IMPACT OF MACRO AND MICRO FACTORS ON PROVISION FOR CREDIT RISKS OF COMMERCIAL BANKS IN VIETNAM: APPROACH ON PYTHON PROGRAMMING PLATFORM

Hoang Thi Thanh Hang¹, Nguyen Thi Kim Phung², Tran Trong Huy³, Lê Thị Phương Liên⁴

¹²³⁴ Ho Chi Minh University of Banking, Vietnam

Abstract

This study uses LGBMRegressor (Light Gradient Boosting Machine Regressor) algorithm in machine learning on python platform along with SHAP (Shapley Additive exPlanation) technique to extract information from machine learning model to evaluate the macro and micro factors affecting the provision for credit risks at commercial banks in Vietnam. Data was collected from 30 commercial banks in Vietnam from 2008 to 2020. Research results show that profitability, size, bad debt, credit balance, capital adequacy ratio, economic growth and unemployment rate have an impact on the provision for credit risk. From there, the article proposes some policy implications to limit credit risk at commercial banks in Vietnam.

Keywords

Commercial banks, LGBMRegressor, Provision for credit risks, Risk-taking

1. Introduction

Commercial banks are one of the financial intermediaries performing the important function of providing capital to the economy by attracting idle money in the society for individuals and organizations to borrow. Thereby, they contribute to promoting the development of industries and fields in a country's economy.

In Vietnam, over a long period of operation, the commercial banking system has grown tremendously in terms of quantity, quality and scale. The business activities of the commercial banks are increasingly rich and diversified. In which, credits are the assets that account for the largest proportion and also provide the main source of income for the bank. However, this property also brings significant risks to the bank when customers fail to fulfill their commitments. Credit risk always causes losses to commercial banks (Nguyen Van Thuan & Duong Hong Ngoc, 2015).

According to Corsetti et al (1998), one of the most important causes of the 1997 Asian financial crisis was the high rate of overdue debt at commercial banks. Before the crisis, the rate of overdue debt at commercial banks in Thailand was 13%, Indonesia was 13%, Philippines was 14% and Malaysia was 10%. And credit risks once again caused the global monetary and financial crisis in 2007-2009, with the starting point being the collapse of the US financial system. According to the announcement of the US Federal Reserve, in 2008 there were a total of 26 banks that failed, but in 2009, this number rose to 140 with a series of bankruptcies of financial institutions. These financial institutions have a long history and the most financial potential in the world.

In Vietnam, after the period 2011-2012, was a period of bad debt crisis, the establishment of VAMC Asset Management Company was issued by the Prime Minister on May 31, 2013 on the basis of Decision 843/2013/QĐ in order to deal with bad debts in the system of commercial banks has achieved a number of important results such as: eliminating the risk of banking system failure; liquidity was maintained; bad debt on the balance sheet was controlled in stable level, thereby improving the balance sheet, financial capacity as well as the scale of operation of commercial banks, especially weak commercial banks. Although the policies of the State Bank have received positive signals, the consequences and lessons of the bad debt crisis in 2012 are still worth pondering.

While there are many articles abroad that have focused on studying the factors affecting loan provisions (Mazreku et al., 2018; Zheng et al., 2019); Ozili, 2018), there are very few articles on this topic in Vietnam. Research by Nguyen Thi Thu Hien & Pham Dinh Tuan (2014), Nguyen Van Thuan & Duong Hong Ngoc (2015)
encountered limitations in data size (only 5-year research period from 2008 to 2013) and research methods. Similarly, the study of Nguyen Thi Kim Phung & Nguyen Thi Nhat Tan (2022) also has limitations on data processing method.

Recognizing the importance of credit risk control and in order to overcome the limitations of previous studies, the authors have studied the topic "Impact of macro and micro factors on provision for credit risks of commercial banks in Vietnam: Approach on python programming platform". Thereby, the study proposes some policy implications to effectively manage credit risk at commercial banks in Vietnam.

2. Theoretical Basis and Empirical Evidence

2.1. Theoretical basis of credit risk provision in Vietnam

2.1.1. The concept of credit risk provision

According to Circular 11/2021/TT-NHNN dated July 30, 2021 of the SBV, stipulating the classification of assets, the level of deduction, the method of making provision for risks and the use of provisions to handle risks in the operation of credit institutions and foreign bank branches, credit risk provision is defined as follows: “Risk provision is the amount set aside and accounted for in operating expenses to make provision for possible risk can happen for debts of credit institutions and foreign bank branches.”

Accordingly, risk provisions include two types:

- Specific provision is the amount set aside to provide for possible risks for each specific debt.
- General provision is the amount set up to provide for possible risks that have not been determined when making specific provisions.

2.1.2 Subjects of provisions

According to the provisions of Chapter 1, Article 2 of Circular 11/2021/TT-NHNN dated July 30, 2021 of the State bank of Vietnam, the subjects that must classify debts and make provisions include:

- Credit institutions, including: commercial banks and non-banking credit institutions, except credit institutions under special control, shall comply with the provisions of law on special control of credit institutions;
- Foreign bank branches.

2.1.3. Method of provisions

The basis of provision for credit risk is currently established by commercial banks in Vietnam based on Circular 11/2021/TT-NHNN dated July 30, 2021 of the State Bank. According to regulations, the level of provision for credit losses depends on the specific debt group, details are shown in Table 1 as follows:

<table>
<thead>
<tr>
<th>Debt group</th>
<th>Group’s name</th>
<th>Specific provision ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Qualified debts</td>
<td>0%</td>
</tr>
<tr>
<td>Group 2</td>
<td>Debts needing attention</td>
<td>5%</td>
</tr>
<tr>
<td>Group 3</td>
<td>Debts below standard</td>
<td>20%</td>
</tr>
<tr>
<td>Group 4</td>
<td>Doubtful debts</td>
<td>50%</td>
</tr>
<tr>
<td>Group 5</td>
<td>Debts likely to lose capital</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1. Debt classification and provisions according to Circular 11/2021/TT-NHNN

In table 1, debts belonging to group 3, group 4 and group 5 are considered as bad debts. Provisions to be made include:

Specific provision: Specific provision is calculated on a customer-by-customer basis, including a specific total provision for all debts owed by the customer. In which, the specific amount of provision to be deducted from the customer for the principal balance of the customer’s i-th debt (Ri) is calculated according to the following formula:

\[ Ri = (Ai - Ci) \times r \] (1)

In which: Ai: i-th Principal balance of debt; Ci: Deductible value of security assets, financial lease assets, negotiable instruments, other valuable papers in the discounting and resale of Government bonds (hereinafter collectively referred to as collaterals) of the i-th debt; r: Specific provision rate.
**General provision:** The amount of general provision to be deducted is determined at 0.75% of the total balance of debts from group 1 to group 4, minus a number of amounts as prescribed.

### 2.2. Overview of previous studies

Currently, there have been many studies in the world referring to factors affecting credit risk of commercial banks. Studies are carried out at the level of countries or groups of countries around the world. Some typical studies include Hasan & Wall (2004), Chen (2005), Foos & Weber (2010), Shour (2011), Mongkonkittichai (2012), Soedarmono & ctg (2017). Depending on each study, the authors have shown the influence of micro, macro factors or both on the credit risk provision of commercial banks.

In Vietnam, there have been many studies on credit risk provisions. However, studies that clearly show the factors affecting credit risk reduction in Vietnamese commercial banks are still limited. Typical studies related to this topic include Nguyen Thi Thu Hien & Pham Dinh Tuan (2014) studying the factors affecting credit risk provision in the Vietnamese commercial banking system and Nguyen Van Thuan & Duong Hong Ngoc (2015) analyze factors affecting credit risk of Vietnamese commercial banks. The most recent is the study of Nguyen Thi Kim Phung & Nguyen Thi Nhat Tan (2022) studying the factors affecting the credit risk provision of commercial banks in Vietnam.

In general, up to now, not many studies on the topic of credit risk provisions in Vietnam have been carried out. And most of the studies use traditional estimation methods in econometrics such as OLS, FEM, REM, GMM or DFGLS models. This study, in addition to inheriting previous research results in Vietnam and the world, the authors focus on overcoming limitations from the previous study to clarify the impact of factors related to credit risk provisions of commercial banks in Vietnam. We use LGBMRegressor (Light Gradient Boosting Machine Regressor) algorithm in python-based machine learning along with SHAP (SHapley Additive exPlanation) technique to extract information from machine learning models (techniques to extract general insights from a machine learning model) to evaluate macro and micro factors affecting credit risk provision (RR). On the basis of the obtained data, the system will process to select the estimation method with the most optimal results compared to the remaining methods. Therefore, the results obtained from the study will be more convincing.

### 3. Research Methods

#### 3.1. Research model and data

**3.3.1. Research model**

The research model is inherited by the authors from previous studies. In which, the model includes both micro and macro factors affecting credit risk provision. The research model used is as follows:

\[ LLP = f(LLP_{t-1}, ROA, SIZE, NP, CE, CAR, GDP, UNR) \]  

In which: \( LLP \): Provision for credit risk; \( LLP_{t-1} \): 1-period delay of credit risk provision; \( ROA \): Return on total assets; \( SIZE \): Bank size; \( NP \): NPL ratio; \( CE \): Credit balance to total assets; \( CAR \): Safety factor of minimum capital; \( GDP \): Economic growth; \( UNR \): Unemployment rate

Among the factors affecting credit risk prevention can be divided into two groups:

- **Group 1:** Group of factors representing bank characteristics including return on total assets (ROA), bank size (SIZE), bad debt ratio (NP), credit balance ratio on total assets (CE), minimum capital adequacy (CAR).
- **Group 2:** The group represents macro factors including growth rate (GDP) and unemployment rate (UNR).

#### 3.3.2. Research data

The study collects data of 30 banks in the period from 2008 to 2020 with the aim of improving the quality of the research model through increasing the sample size and observation. The calculation criteria are taken from reliable sources such as financial statements, management reports of commercial banks. Macro data sources are taken from two sources, the General Statistics Office of Vietnam and the World Bank. The research data is shown in Table 2 below:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Name</th>
<th>Definition Measurement</th>
<th>Prior Research</th>
<th>Expectations Sign</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Provision for credit losses</td>
<td>Provision for bad debts divided by total credit balance</td>
<td>Zheng &amp; ctg (2019); Ozili (2018); Nguyen Thi Thu Hien &amp; Pham Dinh Tuan (2014); Nguyen Thi Kim Phung &amp; Nguyen Thi Nhat Tan (2022).</td>
<td></td>
<td>Audited financial statements of commercial banks</td>
</tr>
</tbody>
</table>

28 | Impact of Macro and Micro Factors on Provision for Credit Risks of Commercial Banks: Hoang Thi Thanh Hang et al.
3.2. Research Methods

The research is done on Python 3.6.8 programming language along with libraries and machine learning algorithms. The model is performed through the following steps:

- Step 1: Performing descriptive statistics, clean data
- Step 2: Finding an algorithm that fits the model and data, based on the python programming language
- Step 3: Determining the performance and reliability of the model, regression coefficients, visualize the results.

4. Research Results and Discussion

4.1. Research results

- Descriptive statistics of research data

The research data is summarized in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>lrp</th>
<th>roa</th>
<th>size</th>
<th>np</th>
<th>ce</th>
<th>car</th>
<th>gdp</th>
<th>unr</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>390</td>
<td>390</td>
<td>390</td>
<td>390</td>
<td>390</td>
<td>390</td>
<td>390</td>
<td>390</td>
</tr>
<tr>
<td>mean</td>
<td>0.013</td>
<td>0.016102</td>
<td>32.1396</td>
<td>0.02342</td>
<td>0.55615</td>
<td>0.1419</td>
<td>0.0593</td>
<td>0.0157</td>
</tr>
<tr>
<td>std</td>
<td>0.0055</td>
<td>0.010841</td>
<td>1.306766</td>
<td>0.01638</td>
<td>0.13546</td>
<td>0.07</td>
<td>0.01056</td>
<td>0.00427</td>
</tr>
<tr>
<td>min</td>
<td>0.00944</td>
<td>0.009121</td>
<td>31.13376</td>
<td>0.01399</td>
<td>0.45692</td>
<td>0.10219</td>
<td>0.05422</td>
<td>0.0116</td>
</tr>
<tr>
<td>25%</td>
<td>0.01147</td>
<td>0.014587</td>
<td>32.20562</td>
<td>0.02075</td>
<td>0.57411</td>
<td>0.121</td>
<td>0.0621</td>
<td>0.0174</td>
</tr>
<tr>
<td>75%</td>
<td>0.0152</td>
<td>0.021254</td>
<td>32.95405</td>
<td>0.02729</td>
<td>0.65904</td>
<td>0.15753</td>
<td>0.06679</td>
<td>0.0185</td>
</tr>
<tr>
<td>max</td>
<td>0.03974</td>
<td>0.065965</td>
<td>34.98866</td>
<td>0.14018</td>
<td>0.85168</td>
<td>0.70434</td>
<td>0.0708</td>
<td>0.0239</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics of variables in the research sample

Source: Summary statistics in Python

According to the above results, the credit risk provision of 30 Vietnamese commercial banks has an average value of 0.013, equivalent to a rate of 1.3% with a standard deviation of 0.006 (0.6%). The return on total assets of the group from 2008 to 2020 has an average value of 1.61%, which is the average rate of return. The average bad debt
ratio of the 30 banks in the sample is 2.34%. Volatility in bad debt partly explains the fluctuation of provision for credit risk. The scale of Vietnamese commercial banks tends to increase steadily from 2008 to 2020. The scale factor is calculated based on the natural logarithm of total assets with the average value of 2008 is 30.73 to 33.06 in 2020. This shows that banks have policies to expand scale to increase competitiveness as well as dominate market share. The ratio of outstanding loans in the average of 30 banks from 2008 to 2020 is 55.6%. The capital adequacy ratio represents the risk absorption capacity of commercial banks, at an average of 14.1%/year. Vietnam's GDP volatility from 2008 to 2019 shows stability in the macro economy as the GDP growth rate did not see any sharp decline in the rate during this period. Although the unemployment rate fell sharply in the period from 2009 to 2010 but soon after this rate rebounded and fluctuated around 2%. However, the picture of the economic situation in Vietnam as well as in the world became worse when the Covid-19 epidemic appeared, it is easy to see that Vietnam's GDP dropped from 2019 to 2020 to only 2.91%, while unemployment rate increased sharply from 2018 to 2020 at 2.39%.

- Data balance

Because the data has many outliers and is not balanced among the variables, it has been normalized using the Robust Scaler method, after normalizing the data is more balanced and convenient for building models.

- Correlation coefficient matrix

The correlation coefficient matrix of the variables in the research model is shown in Figure 1. The correlation coefficient matrix gives an overview of the correlation between the variables in the model as well as the degree of correlation. The results of the correlation matrix in Figure 1 show that there is no close relationship between the independent variables in the research model.

- Model results:

On the basis of the database has been cleaned and normalized to apply to suitable algorithms. Some algorithms used such as Multiple linear regression, LassoLarsCV, Ridge, SVM, ExtraTreesRegressor give bad results, model performance is not high (R Square from 33% to 53%). The algorithm that gives good results with high model performance is LGBMRegressor (R_Square = 90.86%, MSE = 0.08917). The results of the LGBMRegressor model are shown in Figure 2 and Figure 3:

![Figure 1. Correlation coefficient matrix](Source: Heat.corr in Python)

![Figure 2. LGBMRegressor](Source: LGBMRegressor in Python)

![Figure 3. Shape values](Source: shape package in python)
The results from the model in Figure 2 and the shap values in Figure 3 show that the variables UNR, CE, NP, NP, ROA, GDP, CAR have a positive effect on provisions for credit risk (LLP), while the SIZE variable has a negative effect on provision for credit risk (LLP).

4.2. Discuss the results

- Return on Investment - ROA

The coefficient of the ROA variable is 0.13, which gives us the conclusion that there is a positive impact of the rate of return on the provision for credit risk. Vietnamese commercial banks with high profitability will tend to increase provisioning for LLP loans. Banks' profitability is high. Banks tend to increase reserve ratios to minimize fluctuations in earnings or to avoid regulatory scrutiny (Koch & Wall, 2000). Suhartono (2013) studied at Indonesian banks, banks with high ROA are also able to accept more risks. Research by Nguyen Thi Thu Hien & Pham Dinh Tuan (2014) also gives similar conclusions about the supply-side impact of ROA on credit risk provision. From the obtained results, the author comes to accept the hypothesis that the return on total assets (ROA) has a positive impact on the provision for credit risk (LLP).

- Bank size - SIZE

The variable SIZE represents the size of commercial banks with a coefficient of -0.3. The conclusion is that there is evidence for the impact of size on LLP and this effect is inverse. This is in contrast to the results of some previous studies. Anandarajan et al (2003) argue that large banks have a higher level of business activity than smaller banks and will keep more provisions to compensate for the increase in their business activity, i.e. There is a positive relationship between bank size and loan reserve. Suhartono (2013) studied at Indonesian banks and concluded that bank size has an impact on increasing credit risk provision in Indonesian commercial banks, the coefficient of scale has a positive sign and is significant at 1%. Abdullah (2014) studies the factors affecting risk tolerance (loan provision) for commercial banks in Malaysia, showing that size has a positive effect on loan provision with significant level coefficient at 1% in GMM regression. Ozili (2018) showed similar results in a study in 19 countries in Africa. Nguyen Thi Thu Hien & Pham Dinh Tuan (2014) showed that bank size has an impact on increasing credit risk provision in Vietnamese commercial banks. Although contrary to the results of previous studies, this result is also acceptable in the context of Vietnam in the period 2008-2020. The larger the size of the bank, the longer the operation time. Along with this problem, the better the experience in risk management and control, the more risks are minimized, including credit risk.

- Bad debt ratio – NP

The research results show that the bad debt ratio is 0.15, which proves that the bad debt ratio has a positive impact on credit risk. Most prior empirical studies have found evidence of a positive effect of bad debt ratio on LLP. Research conducted on European banks by Hasan & Wall (2003) shows that bad debt ratio has a positive effect on loan provisions. Nguyen Thi Thu Hien & Pham Dinh Tuan (2014) research for banks in Vietnam from 2008 to 2012 shows that the results show the positive impact of bad debt ratio on risk provision based on OLS, FEM regression, REM. Nguyen Thi Kim Phung & Nguyen Thi Nhat Tan (2022) also has the same result. Abdullah (2014) results demonstrate the close relationship between bad debt and provision for credit risk, the coefficient of bad debt (NPL) has significance up to 1% in GMM regression. Ozili (2018) concludes that banks with high NPL ratio will make large provision for loans than other banks. The results in this study are no exception. Conclusion Vietnamese commercial banks with high NPL ratio must definitely make a higher provision for credit risk to prevent risks caused by credit activities.

- Credit balance ratio – CE

The ratio of credit outstanding to total assets CE has a coefficient of 0.18. This implies a positive impact of the CE to LLP ratio. The results show that banks with large credit balances on total assets will set up a lot of credit risk. This result is consistent with a few previous studies such as Bouvatier & Lepetit (2008) which indicated that a higher loan-to-asset ratio implies a higher default risk and banks have a higher default risk. on the overall loan portfolio will retain more provisions to offset the risk of the loan portfolio. However, Packer & Zhu (2012) investigated the provisioning activities for loans of 240 banks in 12 Asian economies during the period 2000-2009, which recorded a negative relationship between outstanding credit and loan provisions for banks in China, India and Japan. In Vietnam, the conclusion of Nguyen Thi Thu Hien & Pham Dinh Tuan (2014) shows that the ratio of outstanding loans to total assets has a negative effect on credit risk forecast. The explanation for this negative effect is that Vietnamese banks have diversified their loans well across different asset classes, thereby reducing credit risk and provisioning. Thus, with different research stages gave different results. And the research results in this article of the authors are acceptable as the more Vietnamese commercial banks lend, the higher the credit risk, so the more provision must be made.
**-Capital adequacy ratio — CAR**

The minimum capital adequacy ratio CAR has a GMM regression coefficient of 0.01. This result leads to the conclusion that banks with high CAR tend to make lower provision for credit risk. According to Suhartono (2013), research finds that strong capital has a negative impact on credit risk or loan provision. It means that banks with higher capital have lower credit risk as they make smaller loan loss provisions. The trend is that a well-capitalized bank usually controlled by a good management team has a lower credit risk. On the other hand, banks with lower capital will bear more credit risk because of their willingness to accept high-risk loans in exchange for profits. Research by Zheng et al. (2019) for 22 commercial banks in Pakistan from 2010 to 2017 found evidence of negative impact of capital adequacy ratio on credit risk provision through GMM regression.

**-Economic growth - GDP**

The cyclical relationship between economic growth and credit risk provision (LLP) has been confirmed by many previous studies. The research results show that the GDP coefficient is 0.06, showing that the above view is accepted. Research results are in contrast to the study of Packer & Zhu (2012), Suhartono (2012), Abdullah (2014). This reflects that during the growth period, Vietnamese commercial banks tend to increase credit risk provisions due to negative assessment of each different situation. As the economy grows hotter, the risk of defaulting businesses increases. And loans are prone to falling out of control. Since then, banks have to make higher provisions when classifying customers.

**-Unemployment rate – UNR**

This study has found evidence of a positive effect of UNR unemployment rate on LLP loan provision, the coefficient of unemployment rate is 0.48 with the largest value compared to the rest of the independent variables in the model. Figure. The results are in contrast to some previous studies. Louzis et al.(2010) research in Greece again shows that when the unemployment rate increases, the bad debt ratio decreases. Kumar et al. (2018) find that an increase in unemployment reduces NPLs in South Pacific banks. However, this result implies that in the period when unemployment rate increases and the economy is at risk of instability, banks in Vietnam will increase provision for credit risk to prevent bad situations that could happen. happen.

### 5. Conclusions and Policy Implications

The research results obtained from the results of the LGBMRegressor model have concluded that the factors affecting the credit risk provision of the group of 30 Vietnamese commercial banks in the period from 2008 to 2020 have shown that the rate of return, size, bad debt, credit balance, capital adequacy ratio, economic growth and unemployment rate all have an impact on provision for credit risk. Among the macro factors, the factor that has the most impact on the credit provision is the unemployment rate. Among the micro factors, the factor that has the most impact on the credit provision is the bank scale. From the research results, the authors propose some policy implications to limit credit risk as follows:

- For banks with large total assets, the credit risk provision is less. Banks should note that when access to capital is abundant due to their large scale, credits will increase with this. Banks need to focus on increasing the size of total assets, but also need to strengthen management and strict control of loans to ensure sustainable growth, without adversely affecting banking business results in general and credit risk in particular.

- Firstly, when banks have larger asset size, this result confirms that banks have more risk provisioning capacity than other banks, which is reasonable when banks Large banks often have access to abundant capital and therefore have a higher provision for credit risk than small and medium-sized banks.

- Secondly, the research results show that Vietnamese commercial banks tend to increase their credit risk provisions when the economy has a high unemployment rate, which may also be the result of Vietnamese commercial banks making High-risk loans during this period lead to an increase in the bad rate, thereby affecting the credit risk provision. This conclusion helps the author to recommend that the State bank of Vietnam needs to increase supervision activities as well as use monetary policy tools to control the quality and lending activities of Vietnamese commercial banks in the period of economic changes employment-related bad.

Although the research topic has tried to collect data of 30 banks for a long time from 2008 to 2020, besides updating more factors affecting LLP loan provisions in recent years. However, there are still limitations that are not deeply exploited macro factors affecting credit risk provision and have not collected data of foreign-owned commercial banks in Vietnam. The following studies can suggest this direction to further clarify the problem of credit risk prevention at commercial banks in Vietnam.
References


