

UDC 336.71

DOI: 10.18523/2519-4739.2021.6.1.73-83

V. Filatov, A. Kaminsky

APPLICATION OF THE SCORING APPROACH TO MONITORING FUNCTION OF CENTRAL BANK CREDIT REGISTRY

The Central Bank Credit Registry was established in Ukraine in 2018. The two key functions which are fulfilled by Credit Register are monitoring and credit information sharing. This paper is devoted to applying a scoring approach for monitoring function realization in segments of individuals. The logic of using scoring tools to monitoring is based on an objective to create an effective form which reflects the dynamic of the above-mentioned segment. Data mining procedures for Credit Registry were realized and most significant characteristics were chosen. Correlation analysis for characteristics was applied. Different approaches to construct scoring for monitoring functions were analyzed. Namely, logistic regression, Machine Learning, method grounded on tree created by the XGBoost algorithm. Last method demonstrated the best efficiency for scoring construction and can be developed for implementation.

The views expressed are those of the authors and do not necessarily reflect those of the National Bank of Ukraine.

Keywords: credit risk, credit registry, central bank, scoring.

JEL classification: G21

Introduction and research problem. Financial stability represents one of the most important components of an economy healthy development. The stable financial system ensures the efficiency of financial intermediation functioning, allows to promptly indicate financial risks and manages them. In conditions of stability, the system provides endogenous mechanisms to absorb shock effects in the economy. In addition, it allows to adjust the price fluctuations of real and financial assets, which objectively preserves monetary stability. Thus, financial stability, in essence, protects the real economy from a wide range of destructive impacts.

One of the key components of financial stability is the stability of the credit market. This is especially true for the Ukrainian financial system in which the credit market occupies a dominant position. The instability has been demonstrated many times through the genesis of the Ukrainian credit market. One of the main reasons for this instability was the high level of non-performing loans (NPL) in the banking system accumulated over the years. Moreover, the NPL problem is also relevant today. At the beginning of 2021, the share of NPL in the banking system of Ukraine was 41 %. NPL distribution is not uniform throughout the system. The segment of lending to legal entities includes 46 % of NPL. Analogical figure for the segment of

lending to individuals is 25 %. Furthermore 70 % of NPL falls on state banks portfolios (about 42 % refers to bank “Privat”).

According to the National Bank of Ukraine point of view, the high level of NPL is the result of credit expansion in previous periods. This credit expansion was accompanied by relatively poor systems of borrower’s solvency assessment. The rights of creditors were poorly protected. In addition, there was a fairly common practice of lending related persons, which was negatively manifested in crisis conditions.

However, this NPL does not restrain the development of lending due to full provision. The potential risks and future default are a threat to financial stability.

To a large extent, this led to the adoption of the Law “On Amendments to Certain Laws of Ukraine on Establishing and Maintaining of the Credit Register of the National Bank of Ukraine and Improving Credit Risk Management of Banks”. According to this Law the facilitating financial stability, banking supervision and bank’s credit risk management of Credit Risk is seen in the context of Ukrainian national security in the economic sphere (Article 67 of this Law).

Credit Registry was created at the National Bank of Ukraine in 2018 and started in 2019. Monitoring function and information sharing are key goals of it.

The first function directly finds an embodiment of banking supervision. The second function ensures the efficiency of credit risk management due to the reduction of information asymmetry.

It should be noted that approaches to information sharing at modern credit markets are quite diverse. Basically, it corresponds to some form of “complementation scheme” of functioning of the Central Bank credit registry and such institutions as bureaus of credit histories (BCH, credit bureaus). In general, the Central Bank Credit Registry and the BCHs operate with similar information about credit information, but there are some differences. Basic differences in goals of collecting information. Goals of the Central Bank are better insights into national trends around lending and providing borrowers with credit reports, which are detailing to (typically) high amounts loans. BCHs are more focused on the consumer loans segment. They also produce different estimations (scorings, alerts, antifraud systems and other) for creditors.

Since 2006, the market of credit bureaus has been developing in Ukraine. Kaminsky (2013) analyzed the genesis of credit bureaus in Ukraine. The results of it correspond to covering the consumer lending segment (dominantly unsecured loans) at almost 100%. At the same time, in the segment of corporate loans, loans to small and medium businesses such as covering are significantly less. Therefore, in our opinion, the abovementioned Law on the Credit Register found in many ways the optimal solution for the system of credit information sharing. In many ways, these two institutions organically complement each other in an information aspect. It should be noted that some monitoring functions can be constructed by the credit bureau data. Kaminsky (2015) described the conception of benchmark, which can be realized on bureaus data.

Our research is focused on the monitoring function of the Credit Registry of the National Bank of Ukraine. Namely, in the form of such monitoring. Research is devoted to construct effective tools for the realization of the monitoring function of the Credit Registry. The proposed approach is based on the use of scoring methodology. The advantage of this approach is to integral representation of the credit risk level of the portfolio presented in the Credit Registry. The dynamics of the scoring values distribution, in our opinion, effectively implies changes in the bank lending market. And this, in turn, allows the NBU insight in trends around lending and to approve timely decisions. Qualitative monitoring is one of the building blocks of financial stability.

We have applied various methods of scoring construction and the most efficient method was found.

Recent publications analysis. Credit Registry data is a powerful data source in many countries, both in Europe and globally. However, the introduction of the Central Bank Credit Registry (CR) had a different effect on the market. Aspects of CR’s functions and roles are subject for discussions. Thus, researchers Ralph De Haas, Matteo Millone, Jaap Bos (2016) based on the data from Bosnia and Herzegovina showed that establishing the CR reduced the real credit risk for new loans, as well as reduced interest rates for recurrence customers with a good credit history. Allen N. Berger et al. (2011) assessed the effect of collateral on credit policy and concluded that data from the Credit Registry helped to distinguish the effect of collateral from other effects clearly. Moreover, Bennardo et al. (2010) showed that collateral for the introduction of centralized CR increases the credit availability.

On the other hand, the monitoring function of the CR may significantly increase the effectiveness of Central bank regulatory policy. Researchers from the Bank of France, Dietsch and Welter-Nicol (2014), investigated how involving the levels of loan to value (LTV) and debt service to income (DSTI) affect the quality of new loans. In turn, Uluc and Wieladeck (2015) assessed the effect of introducing a countercyclical capital buffer on mortgages based on Credit Registry data. In particular, the authors showed that an increase in capital requirements by 100 basis points leads to a decrease in the average loan volume by 5.4 %.

Konečný and others (2015) describes approaches of Using the Czech Central Credit Register for Financial Stability Purpose. In particular, there are the following ways for using: monitoring the level of defaults, indicators of credit standards, the ratio of non-performing loans to all loans, as well as monitoring the classification of debtors between banks.

Doko F. et al. (2021) presents applying different machine-learning models to create an accurate model for credit risk assessment using the data from the real credit registry dataset of the Central Bank of Republic of North Macedonia. Moreover, they used machine-learning techniques to gain the most optimal model. In particular, they tested the following tools of Machine Learning: logistic regression, support vector machines, random forest, neural network, decision tree and concluded that decision tree is the most efficient in their case.

Recently, there has been a significant improvement in credit scoring and credit risk modeling. Most progress was observed in the sector

of peer-to-peer lending. Klimowicz A. et al. (2021) described the method for choosing the optimal cut-off point for credit scorecards with an application of Machine Learning for that sector.

Kaminsky (2012) presented the overview of applying scoring tools in credit risk-management.

Biecek P. et al. (2021) observed that rapid improvements give the opportunity to build a much more accurate model even on a 5-year horizon. However, the authors also noted that Machine Learning and artificial intelligence methods are challenges for micro-prudential monitoring by regulators. At the same time, according to a recent publication of the Bank of England, two-thirds of financial services in the United Kingdom use Machine Learning. In those dynamic circumstances, regulators should develop more advanced monitoring tools to estimate the actual situation on the market.

Research goals and questions. Our paper aims to develop the stable and admissible credit assessment model of individual borrowers for monitoring purposes based on the Credit Registry data. There are two research questions, such as 1) estimate the admissibility of using Credit Registry data for adequate and comprehensive monitoring; 2) construct the best-fit scoring model for effective monitoring of Credit Registry.

Main findings

1. Data

In our paper, we used the data from the Credit Registry of the National Bank of Ukraine. We chose to use the data of credits issued after 01.01.2017 due to a significant structural change in provision policy at the end of 2016.

We consider using only a shortlist of the essential variables from the Credit Registry for our paper's purpose (Table 1). In our following research we

plan to apply different data mining procedures and to explore more information from Credit Registry.

The initial data consists of 14 indicators. Based on the benchmark variables, we create the list of variables for modeling purposes. The logic behind that is to generate factor or level variables that predict the default with high accuracy.

First, we derived our primary variable: Default. It is a dummy dependent variable, where 0 is the designated absence of default, and 1 is the situation with default. Based on the number of past-due days of the debt indicator, we assign 1 to observations with over 90 days of overdue and 0 for other classes.

Second, we derived the "maturity" of the loan variable. Using contract expiration date and date of issue, we found the number of days between these dates. The maturity variable consists of a number of days in the contract.

Third, we derived the "overdue of interest payment ratio" and "overdue of debt ratio". In nominator of ratios, we used the level of overdue of interest payment and overdue of debt, respectively; however, we used the level of credit risk for the denominator of the ratio due to the complexity of that indicator.

We created a sort of dummies in the next step: existing unproved income and currency denomination. Dummy of "unproved income" is 1 where unproved income exists, 0 – otherwise. The Dummy of "currency of loan" is 1 where it is a national currency (UAH), 0 – otherwise.

The final list of indicators for testing is in Table 2.

1.1. Data cleaning

Data cleaning is one of the most accountable steps. We should consider several issues of data cleaning: 1) typo or technical errors in data; 2) inconsistency of data; 3) outliers.

Firstly, we observed each variable for some potential typo or technical errors. We have analyzed

Table 1. List of indicators from Credit Registry

Indicator	Level of measurement
Currency of loan	Code of currency
Date of issue	Date
Contract expiration date	Date
Credit risk	UAH
Financial class of borrower	Code of class (from 1 to 5)
Corrected class of borrower	Code of class (from 1 to 5)
Value of debt	UAH
Overdue of debt	UAH
Overdue of interest payment	UAH
Number of past-due days of debt	Number of days
Number of past-due days of interest payment	Number of days
Proved income	UAH
Unproved income	UAH
Interest rate	Rate

Source: Credit Registry of the National Bank of Ukraine

Table 2. List of indicators for testing

Indicator	Level of measurement	Code name
Dummy of currency	Dummy (0 -1)	Dummy_UAH_not_UAH
Default	Dummy (0 -1)	Default
Maturity	Days	Maturity
Interest rate	Rate	Interest_rate
Credit risk	UAH	Credit_risk
Financial class of borrower	Code of class (from 1 to 5)	Fin_class
Corrected class of borrower	Code of class (from 1 to 5)	Corrected_class
Overdue of debt ratio	Ratio	Overdue_debt_ratio
Overdue of interest payment ratio	Ratio	Overdue_interest_ratio
Number of past-due days of debt	Number of days	Overdue_days_debt
Number of past-due days of interest payment	Number of days	Overdue_days_interest
Proved income	UAH	Proved_income
Unproved income	UAH	Unproved_income
Dummy of existing of unproved income	Dummy (0-1)	Dummy_Unproved_income

each suspicious case and we reduced observation with obvious errors. For instance, there could not be a real interest rate – 500 percent point.

For the next step, we see the consistency of data in case of strong outliers. The borrower with a UAH 1 billion loan amount and only UAH 10,000 proved income is a candidate for dropping.

Thirdly, we see the distribution of indicators for cutting rate optimization. Our target is to have unbiased estimation with no effect of strong outliers. For this purpose, by the rule of thumb, we drop 5 % of observations with the highest level of Credit risk, Proved income, and Unproved income. We also tested 1 % and 10 % of the cutting edge, however with 10 %, we drop a significant portion of regular values, and with 1 %, we have not cut all radical outliers in the sample.

Finally, we keep observation with no “N/A”. It reduces the 65 % of observation; however, it is the most straightforward way with stable results. In the following paper, we plan to make a more advanced assumption about that issue.

2. Methodology

2.1. Optimal model specification

2.1.1. Correlation between indicators

According to Credit Registry data, we have several indicators that we explored in the previous section. However, due to potential issues of multicollinearity and endogeneity we could not include all indicators in the model.

To start with, we explored the interdependency between indicators using the simple correlation between them. In Figure 1 below, we can see the correlation plot of indicators, where red color means perfect positive correlation (~100 %) and blue color for a perfect negative correlation.

First, we observed that there is expected multicollinearity between a number of past-due days of debt and the number of past-due days of interest payment. Moreover, these variables have a direct impact on the Default variable by rule. A high correlation between Default and noted variables is practical evidence of that. Therefore, there is an endogenous issue. We should exclude both indicators from our shortlist.

Second, there is perfect multicollinearity between the overdue debt ratio and overdue interest payment ratio. There could also be an endogenous issue of these variables due to strong indirect relation to Default. We chose to exclude both of them.

Third, we also excluded the financial class and corrected class of borrower due to endogenous issues. There are five classes for borrowers, where the fifth class means Default. High correlation is additional proof of that.

It is important to note that we do not observe multicollinearity between the values of the borrower’s unproved income and a dummy of existing an unproved income. We invigilate medium correlation between variables; therefore, theoretically, it could be in one specification simultaneously.

Besides that, there are other insights from the correlation plot. In particular, an inverse correlation between proved income and unproved income. The more proved income associates with less unproved income and vice versa. The reason for that could be in a significant share of the shadow market, where salary is not official and without taxes.

2.1.2. Information value (IV)

For the next step of selection, we used the Information value of indicators. The concept of information value shows the predictive power of a

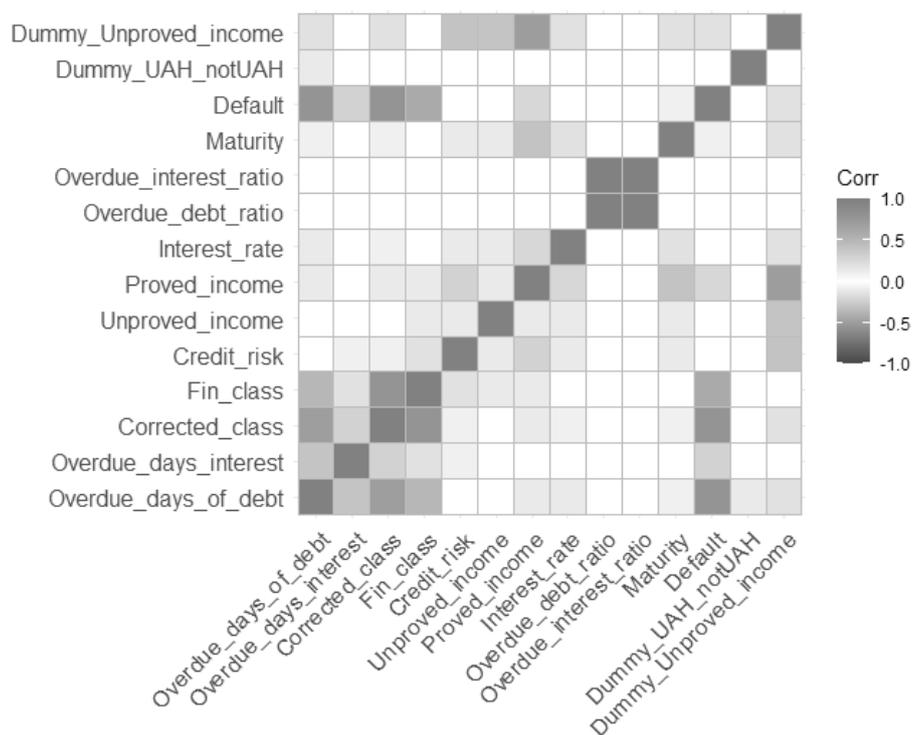


Fig. 1. Correlation plot of indicators

particular indicator, where we have “Good” observations (Non-Default) and “Bad” observations (Default).

$$IV = \sum(DistributionGood_i - DistributionBad_i) \times \ln\left(\frac{DistributionGood_i}{DistributionBad_i}\right). \quad (1)$$

The rule described by Siddiqi N. (2012) leads upon Table 3.

Table 3. Information value and estimated predictive power of indicator

Information Value	Predictive Power
< 0.02	Useless for prediction
0.02 to 0.1	Weak predictor
0.1 to 0.3	Medium predictor
>0.3	Strong predictor

Using this concept, we estimated the information values of the indicators:

Table 4. Information value of indicators with Default as response variable

№	Indicators	Information value
1	Maturity	2.769
2	Interest rate	1.752
3	Unproved income	1.733
4	Proved income	0.799
5	Dummy of existing of unproved income	0.787
6	Credit risk	0.595
7	Dummy of currency	0.002

We observed satisfactory results with a very high IV level for Maturity and relatively low IV for Dummy of currency. At this step, we chose to keep this variable in our list; however, we consider this peculiarity in future phases.

2.1.3. Stepwise selection

The next step of the selection procedure is stepwise selection. There are forward stepwise selection and backward elimination selection. There is adding a new variable in the model and estimation of the efficiency in each specification in the forward approach. In the backward elimination approach, there is a dropping out of variables one by one.

We used bidirectional elimination that compose both methods. There is logistic regression as a method in our stepwise selection and the Default variable as a dependent variable.

To differentiate models’ efficiency, we used Akaike Information Criterion (AIC), which indicates a better-fit model.

The formula is next:

$$AIC = 2K - 2\ln(L), \quad (2)$$

where K – number of independent variables; L – log-likelihood estimate.

According to stepwise selection, we chose first specification, which we call Model 1:

$$Default \sim Dummy_Unproved_income + Maturity + Interest_rate + Proved_income + Unproved_income + Credit_risk$$

According to the stepwise selection, we excluded the variable *Dummy_UAH_notUAH*. This indicator also had a lower information value. We could not conclude that this factor is insignificant; the reason for low explanatory power could be bias due to cleaning of data or some structural peculiarities of that sub-sample.

The result of logistic regression is on Table 5.

Table 5. Logistic regression with specification Model 1

Predictors	Default	
	Log-Odds	p
(Intercept)	-4.04763590	<0.001
Dummy_Unproved_income	1.55337086	<0.001
Maturity	0.00022053	<0.001
Interest_rate	-0.02799737	<0.001
Proved_income	-0.00000034	<0.001
Unproved_income	-0.00000109	<0.001
Credit_risk	0.00000032	<0.001
R ² Tjur	0.094	

Table 6. Specification with quadratic form of indicators

Quadratic form of indicator	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII
Maturity	***				-	-	-	***	***				
Interest rate		***			***	***	***			***	***	***	
Proved income			***		***			***		***		***	***
Credit risk				**	***	***			**	***	***		***
AIC	2891	2587	2871	2887	2553	2578	2589	2868	2883	2551	2576	2566	2859

Note: “***” – significance at 0.001, “**” – significance at 0.01, “*” – significance at 0.05, “.” – significance at 0.1, “-” – not significant.

We observed that all variables are significant. *Dummy* of existing unproved income has a positive dependency to Default. Therefore, if the borrower assigns unproved income with more probability, this borrower will default. Maturity has a positive association with Default, as expected. However, interesting is the relation between proved income and unproved income. One additional hryvnia of unproved income decreases the probability of default more than one proved hryvnia. However, we should remember that some effect of unproved income is related to the dummy of existing unproved income, which is positive.

The interest rate has a negative relation to Default; with a higher Interest rate, the probability of default decreases. It is not intuitive; therefore, our Model 1 has some systemic weaknesses.

2.1.4. Tuning of model

Before running Model 1, we did some data cleaning and dropped out of variable procedures. Some of the steps we did manually, some of them were automatic. However, in Model 1, we assume only linear relatedness between factors and

Default. In fact, there is much-complicated interconnectedness.

In the tuning procedure, we test the quadratic form of each variable to find the best specification with the available set of variables.

Before that, we excluded unproved income variables due to possible multicollinearity with *Dummy* of existing unproved income and strange result as a consequence in Model 1 estimation.

To test the best specification, we added a quadratic form for each variable and paid attention to 1) the significance of variables, 2) the AIC of the model.

As a result, we chose the tuning version of Model 1 with a quadratic form of Interest rate, Proved income, and Credit Risk.

In this specification, the model reflects most of our intuitive expectations. In particular, a lower Interest rate does not mean better conditions to back credit. The low interest rate could be the reason for overheating

the loan market, where the approval thresholds for the borrower are much lower than in standard time.

Hence, our conceptual Model 2 will have the next specification:

$$\text{Default} \sim \text{Dummy_Unproved_income} + \text{Maturity} + \\ + \text{Interest_rate}^2 + \text{Interest_rate} + \text{Proved_income}^2 + \\ + \text{Proved_income} + \text{Credit_risk}^2$$

The result of logistic regression of this model is on Table 7.

Table 7. Specification with quadratic form of indicators

Predictors	Default	
	Log-Odds	p
(Intercept)	-3.28780846396040	<0.001
Dummy_Unproved_income	1.29915121612336	<0.001
Maturity	0.00028699451338	<0.001
Interest_rate^2	0.00489521222754	<0.001
Interest_rate	-0.23026883780630	<0.001
Proved_income^2	0.00000000000004	<0.001
Proved_income	-0.00000078716138	<0.001
Credit_risk^2	-0.00000000000004	<0.001
Credit_risk	0.00000081522633	<0.001
R ² Tjur	0.186	

Compared to Model 1, we observed two times higher R^2 . All variables are significant. However, the quadratic relation of explanatory variables is much more important. We observed the hump-shaped link of Credit risk. There is the point in which the probability of default is the highest; however, after that point increasing credit risk associates with decreasing the probability of default. For interest rate and proved income, we observed convex form relation, which is opposite to hump-shape. It means increasing interest rate and proved income associated with decreasing the probability of default, however only for critical points. After the critical point, increasing of both indicators associates with the increasing risk.

The relation of proved income with the probability of default is not linear, and relatively high income could be more risky than the average one. The reasons could be such as 1) fake information about proved income, 2) borrowers connected to the bank, or 3) weaknesses of the judgment system. For policymakers, there could be powerful insight.

2.2. Alternative model: Machine-learning model

In these latter days, there is a significant improvement in the credit risk assessment. According to the Bank of England, two-thirds of financial services in the United Kingdom use Machine Learning somehow (Bank of England, 2019). The main drivers of that are much more data and technical improvement of econometrics methods. Therefore, using only the Logit model for estimation of the credit risk of borrowers is insufficient.

We tested the Machine Learning approach. There are several methods, which are relevant in our case. In particular, Gradient Boosting, Extreme gradient boosted decision trees, k Nearest Neighbours, Support vector machine, Neural Network, Naive Bayes,

Decision Tree, Random Forest, and Latent Dirichlet allocation. These methods are the most applicable for Credit risk assessment with the binary dependent variable. In our subsequent work, we plan to make our testing to choose the best fit model for our data; however, in this paper, we decided Extreme gradient boosted decision trees (XGBtrees) for several reasons. First, according to Beeravalli V. (2018), the XGBtrees is one of the most balanced methods relatively. It has good accuracy, medium sensitivity, medium specificity, and well-balanced accuracy. Second, the all-upward mention list of methods has similar efficiency results due to Beeravalli V. (2018).

The method of XGBtrees is described in the paper of Chen T. (2016). We use standard parameters of it.

To test the efficiency of the method, we split data into training and testing sub-samples. There are no direct rules for the ratio of data split into training sets; however, standard practices use 70 %/30 % or 67 %/33 % (Brownlee, 2019). We chose the 67 %/33 % rule.

We would test Model 1 and Model 2 specifications with the XGBtrees method.

2.3. Comparison of models and methods

To compare models and methods, we use several metrics of efficiency. R^2 shows the goodness-of-fit, Area under Receiver Operating Characteristic (AUROC) curve shows the tradeoff between specificity and sensitivity of operators, Root Mean Square Error measures the average distance from the predicted point and actual point, F1 score shows the balance between the precision and recall. There are the most common metrics in binary classification models.

We use training sub-sample to estimate the parameters and testing sub-sample to see the out-of-sample effectiveness.

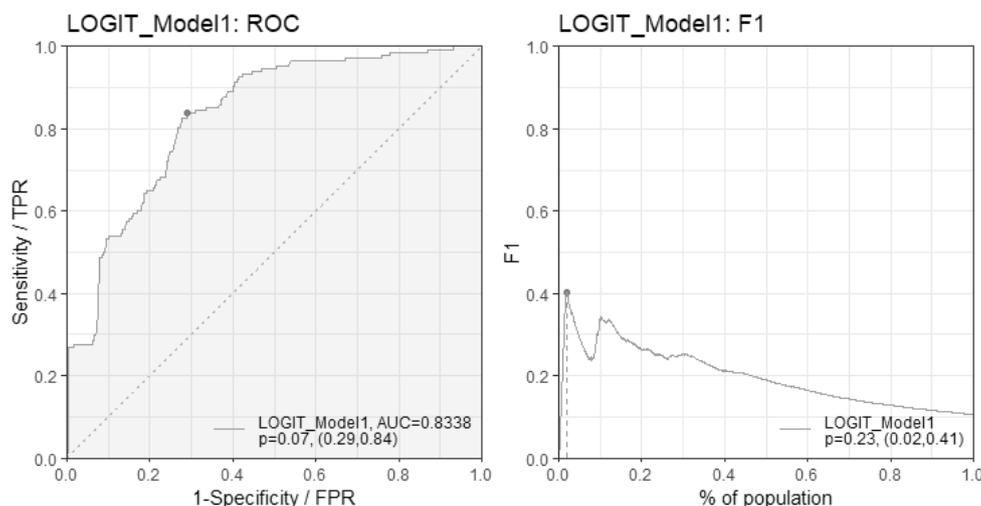


Fig. 2. Efficiency of Logit with specification of Model 1

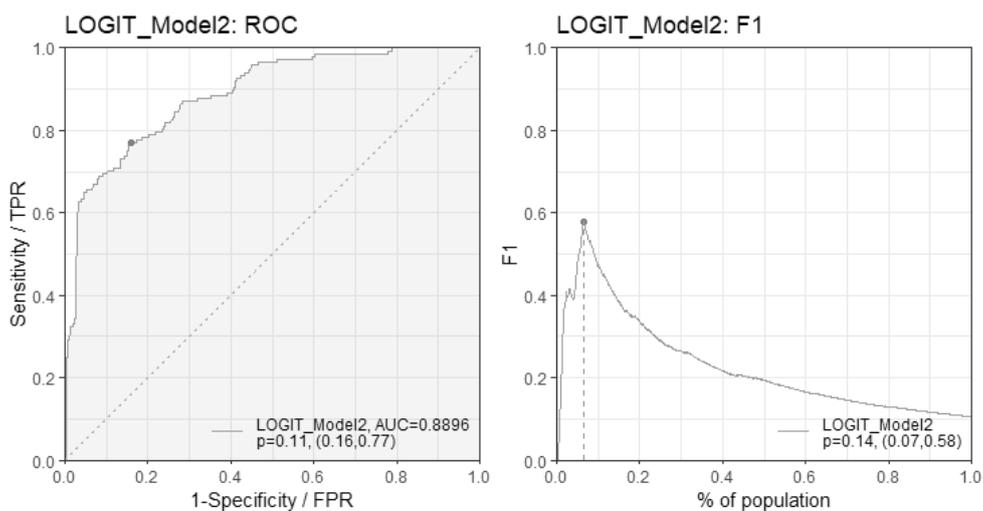


Fig. 3. Efficiency of Logit with specification of Model 2

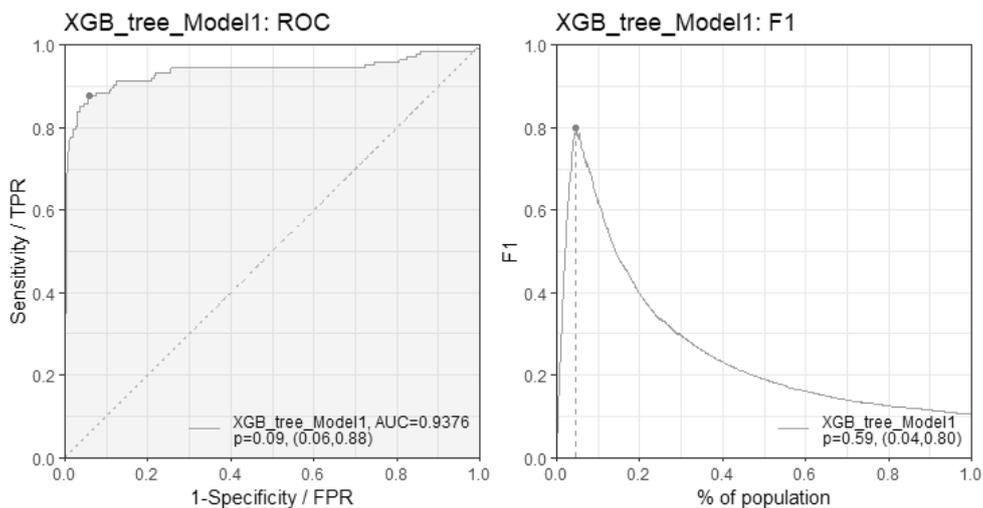


Fig. 4. Efficiency of XGBtree with specification of Model 1

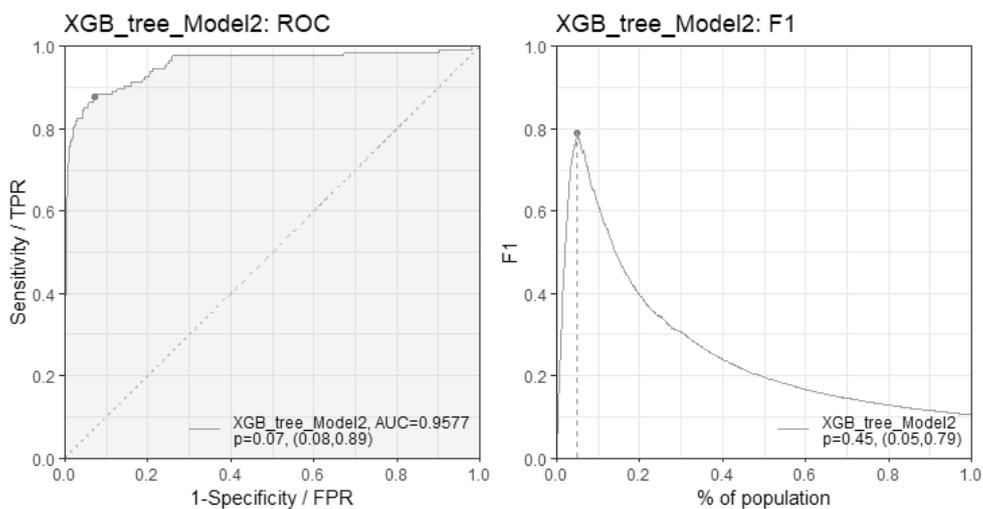


Fig. 5. Efficiency of XGBtree with specification of Model 2

The AUROC and F1 scores are acceptable. We can conclude that this specification and model could be used to estimate default probability; however, we should consider potential bias.

Nevertheless, the significant plus of this model is straightforward interpretation, in particular for policymakers.

Compared to Fig. 2, Model 2 is more effective in the prediction of default according to both metrics AUROC and F1 score. Moreover, the F1 score is one and a half higher than in the previous case.

Surprisingly, using the XGBtree method, we observed much higher efficiency compared to the Logit method. AUROC is around 0.94, which means too significant classification power, while the F1 score is 0.8, which means the same. Moreover, the F1 score is two times more than in the Logit method, and the graph of this metric is more stable than in the Logit approach (Fig. 2).

The apparent weakness of this approach is the inability to interpret and decomposition. On the one hand, it could work in aggregate format for simple monitoring. On the other hand, it could be insufficient for deep policy analysis.

XGBtree method with the specification of Model 2 gives us the best result according to the AUROC metric. However, the F1 score is only close to the result of Model 1 with the XGBtree approach.

Table 8. Comparison of methods

Method	Model	RMSE	R ²	AUROC	F1
Logit	Model 1	0.216	0.116	0.833	0.41
	Model 2	0.202	0.227	0.89	0.58
XGBtree	Model 1	0.14	0.627	0.938	0.8
	Model 2	0.141	0.623	0.958	0.79

3. Results. Using both methods and models, we made the distribution of alternative scorecards with a range from 0 to 1 where a higher value means more risks.

We predict the dependent variable in each estimation for each observation using estimation parameters for creating the scorecard.

As a result, we get some value from $-\infty$ to $+\infty$. We have normalized these values to the 0-1 range for better visual interpretation using the MINMAX approach.

The formula of MINMAX provides “natural” normalization:

$$\frac{(X_{i,t} - MIN_i)}{(MAX_i - MIN_i)} \quad (3)$$

where MIN_i – minimum value of estimation i ; MAX_i – maximum value of estimation i .

This form of presentation is more sophisticated for policymakers and stakeholders.

It is important to note that we do not show the probability of default distribution; it is only normalized values of prediction. It is a quasi-scorecard.

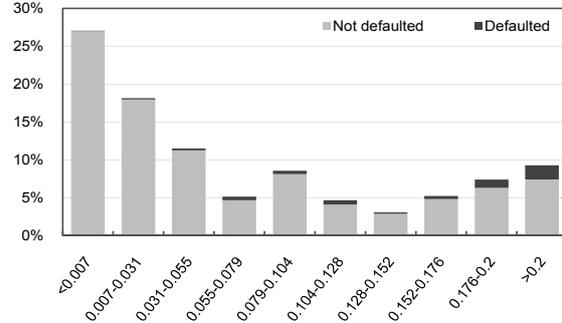


Fig. 6. Distribution of Defaulted and Not Distribution of Defaulted and Not defaulted credits of Model 1 by XGBtree

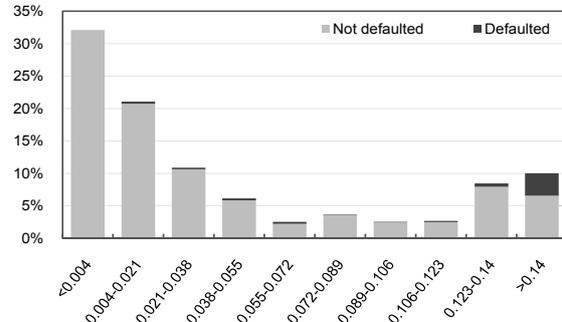


Fig. 7. Distribution of Defaulted and Not defaulted credits of Model 2 by XGBtree

- On these figures, we observed several points:
- The distribution in both models is not monotonic; there are some hikes in middle bins.
 - In Model 1, there are many defaults in the middle group of bins. It means that the classification feature has some weaknesses.
 - In contrast, most of the defaults in Model 2 are in the last bin. However, we observed that the ratio of defaults is higher in bin (0.055, 0.072] than in (0.089, 0.106] or (0.106, 0.123]. There is no monotonic distribution of defaults.
- On these figures, we observed acceptable results:
- There is the monotonic distribution of credit across different bins.
 - There is no significant visual difference between Model 1 and Model 2.
 - Most of the defaults are in the last bin in both models.

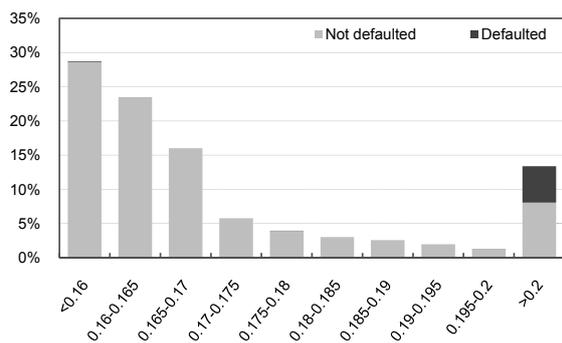


Fig. 8. Distribution of Defaulted and Not defaulted credits of Model 1 by XGBtree

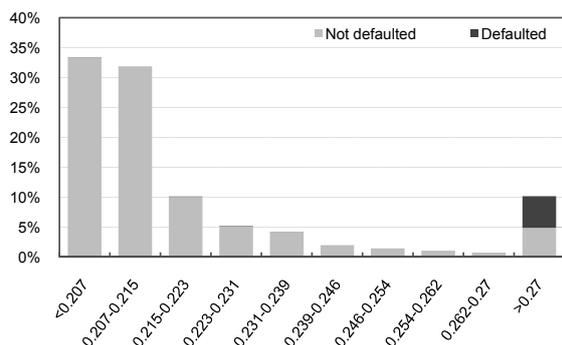


Fig. 9. Distribution of Defaulted and Not defaulted credits of Model 2 by XGBtree

– The distribution of scores is a two-tailed form, with a relatively low number of borrowers in the middle group of bins.

Finally, according to the comments mentioned above, we chose the XGBtree method with the specification of Model 1 for monitoring purposes. In that case, we have the most balanced results.

Conclusions and further research proposals. In this paper, we explored the Credit Registry data and the ability of that data for monitoring purposes of the Central Bank.

At first, we showed that this data is a good source for monitoring purposes and could be used by the regulator on an ongoing basis. We have used only a partial list of potential variables in our estimation; however, it is possible to expand the list and develop a more advanced model in our following paper.

Second, we tested the Logit method and Machine Learning method to estimate the difference in effectiveness for our purposes. According to monitoring

purposes, we have the following recommendations. In a deep analysis of factor dependency, we recommend using simple Logit regression with some specification tuning. For instance, it could be helpful in DSTI, DTI calibration, where the income factor effect played a key role. At the same time, in the case of systemic risk accumulation monitoring, the Machine Learning method could be much more efficient and valuable. For instance, increasing the ratio of borrowers in middle bins across time could signal a change in the bank’s risky behavior.

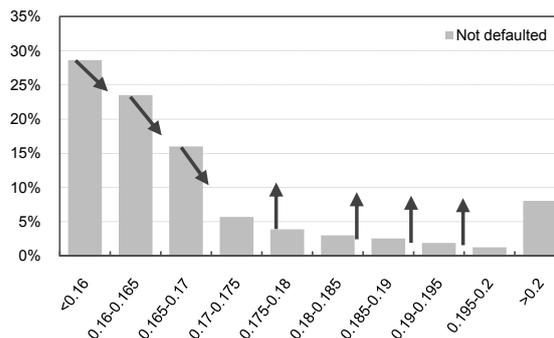


Fig. 10. The hypothetical pattern of risky dynamic

Finally, we also made fewer valuable conclusions for policymakers, however more interesting for credit risk modeling based on Credit Registry data.

In particular, we found that tuning of the model is crucial for the Logit model. The relatedness between indicators is not monotonic and linear in most cases; therefore, we need to test at least a quadratic form of variables. Nevertheless, there is not an essential step for Machine Learning models.

Surprisingly, even the basic list of variables from the Credit Registry and basic models could significantly predict default. It is also a signal for banks to use Credit Registry data more intensively.

There are also some insights for policymakers. The factors of income, maturity, and interest rate are significant. Moreover, the dependency with Default is not linear. However, such factors as the currency of credit or unproved income should be the subject of the following research due to the non-obviousness of their effect.

Scoring based monitoring of the Central Bank Credit Register can be efficiently applied to regulators for developing adaptive policies.

References

Bacham, D., & Zhao, J. (2017). Machine Learning: challenges, lessons, and opportunities in credit risk modeling. *Moody’s Analytics Risk Perspectives*, 9, 30–35.

Bank of England, Financial Conduct Authority. (2019). Machine learning in UK financial services.

Beeravalli, V. (2018). Comparison of Machine Learning Classification Models for Credit Card Default Data. <https://medium.com/@vijaya.beeravalli/comparison-of-machine-learning-classification-models-for-credit-card-default-data-c3cf805c9a5a>.

Bennardo, A., Pagano, M., & Piccolo, S. (2015). Multiple bank lending, creditor rights, and information sharing. *Review of Finance*, 19, 519–570.

Berger, A. N., Scott Frame, W., & Ioannidou, V. (2011). Tests of ex ante versus ex post theories of collateral using private and public information. *Journal of Financial Economics*, 100(1), 85–97.

Biecek, P., Chlebus, M., Gajda, J., Gosiewska, A., Kozak, A., Ogonowski, D., Sztachelski, J., & Wojewnik, P. (2021). Enabling Machine Learning Algorithms for Credit Scoring--Explainable

- Artificial Intelligence (XAI) methods for clear understanding complex predictive models. *arXiv preprint arXiv:2104.06735*.
- Bos, J. W., De Haas, R., & Millone, M. (2016). Show me yours and I'll show you mine: sharing borrower information in a competitive credit market. *BAFFI CAREFIN Centre Research Paper*, 2015-8.
- Brownlee, J. (2019). *XGBoost With Python*. Machine Learning Mastery.
- Dietsch, M., & Welter-Nicol, C. (2014). Do LTV and DSTI caps make banks more resilient? *Débats économiques et financiers*, 13, Banque de France ACPR.
- Doko, F., Kalajdziski, S., & Mishkovski, I. (2021). Credit Risk Model Based on Central Bank Credit Registry Data. *Journal of Risk and Financial Management*, 14(3), 138. <https://doi.org/10.3390/jrfm14030138>
- Gareth, J., Witten, D., Hastie, T., & Tibshirani, R. (2014). *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated.
- Hurvich, C. M., & Tsai, C. L. (1990). The impact of model selection on inference in linear regression. *American Statistician*, 44, 214–217.
- Kaminsky, A. (2013). Genesis and Structure of Credit Bureaus System in Ukraine. *Bulletin of Taras Shevchenko National University of Kyiv. Economics*, 9(151), 31–34. <http://dx.doi.org/10.17721/1728-2667.2013/151-10/7> [in Ukrainian].
- Kaminsky, A. (2015). Credit bureau benchmarking as a tool for estimation of a bank's position at the market. *Bulletin of Taras Shevchenko National University of Kyiv. Economics*, 1(166), 60–64. <http://dx.doi.org/10.17721/1728-2667.2015/166-1/8>
- Kaminsky, A., & Pysanets, K. (2012). Scoring Technologies in Credit Risk-Management. *Business Inform*, 4, 197–201. https://www.business-inform.net/export_pdf/business-inform-2012-4_0-pages-197_201.pdf [in Ukrainian].
- Klimowicz, A., & Spizewski, K. (2021). Concept of peer-to-peer lending and application of machine learning in credit scoring (No. 2021-04).
- Konečný, T., Plašil, M., Rusnák, M., & Řežábek, P. (2015). Use of the Czech Central Credit Register for Financial Stability Purposes, Czech National Bank Financial Stability Report 2014/2015.
- Law of Ukraine No. 2277-VIII. (2018). On Amendments to Certain Laws of Ukraine on Establishing and Maintaining of the Credit Register of the National Bank of Ukraine and Improving Credit Risk Management of Banks. <https://zakon.rada.gov.ua/laws/show/2277-19#Text>
- National Bank of Ukraine. (2021). Loan Portfolio Quality (NPLs). *Dashboard of Financial Stability Indicators*. <https://bank.gov.ua/en/stability/npl>
- Siddiqi, N. (2012). *Credit risk scorecards: developing and implementing intelligent credit scoring* (Vol. 3). John Wiley & Sons.
- Uluc, A., & Wieladek, T. (2015). Capital requirements, risk shifting and the mortgage market. *Bank of England working papers* 572.

Філатов В. Ю., Камінський А. Б.

ЗАСТОСУВАННЯ СКОРИНГУ ДЛЯ МОНІТОРИНГОВОЇ ФУНКЦІЇ КРЕДИТНОГО РЕЄСТРУ ЦЕНТРАЛЬНОГО БАНКУ

У 2018 р. було створено Кредитний реєстр Національного банку України, базовими функціями якого є моніторинг кредитної системи та обмін кредитною інформацією між банками. Основна мета роботи полягала в обґрунтуванні скорингового підходу до реалізації моніторингової функції та побудови скорингу на основі даних реєстру. Цей підхід у статті розглянуто в сегменті фізичних осіб, інформація про кредитування яких передається до реєстру.

Першим кроком для досягнення мети дослідження було застосовано процедури Data Mining. Побудовано статистично значущу вибірку, до якої застосовано процедури «очищення» даних щодо викидів та технічних помилок. Другим кроком ідентифіковано бінарну змінну «дефолт» та сукупність характеристик, які можуть бути використані під час побудови скорингу. Третім кроком проведено комплексний аналіз пояснювальних характеристик методами кореляційного аналізу та поетапної селекції (stepwise selection). Проведено тестування нелінійного зв'язку між характеристиками та здійснено оцінку їх інформаційної значущості. Сформовано сукупність значущих характеристик, до якої увійшли такі характеристики, як строк кредиту, відсоткова ставка, підтверджений і непідтверджений дохід тощо. На цій основі здійснено побудову скорингів за допомогою різних методів. Зокрема, застосовано методи логістичної регресії та машинного навчання. Аналіз результатів показав, що застосування методу Extreme Gradient Boosting Trees дає кращі результати скорингового оцінювання кредитів. Зокрема, спостерігаються менші значення середнього квадратичного відхилення та більші значення коефіцієнта детермінації.

Незважаючи на вищу ефективність одного з методів, автори рекомендують використовувати обидва методи для більш повного і ґрунтовного аналізу досліджуваної проблеми. З допомогою методу логістичної регресії можна оцінити залежність між залежною змінною та характеристиками, доступними в реєстрі. З іншого боку, метод Extreme Gradient Boosting Trees ефективніший для побудови саме скорингової моделі, яка пропонується в роботі для здійснення функції інтегрального моніторингу.

Результатами проведеного дослідження є: 1) оцінка значущості даних Кредитного реєстру для здійснення макропруденційного моніторингу; 2) побудований скоринг інтегральної оцінки якості кредитів, ефективність якого визначено компаративним аналізом при застосуванні різних методів побудови.

Ключові слова: кредитний ризик, кредитний реєстр, центральний банк, скоринг.



Матеріал надійшов 30.04.2021