ratings showed close connections with various expert financial indicators used in financial stability reports. But still it was hard to validate the results of classification due to the lack of actual outcomes on defaults.

On contrarily, credit register data may be used to construct credit risk measures entirely based on data. For example, probability of default may be derived as a frequency of accounts with overdue payments that excess certain limit [4]. This approach implies a definition of default in sense of Basel framework [5]. According to the Basel definition, the event of default is defined as 90 days past due in the debt or interest payments by a contract.

The credit register database collects information on credit contracts and collaterals. All changes in contract information or in the history of payments, including debt outstanding, overdue payments (by debt and interest) and group risks for reserves accumulation, are promptly reported by banks to the credit register of National Bank.

All contracts are grouped by borrower (company or individual), type of credit contract, currency, bank that issued the credit.

2 Aggregated measures of default

Now describe the structure of credit register data and corresponding aggregated measures of credit risk which are proposed in this research.

Let $J(i,t) \subset J$ be the set of active credit contracts of a company with number i on date t where J is a set of all credit contracts; $I(t) \subset I$ be the set of companies that has at least one active contract on date t where I is all set of companies in a sample; $a(i) \in \{1,\ldots,A\}$ be the major economic activity of a company i; $b(j) \in \{1,\ldots,B\}$ be a bank number that issued the credit with number $j \in J$.

There are indicators which changes are tracked by credit history: s_{jt} is debt outstanding, p_{jt} is payments overdue. Given the information of the date of last change for each indicator and corresponding amount, we may calculate a number of days overdue for credit contract j on date t, that is denoted as d_{jt}. Actually, it's defined as a number of days in the period when overdue payments p_{jt} were greater than zero from particular date in the past to the last observed date. Also, there are data specific to a credit contract which are currency, amount of credit issued s_j^0. Further on, let s_{jt} represent so called Exposure at Default (EAD) on a contract j on date t.

Now define the rule of default according to the Basel framework for a single contract j:

$$bd_{jt} = 1 \ if \ d_{jt} > 90, \quad else \ \ bd_{jt} = 0 \tag{1}$$

which may be adopted for a company level by using a rule of maximum overdue days by all contracts of a company i, that is $d_{it} = \max\limits_{j \in J(i,t)} \{d_{jt}\}$, and subsequent substitution to formula (1). Note that here the overdue days are calculated based on overdue payments as a whole including overdue payments by debt, interest and service. In fact, the rule (1) means that there is a technical default on a contract, because the actual default can only be a result of judicial procedure.

In reality, the rule (1) should be modified to take into account an amount of overdue payments: if they are too small or if the level of debt outstanding is too low, the positive decision about default according to the rule (1) should be rejected. We can formulize some additional conditions of default, for example, by requiring that overdue payments p_{jt} be greater than a particular amount or ratio of overdue payments p_{jt} to amount of credit s_j^0 be greater than some threshold η that can be expressed by a formula $p_{jt}/s_j^0 > \eta$. For the thresholds expert assessments are usually taken, but we use data driven approach with an aim to get these from distribution of the actual data.

Using the indicators derived above we can calculate Probability of Default (PD) on date t on a level of a company i, an industry a or a bank b that are denoted as pd_{it}, pd_{at}, pd_{bt}, by calculating relative frequency of default occurrences on the data according to the rule like (1).

Another one important characteristic we address is Loss Given Default (LGD) that is, again, calculated on the data. For this task data on collaterals are used. Let K

is a set of all available collateral contracts in database, $K(i,t) \subset K$ is a set of active collaterals that belong to company i as a pledger on date t. For each collateral contract k on date t there are the following indicators: r_{kt} – requirements on the collateral, u_{kt} – the amount of requirement that were reimbursed, v_k^0 – value of collateral. Again, let $b(k) \in \{1, \dots, B\}$ be a bank that issued a collateral contract k. Then, using the data on collaterals we can calculate an approximation of LGD by applying a formula:

$$LGD_{kt} = 1 - u_{kt}/r_{rt}, \tag{2}$$

which provides LGD on a level of a contract. Also we may generalize (2) by calculating LGD on a level of a company, type of industry or a bank. Usually it's done on a level of industry that could be get by aggregating u_{kt}, r_{kt} on a subset of active contracts $K(a,t) \subset K$ on date t that belong to companies which are from industry a.

To generalize, let define aggregated indicators pd_{lt}, lgd_{lt}, ead_{lt}, which are PD, LGD, EAD on date t on a level of aggregation $l \in \{i, a, b\}$, that is on a level of a company, an industry of a bank. The proposed indicators may be used for calculating expected losses from credit portfolio that is expressed by a formula:

$$EL_t = \sum_l PD_{lt} \times LGD_{lt} \times EAD_{lt}, \tag{3}$$

where the sum is done on a set of different industries, types of credit contracts (products), or different types of borrowers.

3 Applications of proposed measures

Credit register data may be very promising for monitoring financial stability of real and banking sectors [6]. This research aims to apply data driven approach to uncover vulnerabilities in real time. Some traditional modeling techniques for panel and macroeconomic data may also be applied on proposed risk measures derived from the data.

We can propose further applications of the risk measures outlined above.

1. Analytical reporting for monitoring financial stability. It was the first intention to do this research. It's expected that the methodology of calculating risk measures developed on a test sample would be applied to the whole database to get timely reports on the current situation.

2. Mathematical modeling of credit risk measures. For example, a logit model to forecast probability of default of companies at one-year horizon could be developed based on classification variable derived from the data on overdue payments. The developed scoring model may help to validate scoring models of commercial banks.

References

[1] Malugin V., Hryn N., Novopoltsev A. (2014). Statistical analysis and econometric modelling of the creditworthiness of non-financial companies. *International Journal of Computational Economics and Econometrics*. Vol. **4 (1/2)**, pp. 130-147.

[2] Malugin V.I., Novopoltsev A.Yu., Hryn N.V., Milevsky P.S. (2018). Statistical analysis of creditworthiness of real sector of Belarusian economy with microdata *Bank Bulletin. – A special issue "Bank Research"*. No. **14**. (in Russian)

[3] Malugin V.I., Novopoltsev A.Yu., Hryn N.V., Pashkevich A.V. (2019). Development of model tools for assessing and predicting the credit risk of banks at the micro level *Bank Bulletin. – A special issue "Bank Research"*. No. **17**. (in Russian)

[4] Antunes A., Gonçalves H., Prego P. (2016) Firm default probabilities revisited. *IFC-ECCBSO-CBRT Conference on "Uses of Central Balance Sheet Data Offices information"*. Ozdere–Izmir, Turkey, 26 September 2016. Mode of Access: https://www.bis.org/ifc/publ/ifcb45i.pdf. Date of Access: 11.03.2019.

[5] Basel II: Revised international capital framework. Bank of International Settlements. Mode of Access: https://www.bis.org/publ/bcbsca.htm. Date of Access: 11.03.2019.

[6] Navapoltseu A.Yu. (2018) Using microdata on companies to monitor financial stability. *The conference of young researchers in economics and finance*. Minsk, Belarus, December 22, 2018. Mode of access: http://eng.beroc.by/conference/conference-of-young-researches/2018/. Date of Access: 11.03.2019.