วารสารบริหารธุรกิจเทคโนโลยีมหานคร
MUT Journal of Business Administration
ปีที่ 17 ฉบับที่ 2 (กรกฎาคม – ธันวาคม 2563)
Volume 17 Number 2 (July – December 2020)

Predictive Power of Credit Scoring Questions: Case of SMEs Loan Approvals

อำนาจในการพยากรณ์ของคำถามคะแนนเครดิต: กรณีการอนุมัติเงินกู้ SMEs

Received: *September 8, 2020*
Revised: *December 2, 2020*
Accepted: *December 4, 2020*

Nopmanee Parungrojrat นพมณี การุ่งโรจน์รัตน์ [1,*]

[1] Assistant Professor, Department of Management, Faculty of Business Administration, Mahanakorn University of Technology, Bangkok, Thailand
[1] ผู้ช่วยศาสตราจารย์ สาขาวิชาการจัดการ คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีมหานคร กรุงเทพมหานคร ประเทศไทย

ABSTRACT

This research paper aimed to examine predictive power of credit scoring questions being used for screening small and medium enterprises (SMEs) loan applicants. The questions and data of 619 applicants from a non-disclosed leading bank for SMEs in Thailand were explored. Firstly, the exploratory factor analysis was employed. The results showed that from the 12 questions being used in the credit scoring system, there were 5 extracted factors. Secondly, in order to prove predictive power of the questions, the 5 extracted factors were used as predictors or explanatory variables in a Binary Logit model for the non-performing loans (NPLs) classification, i.e. being NPLs or normal debt status. Using step-wise regression technique, 3 factors remained in the model. The 3 factors were "Time and Size", "Debt Records", and "Debt Burden". Consequently, the implication was that for screening SMEs application,

* E-mail address: nopmanee555@gmail.com

credit providers should pay more attention to the questions related to the 3 factors. Meanwhile, as suggested by the model, for the excluded factors, i.e. "Income Reliability" and "Business Prospective", these factors may also still be considered but with comparatively less importance. Although the overall model predictive accuracy was at 69.3%, the percentage was at 31.5% for correctly predicted NPLs cases. Thus, more studies on the credit scoring questions and also more varieties of the models are needed.

Keywords: Predictive Power, Credit Scoring, SMEs, Loan Approval, NPLs

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบอำนาจในการพยากรณ์ของคำถามการให้คะแนนเครดิตที่ใช้ในการคัดกรองผู้ขอสินเชื่อสำหรับวิสาหกิจขนาดกลางและขนาดย่อม (SMEs) โดยใช้คำถามและข้อมูลของผู้สมัคร 619 รายจากธนาคารชั้นนำเพื่อ SMEs ในประเทศไทย ที่ไม่เปิดเผยชื่อ ในขั้นแรกได้ใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ ซึ่งผลลัพธ์แสดงให้เห็นว่า จากข้อคำถาม 12 ข้อที่ใช้ในระบบการให้คะแนนเครดิต มีปัจจัยสกัดได้จำนวน 5 ปัจจัย ในขั้นต่อมา เพื่อพิสูจน์อำนาจในการพยากรณ์ของข้อคำถามที่ใช้ ปัจจัยที่สกัดได้ทั้ง 5 ถูกใช้เป็นตัวทำนายหรือตัวแปรอธิบายในแบบจำลอง Binary Logit สำหรับการจัดประเภทสินเชื่อที่ไม่ก่อให้เกิดรายได้ (NPLs) เช่นเป็น NPLs หรือสถานะหนี้ปกติ จากการใช้เทคนิควิเคราะห์ถดถอยแบบ Stepwise ปัจจัยที่สกัดได้ 3 องค์ประกอบยังคงอยู่ในแบบจำลอง ปัจจัยทั้ง 3 คือ "เวลาและขนาด" "ประวัติการชำระหนี้" และ "ภาระหนี้" ดังนั้นการนำไปใช้ก็คือ ในการคัดกรองผู้สมัครสินเชื่อ SMEs ผู้ให้บริการสินเชื่อ ควรให้ความสำคัญกับข้อคำถามที่เกี่ยวข้องกับ 3 ปัจจัยนี้ ในขณะเดียวกันตามที่แบบจำลองแนะนำ สำหรับปัจจัยที่ได้ตัดออกจากแบบจำลอง ซึ่งก็คือ "ความแน่นอนของรายได้" และ "แนวโน้มธุรกิจ" ปัจจัยเหล่านี้อาจยังคงได้รับการพิจารณา แต่มีความสำคัญน้อยกว่า แม้ว่าความแม่นยำในการทำนายแบบจำลองโดยรวมจะอยู่ที่ 69.3% แต่สำหรับกรณีของการคาดการณ์ NPLs ได้ถูกต้องนั้น อยู่ที่ 31.5% ดังนั้นจึงจำเป็นต้องมีการศึกษาเพิ่มเติมเกี่ยวกับคำถามการให้คะแนนเครดิต และรูปแบบของแบบจำลองที่หลากหลายมากขึ้นในอนาคต

คำสำคัญ: อำนาจในการพยากรณ์, คะแนนเครดิต, ธุรกิจขนาดกลางและขนาดย่อม, การอนุมัติเงินกู้ สินเชื่อที่ไม่ก่อให้เกิดรายได้

Introduction

According to ADB report (Korwatanasakul and Paweenawat (2020)) SMEs businesses are accounted for 45% of Thailand's 2018 GDP, employed 14 million workers, approximately. For the year 2019, the number of firms was roughly 3 million companies and mainly were in service (42%) and trading (40%) sectors. With more varieties of new businesses, new models of entrepreneurship, and quick technology disruption replacing the old- outdated- slow- adjusting firms, the viability of these businesses becomes more questionable from any creditors' point of view. Non-performing loans (NPLs) are costly to any credit providers. Systems and procedures in screening and approving new (or existing customer) loan applications are then vital. Credit scoring is one widely used in the business. With carefully crafted questions, the integrity of the applicants and throughout the line of credit provider's staff, and of course the overall supported business conditions, the scoring process helps creditors in avoiding bad debts. Even with the system forced, for Thailand's case, Bank of Thailand reported gross NPLs was 460,961 Million Baht or 3.14% to total loans, at the end of December 2019. Nowadays, in a present of COVID-19 pandemic, the number is expected to arise. Moreover, the situation of the sampled debt portfolio given, from a non-disclosed leading bank for SMEs in Thailand, was far worst. Out of 619 approved debt applicants, 213 cases turned to be the NPLs. Therefore, the question arises i.e. "What went wrong in the credit scoring system?". This research explores some parts of it, especially the questions being used in the scoring system especially the predictive power of the questions in forecasting (and of course preventing) the NPLs. There are then 12 questions of SMEs loan application, being used for credit scoring in a non- disclosed leading bank for SMEs in Thailand, examined.

Research Questions

From the introductory mentioned, hence, the main research question is that whether the questions being used for current SMEs loan application are suitable for screening potential bad debts. In addition, which of the questions are the vital items for creditors to be carefully considered in loan approval process.

Research Objectives

1. To factorized SMEs loan application questions for dimension reduction and considering potential redundancy of the questions.

2. To explore the predictive power of the factorized questions as predictors of the selected debt classification model-the Binary Logit Model.

Expected Benefits

Results found from this work can be utilized as a tools to improve SMEs loan application. Also, with known predictive power of groups of questions, in preventing future bad debts, the credit scoring system and question weighting assignment can then be adjusted accordingly.

Literature Reviews

Thomas *et al*. (2002) summarizes that credit scoring technique can be considered as a tool for borrower risk assessment. There are two techniques being used i.e. application scoring and behavioral scoring for assessing new applicants and existing lender's customers respectively. Numerical credit scoring was introduced in around 1930s by mail order companies solving inconsistencies among their credit analysts. Shortly after the second World War, statistical models with statistical classification techniques were being used in lending decisions. Nowadays, the techniques being employed range from discriminant analysis, classification trees, nearest-neighbor approach, logistic regression, neuron networks, to the machine learning. Data required for consumer credit rating, for example, are the applicant's age, annual income, home owning status, occupation, time with bank, time with employee, and credit purpose (Hand and Henley, 1997). Meanwhile, for commercial or business credit rating additional data on the business finance and the owners' experience are required.

In developing the statistically-correct credit scoring model, a probability sampling on historic applicant population is needed. However, only the successful applicants' data are available and being used creating the sampling bias. Additionally, using stepwise procedure in entering explanatory variables to the model may be able to avoid multicollinearity problem. Unfortunately, a variable with good predictive ability but highly correlated to an entered variable will not enter the final equation (Capon, 1982). Thus, in this research, factorized applicants' data, using factor analysis technique, even non- orthogonal rotation- allowing some degree of correlation between factored variables but not yet triggering the multicollinearity problem, may solve the problem.

Turning to empirical studies, starting with credit scoring and credit availability, Frame *et al*. (2001) explored model explaining relationship between credit scoring and credit availability for small businesses using data from large U.S. banking organization. It was concluded that, from the structural model used, there was evidence confirming that credit scoring increases credit availability for small businesses. Hence, the use of credit scoring had a positive and statistically significant effect on sampled banks' portfolio share of small-business loans. For the US small business cases, Berger and Frame (2007) defined credit scoring as a statistical approach to predicting the probability that a credit applicant will default or become delinquent. Survey finding concluded that large banks have a comparative advantage in transactions lending technologies based on the analysis of quantitative information. Meanwhile, small banks have a comparative advantage in relationship lending based on qualitative information. It also concluded that small business credit scoring increased small business credit availability.

For models evaluation, Kwan and Tan (1986) evaluated a credit scoring model using discriminant analysis on loan portfolio of a major local bank in Singapore. Both financial and non-financial data inputs were being used. The financial data were categories of financial ratios such as cash flow, activity, profitability, leverage, and liquidity ratios. Meanwhile, the non-financial data included the management record, the prospects for the growth and further development of the business, the firm's policies in key functional areas, and the skills and competitive advantage of the company as a whole. The results showed that there was a 96% of classification accuracy.

Hoffmann *et al*. (2002) compared two types of fuzzy classifiers for credit scoring, the first classifier being Genetic Fuzzy and the second classifier being a Neuro fuzzy by using credit scoring from a major Benelux financial institution. Results concluded that the Genetic Fuzzy classifier yielded the best classification accuracy. However the method was less intuitive and humanly understandable. In contrary, the Neuro fuzzy classifier yielded inferior classification performance. Nevertheless the classifier rule base was both compact and comprehensible for the credit scoring expert. Later, Bijak and Thomas (2012) examined the effectiveness of segmented customer data on credit scoring, applying data provided by two major UK banks and one of the European credit bureaus. It was concluded that customer segmentation did not always improve model performance in credit scoring. Using SMEs data from one Tunisian commercial bank, Khemais *et al*. (2016) concluded that for credit risk prediction logistic regression performed slightly better than discriminant analysis technique. Recently, Boughaci and Alkhawaldeh (2018) evaluated range of machine learning classifiers on credit scoring on the data (German, Australian, Japanese, Polish, Indian Qualitative Bankruptcy and Taiwan default of credit card clients) from Machine Learning Repository of the University of California at Irvine. It was concluded that the Bayes Net and Boosting classifiers generally were the effective tools for credit scoring.

For a list of considered variables in credit scoring and loan approval process, Table 1 below shows gathered variables from the literature reviews.

Table 1: List of Considered Variables in Credit Scoring

Author (Year)	Types of Loan	List of Considered Variables
Orgler (1970)	Commercial	Liquidity: Current Assets/Current Liabilities; Working Capital; Cash/Current Liabilities; Inventory/Current Assets; Quick Ratio; Working Capital/Current Assets. Profitability: Net Profit/Sales; Net Profit/Net Worth; Net Profit/Total Assets; Net Profit (More Or Less Than 0); Net Profit. Leverage: Net Worth/Total Liabilities; Net Worth/Fixed Assets; Net Worth/Long-Term Debt; Net Worth (More Or Less Than 0). Activity: Sales/Fixed Assets; Sales/Net Worth; Sales/Total Assets; Sales/Inventory; Sales/Receivables.
Capon (1982)	Consumers	Zip Code, Bank Reference, Type of Housing, Occupation, Time at Present Address, Time with Employer, Finance Company Reference, Cards (Department Store/Oil/Major Credit Card)
Kwan and Tan (1986)	Commercial	Financial Indicators: Liquidity Ratios, Leverage Ratios, Activity Ratios, and Profitability Ratios Non-financial Indicators: the management record, the prospects for the growth and further development of the business, the firm's policies in key functional areas, and the skills and competitive advantage of the company as a whole.
Hand and Henley (1997)	Consumers	Time at Present Address, Home Owner Status, Postcode, Telephone, Applicant's Annual Income, Credit card, Type of Bank Account, Age, Type of Occupation, Purpose of Loan, Marital Status, Time with Bank, and Time with Employer

Table 1: (Continue)

Author (Year)	**Types of Loan**	**List of Considered Variables**
Hoffmann *et al*. (2002)	Consumers	Identification Number, Income, Age, Profession, Marital Status, Property, Economical Sector, Employment Status, Title/Salutation, Term, Purpose Amount of Loan, Monthly Payment, Amount on Purchase Invoice, Percentage of Financial Burden, Private or Professional Loan, Savings Account, Other Loan Expenses, Code of Regular Saver, Number of Years Client, Number of Years Since Last, Number of Years Employed Number of Years In Belgium, Number of Years Since Last, Number of Checking Accounts, Number of Term Accounts, Number of Mortgages, and Number of Persons Responsible for
Altman and Sabato (2007)	Commercial	Leverage: Short Term Debt/Equity Book Value Liquidity: Cash/Total Assets Profitability: EBITDA/Total Assets Coverage: Retained Earnings/Total Assets Activity: EBITDA/Interest Expenses
Hu and Ansell (2007)	Commercial	Net Profit Margin, Gearing Ratio, Total Debt/(Total Debt + Market Capitalization), Operation Cash Flow, Payables Turnover, Total Assets, The Five Years Correlation Coefficient Between Government Debt and Total Sales
Falangis (2008)	Consumers	German Dataset: Status Of Existing Account, Credit History, Duration in Month, Purpose, Credit Amount, Savings Account, Present Employment, Personal Status and Sex, Installment Rate in Percentage of Disposable Income, Other Debtors, Present Residence, Property, Age in Years, Other Installments Plans, Housing, Number of Existing Credits at This Bank, Job Dependents, Telephone, and Foreign Worker. Greek Dataset: Age, Resident Status, Occupation, Income, Number of Dependents, Phone, Marital Status, Time in Job, Other Card, Mobile Phone, Job Phone, Bank Account, Time in Address, Home Address, and Job Address.

Table 1: (Continue)

Author (Year)	Types of Loan	List of Considered Variables
Moon and Sohn (2010)	Commercial	Technology: Knowledge Management, Technology Experience, Management Ability, Fund Supply, Human Resource, Environment of Technology Development, Output of Technology Development, New Technology Development, Technology Superiority, Technology Commercialization Potential, Market Potential, Market Characteristic, Product Competitiveness, Sales Schedule, Business Progress, Return on Investment. SME Specific Characteristics: Stock List, External Audit, Investment by Foreigners, Professional Manager, Venture Company, INNO Biz, New Company, Production Stage, Joint Company, Category of Business. Economic Indicators: Total business environment index, Economic situations index of SMEs, Economic preceding index, Business survey index, Korean Composite Stock Price Index, Operation index of SMEs, Consumer price index, An earning rate of the national bonds in 3 years, The exchange rate of Won per Dollar.
Khemais *et al*. (2016)	Commercial	Value Added Ratio, Households' Part, Settlement Periods of Supplier Credits, Net Return of Equity, Working Capital, Liquidity Ratio, Repayment Capacity
Ahelegbey *et al*. (2018)	Commercial	Leverage Ratio, Total Asset, Total Liability, Current Ratio, Quick Ratio, ROI, ROE, Asset Turnover, Return on Sales (ROS), Debt Conversion Ratio, Debt Ratio, Return on Capital Employed (ROCE)
Niu *et al*. (2019)	Consumers	Demographic information: Age, Gender, Marital Status, Number of Children, Family Member Counts, Education Status , Income Level, Car Ownership, Income Category, Job Title, House Ownership, Days of Work Registration information: Minutes of Registration, and minutes before the borrower changed the document with which he applied for the loan Loan Information: Loan Amount, Interest Rate, Time, and Repayment Period Social Network Variables: Social Stability, Social Exposure, and Social Quality

Table 1: (Continue)

Author (Year)	Types of Loan	List of Considered Variables
Gou *et al*. (2019)	Commercial	Credit History, Repayment Behavior, Employment Stability, and Other Indicators from UCI machine learning repository
Berg *et al*. (2020)	Commercial	Order Amount, Gender, Age, Credit Bureau Score, Digital Footprint Variables (e.g. Device, OS, E-mail host)

Source: Gathered by Author

From Table 1, it can be seen that there are varieties of considered variables used in credit rating literature. Nevertheless, the variables considered for the commercial loan can be divided into two groups: business environments (e.g. actual industry/business performance and future outlook of the business, economic conditions, government policies) and applicants' qualifications (e.g. the SMEs financial performances, owner's experience in business, debt records of the business and the owner, types of business premise). The next section explains the research framework, accordingly.

Research Framework

From the reviews mentioned this research framework can be illustrated as Figure 1 below:

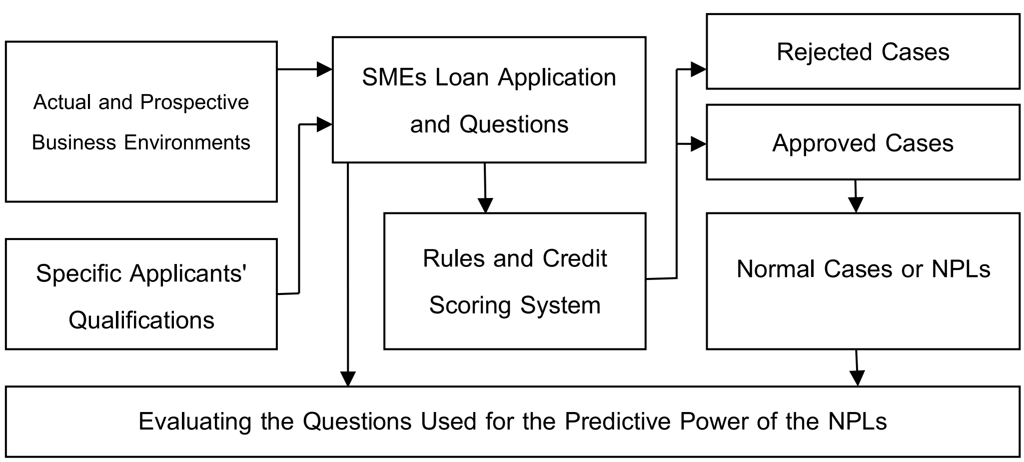

Figure 1: Research Framework

Source: Author

Figure 1 explains the research framework as: from the actual and prospective business environments and specific applicants' qualifications, the loan approval data are gathered from questions being used in the loan application. With the rules and credit scoring system set by creditors, the loan approval results of approving or denying the loan applications are made. Later on, the currently approved loan statuses are evaluated i.e. the normal cases or the NPLs. Hence, the research process here is to evaluate the questions used for the predictive power of the NPLs.

Research Methodology

Allowed to release data of 640 applicants from the year of 2016 from a non-disclosed leading bank for SMEs in Thailand are explored. There are 12 SMEs loan application question being used in credit scoring system, detailed in Table 2.

Table 2: SMEs loan application questions

Number	Details	Scale
Q1	Industry Rating	1 to 5
Q2	Registered Capital	1 to 5
Q3	Income Regularity	1 to 2
Q4	Applicant's experience in business	1 to 5
Q5	Premise Ownerships	1 to 4
Q6	Length of time of the business	1 to 5
Q7	Length of time in contacting with creditor	1 to 6
Q8	Debt repayment records	1 to 4
Q9	Debt restructuring records	1 to 4
Q10	Income Growth Rate	1 to 5
Q11	Debt Service Coverage Ratio	1 to 5
Q12	Debt/Equity Ratio	1 to 5

Source: Non-disclosed leading bank for SMEs in Thailand

In conducting this research, firstly the questions are factorized for dimension reduction. Principle component analysis with Promax rotation, allowing extracted factors to be correlated, is utilized. After obtaining factorized questions, secondly, the factors are then being used as the predictors in debt classification model. The model being employed here is a Binary Logit Model. Therefore, from the model, the sampled applicants are then classified into two classes i.e. the normal debts and the NPLs. Predictability of the predictors and the model can be measured against the current real status of the approved applicants. Below, Figure 2, illustrates the methodology mentioned.

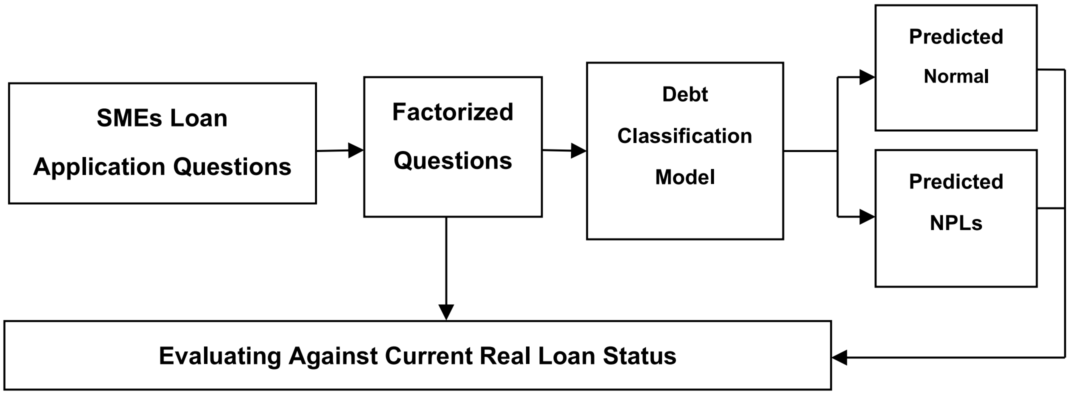

Figure 2: Research Methodology

Source: Author

Research Results

Firstly, the exploratory factor analysis is employed on the 12 question data mentioned. Kaiser-Meyer-Olkin measure of sampling adequacy obtained equals 0.61, regarding as mediocre (Kaiser and Rice (1974)). Meanwhile Bartlett's test of sphericity shown significant result rejecting hypothesis of identity correlation matrix between items being analyzed (Sig. = 0.000). Hence, the sample is suitable for structure detection. Principle component analysis with Promax rotation, allowing extracted factors to be correlated, is utilized. Rotated results can be seen below.

Table 3: Rotated Factors

	Component				
	1	**2**	**3**	**4**	**5**
Q6	**0.915**	-0.104	-0.048	0.098	-0.051
Q4	**0.905**	-0.088	-0.066	0.081	-0.028
Q7	**0.725**	0.209	-0.059	-0.108	-0.012
Q2	**0.386**	-0.023	0.365	-0.335	0.184
Q9	-0.010	**0.872**	-0.001	0.002	0.063
Q8	-0.010	**0.870**	-0.071	-0.019	-0.014
Q3	-0.022	-0.017	**0.748**	0.352	0.097

Table 3: (Continue)

	Component				
	1	2	3	4	5
Q5	0.209	0.106	**-0.497**	0.201	0.055
Q12	0.072	0.086	-0.070	**0.761**	0.059
Q11	-0.008	-0.114	0.188	**0.529**	-0.022
Q10	-0.042	0.106	0.118	0.040	**0.857**
Q1	0.051	0.188	0.455	0.030	**-0.540**

Source: Author Calculation

Extraction Method: Principal Component Analysis.

Rotation Method: Promax with Kaiser Normalization.[a]

a. Rotation converged in 5 iterations.

Result above shows that, from research objective 1, from the 12 questions being used in the credit scoring system, there are 5 extracted factors namely:

- **Factor 1 (Q6, Q4, Q7, Q2): "Time and Size"** indicating length of time in business and in contacting with creditor, owner's experience, and size of the business (which is likely depends of time)
- **Factor 2 (Q9, Q8): "Debt Records"** consisting of debt restructuring and repayment records of the applicants.
- **Factor 3 (Q3, Q5): "Income Reliability"** indicating income regularity and types of premise ownerships. Applicants owned business premise suggests the higher likelihood of income regularity (comparing with rented/temporary/or no business premise)
- **Factor 4 (Q12, Q11): "Debt Burden"** consists of two financial ratio i.e. debt/equity ratio and debt service coverage ratio, signifying debt burden of the business.
- **Factor 5 (Q10, Q1): "Business Prospective"** combines two indicators of industry outlook rating and recent business income growth rate.

Secondly, in order to prove predictive power of the questions, the 5 extracted factors above are then used as predictors or explanatory variables in a Binary Logit model for non-performing loans (NPLs) classification, i.e. being NPLs or normal debt status. Step-wise regression technique is employed. Final model results i.e. parameters estimation and its significance can be seen below:

Table 4: Logistic Regression Results

Remaining Factors	B	S.E.	Wald	df	Sig.	Exp(B)
Factor 1: Time and Size	-0.704	0.094	56.578	1	0.000	0.495
Factor 2: Debt Records	-0.247	0.090	7.463	1	0.006	0.781
Factor 4: Debt Burden	-0.188	0.089	4.450	1	0.035	0.829
Constant	-0.716	0.091	61.202	1	0.000	0.489

Source: Author Calculation

To answer research objective 2, table 4 above shows backward stepwise Binary Logit regression result, after 3 steps i.e. excluding statistically insignificant factors-factor 3 and factor 5, there are 3 factors remaining in the model. The 3 factors are Factor 1: Time and Size, Factor 2: Debt Records, and Factor 4: Debt Burden. Table 5 below shows overall model percentage correct of 69.3%. For normal loans (NPL=0), the correct percentage is at 89.2%. Meanwhile, for case of NPLs (NPL=1), the correct percentage is at 31.5%.

Table 5: Classification Table

Observed		Predicted		
		NPL		Percentage Correct
		0	1	
NPL	0	362	44	89.2
	1	146	67	31.5
Overall Percentage				69.3
Source: Author Calculation (The cut value is .500)				

In summary for the research results against the objectives stated, table 6 is then provided below.

Table 6: Summary of Research Objectives and Results

Research Objectives	Results
1. To factorized SMEs loan application questions for dimension reduction and considering potential redundancy of the questions	There are five factors found i.e. Time and Size, Debt Records, Income Reliability, Debt Burden and Business Prospective. So, application questions can be drop down to these factors preventing redundancy and potentially increase effectiveness of the loan approval process.

Table 6: Summary

Research Objectives	Results
2. To explore the predictive power of the factorized questions as predictors of the selected debt classification model-the Binary Logit Model	From the five extracted factors mentioned, findings for this objective can be summarized as: there are 3 factors significantly explain the likelihood of being future NPLs namely Time and Size, Debt Records, and Debt Burden. Other two factors (Income Reliability and Business Prospective) are proved not the significant factors.

Discussion

In the light of loan approval and credit screening and scoring process, modern literature can be traced back to the 1970s. Factors regarding overall economic situation, industry environment, the financial performance of the applicants' business, owner's experience in business, debt records of the business have been gathered, analyzed, modeled, in predicting the possibility of being future NPLs with no personal discretion. Although the effort of the credit screening process is universally adopted, NPLs, or sometimes even creditors' financial catastrophes happen. Application questions in the credit scoring process become the main tool in gathering the data, henceforth, formulating the research objectives mentioned. With the specific released data obtained from real-world practice, the credit screening questions can then be explored. From the research results obtained, there are some discussion points, according to the research objectives, concerned.

For screening, SMEs application credit providers should pay more attention to the questions related to the 3 factors remaining in the model. For "Time and Size", surviving the business in a length of time, long-term relationship with the creditor, and owner's business experience would indicate the robustness of the business and skill of the applicants in managing the business. For the length of time in business, this can be a good indicator of the survival experience of the applicants. The longer the time in business, the more cases of risky situations being exposed. Nonetheless, every business has its own timing sometimes being called a product or business life cycle. Once the cycle reaches maturity or the business being replaced or disrupted by new superior products or services, business growth ends. Thus, length of time in business is a vital not sufficient factor in the credit screening process. Assigning more weight on the business length of time reduces chances for newly found business or business of the future.

Turning to the next topic related to time, a long-term relationship with creditors, this would indicate business continuity and royalty of the SMEs towards creditors. Meanwhile, new potential customers create new opportunities, also post new risks to creditors' loan portfolio. As credit providers are in the financial intermediary service business, service or brand switching should not be surprised. Customers want the best available service quality. Thus changing from one to another credit provider seems reasonable for good credit customers suffering from their previous service providers. In competition, gaining new customers from competitors would be appreciated by any business. Consequently, a long-term relationship with a credit provider is also a good indicator but new customers should always be welcomed.

Next, the owner's business experience would reveal management skills. The longer the experience, the more practical skill in business management accumulated. Lastly, for this factor, the size of the business also is a good sign of a strong and viable business. With an acceptable debt-leverage level, the size of the business would be an advantage in helping the business to cope with unprepared risk. In conclusion, this result of the importance of Time and Size agrees with previous works mentioned e.g. Falangis (2008), Ahelegbey *et al.* (2018), and Moon and Sohn (2010).

As seen in most of the literature, for "Debt Record", good debt record customer i.e. repayment in time, no debt restructuring records are preferable to any creditor. The reasonably good debt repayment records show the punctuality and discipline of the customers. This also indicates good cash flow management of the applicants. In addition, debt restructuring records, even though sometimes being caused by macro-business-environment or systematic risk, no such records are still being preferable to debt providers. Businesses survived such risk are the stronger ones. Just as mentioned, since all the business record is in the past. There will be no guaranteed future. Yesterday's good business may not be as good tomorrow.

Lastly, "Debt Burden", a low Debt/Equity ratio shows less in financial leverage and risk. A high level of Debt/Equity ratio posts higher risk or difficulty on debt repayment and cash flow management, especially in an economic downturn period. At the same time, the higher the debt service coverage ratio, the higher the ability to repay back the debt. Financial ratios regarding debt burden also can be seen as selected indicators in many previous works e.g. Orgler (1970), Kwan and Tan (1986), Altman and Sabato (2007), Hu and Ansell (2007), and Ahelegbey *et al.* (2018). It should be noted that there is no universal "acceptable" or "good" Debt/Equity ratio. Some industries in practice may have comparatively high leverage e.g. heavy investment industries. Setting the standard acceptable leverage levels across industries is then a real challenge. Even within the same industry, differences in debt burden across companies in the same business can still be sensible and explainable. Hence, high Debt Burden does not necessarily mean bad for business.

Meanwhile, as suggested by the model, the excluded extracted factors are "Income Reliability" and "Business Prospective". The two factors both related to the facts from the past (which do not necessarily imply the continuity into the future) and the discretionary predicted future of the business from the creditor's point of view (which is not a certain fact of the future). These may regard as the weakest links here in the chain of credit scoring questions. However, these factors may also still being considered but with comparatively less importance. Although the result here showing that a discretionary view of the future of the business is statistically insignificant, the view still helps creditors to evaluate the likelihood of the future. So, the question is not whether should one use this discretionary view of the future of the business as an indicator of the credit screening process but how precise the view it is.

Recommendation for Business

From the discussion above, the application questions in the credit scoring process become the main tools in gathering the data. The research results here can be applied to creditors in designing question items. Firstly, the question will be seen in groups e.g. Time and Size, Debt Records, Income Reliability, Debt Burden, and Business Prospective. Within these five groups of questions, the relevant itemized questions will be asked or decided by loan approval officers. The allocated score should be weight mainly on the three significant less discretionary factors i.e. Time and Size, Debt Records, and Debt Burden. However, as discussed above Income Reliability and Business Prospective still also be considered with care.

Recommendation for Further Research

In order to avoid data bias, if possible, both types of SMEs applications i.e. rejected or approved applicants should be included in the study, as suggested by Capon (1982). However, it is understandable that, for the rejected cases, following current information such as current sales, profits, and overall businesses' viability are reasonably difficult to be gathered. Although the sample size in this research work is adequate, larger sample size is recommended for future studies. The larger the sample size, the more complicated model allowed. Furthermore, specific studies based on different types of industries or other types of subcategories of SMEs should be concerned. The differences in business environments and risks exposed in each industry may yield different and interesting results. Lastly, although overall model predictive accuracy is at 69.3%, the percentage is at 31.5% for correctly predicted NPLs cases. Thus, more studies on the validity of credit scoring questions and also more varieties of the models proposed are needed.

References

Ahelegbey, D. F.; Giudici, P. and Hadji-Misheva, B. 2018. **Latent Factor Models for Credit Scoring in P2P Systems**. MPRA Paper No. 92636. Retrieved November 25, 2020 from https://mpra.ub.uni-muenchen.de/92636/

Altman, E. I. and Sabato, G. 2007. **Modelling Credit Risk for SMEs: Evidence from the U.S. Market**. ABACUS. Vol. 43. No. 3. 332-357.

Berg, T.; Burg, V.; Govombovic, A. and Puri, M. 2020. **On the Rise of FinTechs: Credit Scoring Using Digital Footprints**. The Review of Financial Studies. Vol. 33. 2845–2897.

Berger, A. N. and Frame, W. S. 2007. Small Business Credit Scoring and Credit Availability. **Journal of Small Business Management**. Vol. 45. No. 1. 5–22.

Bijak, K. and Thomas, L. C. 2012. Does segmentation always improve model performance in credit scoring?. **Expert Systems with Applications**. Vol. 39. 2433–2442.

Boughaci, D. and Alkhawaldeh, A. A. K. 2018. Appropriate machine learning techniques for credit scoring and bankruptcy prediction in banking and finance: A comparative study. **Risk and Decision Analysis**. Vol. 1. 1-10.

Capon, N. 1982. **Credit Scoring Systems: A Critical Analysis. Journal of Marketing**. Vol. 46. No. 2. 82-91.

Falangis, K. 2008. The use of MSD model in credit scoring. Operational Research. **An International Journal**. Vol. 7. No. 3. 481-50.

Frame, W. S., Srinivasan, A., and Woosley, L. 2001. The Effect of Credit Scoring on Small-Business Lending. **Journal of Money, Credit and Banking**, Vol. 33. No. 3. 813-825.

Guo, S.; He, H. and Huang, X. 2019. **A Multi-Stage Self-Adaptive Classifier Ensemble Model With Application in Credit Scoring**. IEEE Access, Vol. 7. 78549-78559.

Hand, D. J. and Henley, W. E. 1997. Statistical Classification Methods in Consumer Credit Scoring: a Review. **Journal of Royal Statistical Society**. Vol. 160. 523-541.

Hoffmann, F.; Baesens, B.; Martens, J.; Put, F. and Vanthienen, J. 2002. Comparing a Genetic Fuzzy and a Neurofuzzy Classifier for Credit Scoring. **International Journal of Intelligent Systems**. Vol. 17. 1067–1083.

Hu, Y.C. and Ansell, J. 2007. Measuring retail company performance using credit scoring techniques. **European Journal of Operational Research**. Vol.183. 1595–1606.

Kaiser, H. F. and Rice, J. 1974. Little Jiffy, Mark IV. **Educational and Psychological Measurement**. Vol. 34. 111-117.

Khemais, Z., Nesrine, D. and Mohamed, M. 2016. Credit Scoring and Default Risk Prediction: A Comparative Study between Discriminant Analysis & Logistic Regression. **International Journal of Economics and Finance**. Vol. 8. No. 4. 39-53.

Korwatanasakul, U. and Paweenawat, S. W. 2020. **Trade, Global Value Chains, and Small and Medium-Sized Enterprises in Thailand: A Firm-Level Panel Analysis**. ADBI Working Paper Series. No. 1130.

Kwan, K. C. and Tan, J. 1986. Credit Scoring for Commercial Loans: The Case of Singapore. **Asia Pacific Journal of Management**. Vol. 4. No.1 41-54.

Moon, T.H. and Sohn, S.Y. 2010. Technology credit scoring model considering both SME characteristics and economic conditions: The Korean case. **Journal of the Operational Research Society**. Vol. 61. 666-675.

Niu, B.; Ren, J. and Li, X. 2019. Credit Scoring Using Machine Learning by Combing Social Network Information: Evidence from Peer-to-Peer Lending. **Information**. Vol. 10. 397.

Orgler, Y. E. 1970. A Credit Scoring Model for Commercial Loans. **Journal of Money, Credit and Banking**. Vol. 2 No. 4. 435-445.

Thomas, L. C.; Edelman, D. B. and Crook, J. N. 2002. **Credit Scoring and Its Application. Philadelphia**. Society for Industrial and Applied Mathematics.