SME CREDIT RISK MODELLING IN SOUTH AFRICA: A CASE STUDY

1 Mercy Marimo, African Export-Import Bank, Cairo, Egypt
2 Charles Chimedza, University of the Witwatersrand, South Africa
*Corresponding author E-mail: mmarimo@afreximbank.com
1 ORCID ID: 0000-0002-6973-3064
2 ORCID ID: 0000-0001-8980-1610

ARTICLE INFO

Original Scientific Paper
Received: 19.09.2021
Revised: 05.01.2022
Accepted: 11.04.2022

doi 10.7251/ACE2236009M
UDC
336.717.061:330.554(680)

Keywords: *SME scorecard, consumer credit, reject inference, credit scoring*

JEL Classification:
C35,G32

ABSTRACT

This paper examined the funding conundrum by assessing the success rate of applications of small and medium sized enterprises (SME) for commercial bank funding. A quantitative group analysis was done on the overdraft obtained from one of the leading financial institutions in South Africa to determine the drivers of default. The SME scorecard was developed using logistic regression on credit applications over a seven-year observation period to analyse the default experience as part of credit risk management. The robustness, stability and relevance of an application scorecard is enhanced by the reject inference process and inclusion of bureau information. Small businesses operating in the service sector and having a long-standing rapport with the bank can easily access commercial bank funding. SMEs in the construction industry with a high number of credit enquiries are unlikely to survive the stringent conditions of the bank lending criteria. It is the prerogative of the principal business owner to honour their financial obligations across the credit industry if commercial bank funding is desired. Their credit quality, as reflected in the bureau information, forms the fulcrum of the SME application scorecard. The model developed in this study can be used as a tool to reduce defaults and serious delinquencies in boarding new applicants. Furthermore, the model can be applied to determine risk tendency and monitor the performance of SME credit portfolios.

1. INTRODUCTION

Most financing options for small and medium sized enterprises (SME) are limited to inception stages of a business cycle, contributing to seed and initial growth funding. From inception, small businesses are largely funded through personal savings, family, friends, donations, business angels, retained earnings, etc. (Chimucheka & Rungani, 2013). As businesses develop, a wider spectrum of funding is sought to finance operations and rapid growth, one of which is commercial bank loans. However, SME accessibility to commercial funding is limited (Laurens, 2012). Lack of adequate funding is the main operating and growth constraint of SMEs (Berger & Udell, 1998). As SMEs form an integral aspect of most economies globally, it is crucial to investigate this constraint and suggest ways to mitigate this challenge. Given that vast research on SME funding in South Africa is based on cross-sectional data, it is difficult to make a causality claim when establishing the link between SME specific attributes and access to finance (Makina et al., 2015). This study therefore seeks to utilise longitudinal data to perform regression analysis to shed more light on factors affecting the success rate of SMEs in accessing commercial funding.

This paper investigates the determinants or drivers of the success rate of SME access to commercial funding in South Africa. This is done by using an advanced statistical model, making use of a wealth of historic information. To the best of our knowledge, inadequate funding of SMEs from financial institutions in South Africa is a result of a myriad of factors which include: comprehensive enforcement of regulatory requirements, lack of collateral, information asymmetry, moral hazards, lack of sound track records on credit performance, a technological division between lenders and borrowers and lack of financial records. Investigating the financing conundrum of South African SMEs through tracking groups of SME loan applications forms the fulcrum of this study. The main aim of this study is to investigate the success rate of SME application for commercial funding using data obtained from one of the leading financial institutions in South Africa. The research objective is: *To identify drivers of default and determine how the model is affected by introducing bureau data.*

2. STYLISED FACTS ABOUT THE SME SECTOR IN SOUTH AFRICA

The South African government recognizes the importance of the SME sector as evidenced by the establishment of the Ministry of Small Business Development in 2014 responsible for the facilitation, promotion and development of SMEs

(SEDA, 2016). This was set up to implement policies, strategies and programs earmarked to create an enabling environment for SME growth and development. This department operates under various responsible institutions or agencies. The Small Enterprise Development Agency (SEDA) implements SME business strategy. This agency is also involved in the development and implementation of policies and standards as well as integration of government-funded SMEs across the three tiers of government (SEDA, 2016).

The Small Enterprise Finance Agency (SEFA) merged with the South African Micro-Finance Apex Fund (SAMAF) and Khula Enterprise Finance Limited to cater for SMEs with funding requirements for less than or equal to R3 million (SEDA, 2016) through the provision of revolving loans, term loans, bridging finance among other governmental financial support streams. Technical support is provided through the National Youth Development Agency (NYDA) targeted at young South Africans of between the ages of 14 and 35. The National Empowerment Fund (NEF) was also established to provide non-financial support to black-owned SMEs. Despite the unmerited government intervention and support, individual SME growth remained inhibited by various challenges at various scales depending on the size and scope. Commercial lenders are less likely to lend start-ups and informal businesses, which form the greatest proportion of SMEs. In the Gauteng province SMEs are more likely to get funding compared to those in Mpumalanga and Northern Cape provinces. This is mainly due to the predominantly rural nature of the latter provinces and a lack of access to physical infrastructure, and a widening gap in the technological division. Skills shortages, permit delays and high levels of crime are some of the obstacles hampering SME growth across the country (SEDA, 2016).

The SME segment is considered an integral aspect of the economy in South Africa. The Department of Trade and Industry (DTI) developed and published the National Small Business Act (President's Office, 1996). The public and private sectors define SMEs across several economic industries for ease of funding and other forms of support. DTI is a government department which is known to promote structural transformation and economic development. In the Small Business Act (President's Office, 1996), revenue, gross asset value and size, with maximum thresholds of R40 million, R18million and 200 employees respectively segment SMEs.

3. THEORETICAL FRAMEWORK

3.1. Background

Njoku & Odii (1991) used linear regression to model the actual outstanding debt and they identified loan volume, years in business and experience, major occupation, years of formal education, household size, loan period, business size and business output as drivers. On the other hand, Valluri, Raju & Patil (2021) modelled loan churn. They used different classical statistical models and machine learning models, such as logistic regression (LR), linear discriminant analysis (LDA), decision trees (DT) and random forests (RF). They found value in pre-selecting variables. They concluded that the RF classification measures report the strongest performance by using all the variables. Similar methodologies were applied by Fantazzini & Figini (2009) in comparison with a standard logit model which was observed to perform better timeless sample compared to RFs. De Noni, Lorenzon & Orsi (2007) developed a qualitative and quantitative risk model to measure and manage credit risk in SMEs. The model was designed with the ability to encapsulate expert judgement which is not necessarily captured by quantitative methodologies. It is perceived that SMEs' knowledge of credit criteria is somewhat low as commercial banks are not very transparent in this area. Authors argued that the length of operating a business has a significant influence on the evaluation of credit risk factors with more experienced SMEs in operation for over 10 years having a more intense perception of the significance of credit risk (Dvorský, Schönfeld, Kotásková & Petráková, 2018).

3.2. Credit Scoring

Statistical models to evaluate the creditworthiness of applicants can enhance regulatory and legislative requirements. The applications undergo a credit scoring process where ratings are assigned to reflect the ability and willingness of borrowers to repay debt timely and in full. Credit scoring forms the cornerstone of credit risk management by offering a systematic way of assessing the credit quality of obligors and this, in the past led to better credit granting decisions (Wendel & Harvey, 2006). Credit scoring can be divided into two main pillars: the front end (acquisition) and the back end (existing customer). The front end deals with through the door customers where the application scorecard contributes to the credit lending decision process. The application scorecard becomes an important aspect for business acquisitions. The back end process uses the behavioral scorecard to determine the risk levels of existing customers and inform credit risk management and collection strategies (Beck, 2013). To enhance the business acquisition process for lenders, an application scorecard is extensively discussed in this paper. Financial service providers develop scorecards using historic data with an assumption that the historic trends are like future experiences.

3.3. Credit Rationing

Credit rationing is a market imperfection phenomenon where lenders limit the supply of credit to borrowers demanding funds, even if the latter are willing to pay higher interest rates (Mutezo, 2015). Credit rationing occurs due to information and control limitations in the financial markets. This event reflects failure of price mechanisms which in turn miscarries market equilibrium. SME credit industry suffers credit rationing when SMEs fail to provide sufficient collateral to hedge against potential credit losses by the lenders.

4. MATERIALS AND METHODS

4.1. Data Sources

Research data was obtained from one of the leading financial institutions in South Africa. From the main data warehouse, the application, behavioural and performance information on SMEs were extracted and insights were drawn from the empirical data. Bureau data, macroeconomic information and credit industry data were sourced from external institutions like Experian and Moody's for each month of observation.

4.2. Population, Sampling Approach and Sample Size

For the SME application scorecard model development, groups of applications received in each month under observation were tracked to determine the behaviour of applicants over time. In order to check the performance of the models and to ensure that no over/under-fitting occurs, an independent holdout validation data set was set aside from the complete data set. The full data set was randomly split using simple random sampling into a training and a validation data set in the ratio 80:20 respectively. This ratio was exercised in the works of Marimo & Chimedza (2017), Vrigazova (2021) and Visser et al. (2000). The model was developed on the training data set and the performance of the model tested on the independent validation set. The total number of overdraft (OD) applications submitted to the bank by SME applicants during the seven-year period determined the sample size. This study considers applications only up to January 2018 to allow for at least 18 months (February 2018 to July 2019) performance of the loans. Figure 1 shows the population flow of SME loan applications.

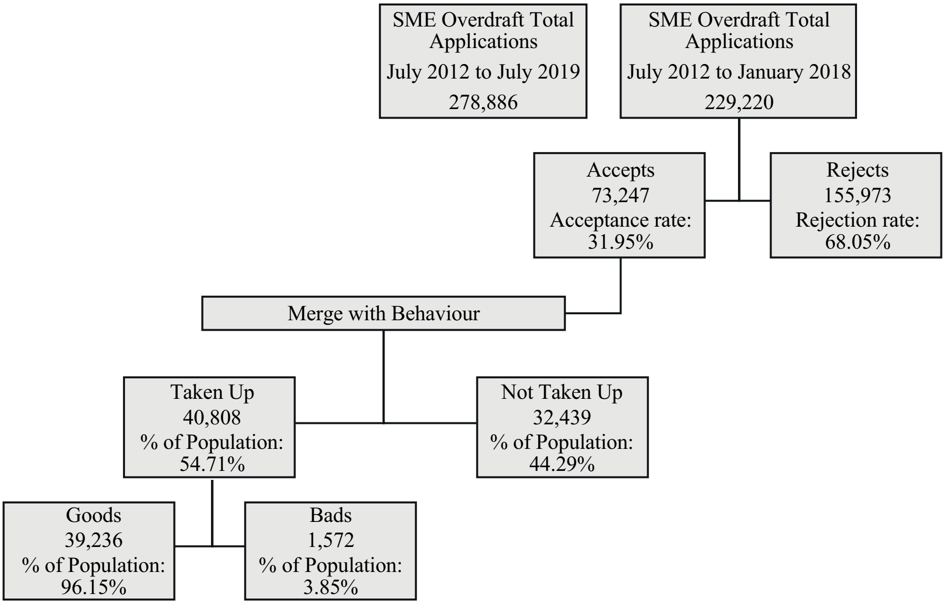

Figure 1: Population Flow

Source: Authors' calculation

It also shows how the data was prepared for the scorecard model development. Of the 73,247 approved applications, 44.29 percent were not taken up by the SMEs. This is mainly due to the cold scoring technique used by the lender to grow the business. Cold scoring is a process whereby potential clients are identified using propensity score models and other internal processes. The lender generates the applications on behalf of the customers. It is the prerogative of the SMEs to take up or decline the offer. The statistics herein show that almost half of the applicants do not take up the offers; this is probably because they do not need the credit facility offered at the time. On average, about 3.85 percent of SMEs default on loans within the first 18 months of their OD credit facility.

Table 1 shows the volumes allocated in each case.

Table 1: Sample Design

Sample	Category	Number of Observations
Development	Accepts	32647
	Rejects	58275
Validation	Accepts	8161
	Rejects	14569

Source: Authors' calculation

4.3. SME Application Scorecard

The relevant information was extracted from the data warehouse of the institution in scope. This entails information on SMEs which submitted applications during the observation period for an overdraft facility. The raw dataset consists of standard variables which cannot be used in modelling but are crucial as identifiers and important indicators for the purposes of segmentation. Standard variables include, ID number, observation month and account number. Sensitive information such as ID number and account number were masked to comply with issues of confidentiality in the business environment and the Protection of Personal Information (POPI) Act (POPI, 2013).

4.3.1. Measuring Default Status

The South African banking industry adheres to the international banking regulations recommended by the Basel Committee on Banking Supervision (BCBS) through the Basel Accords. For the local banks to sustain adequate capital reserves, the Basel Accords are enforced by the South African Reserve Bank (SARB) to ensure sustainability in the event of economic strain. This study therefore follows the default definition as defined in the Basel Accords, Bank for International Settlements (Laurens, 2012) as follows:

“A default is considered to have occurred with regard to a particular borrower when either or both of the two following events have taken place.

- The bank considers that the obligor is unlikely to repay his/her credit obligations to the bank in full.
- The obligor is more than 90 days past the due date on any credit obligation to the banking group.”

This definition of default was used to derive the default status, the dependent or response variable in the application scorecard model building process.

4.3.2. Definition of Variables

Independent variables were extracted from the application tables found in the data warehouse. For the previously accepted applications, the repayment behaviour was tracked in order to establish a link between application variables and default. Additional variables were derived from the readily available raw data if they were deemed to be predictive of loan performance. The choice of variables was mostly driven by available variables some of which were supported by Valluri, Raju & Patil (2021), Njoku & Odii (1991) and Pasha & Negese (2014).

Table 2: Potential Risk Drivers

Variable	Description	Rationale for Consideration
Default Status	A derived binary target variable indicating whether an account defaulted within the outcome period	Binary target variable
Time since last transaction	The time that has elapsed (in months) since the applicants' last credit transaction on their main account with the lender	The higher the number of months since an applicant's last credit transaction, the greater the likelihood of the applicant not having sufficient funds to meet debt obligations and thus the higher the risk of default.
Sector	The industry under which the applicant operates is indicated by this variable	Certain industries tend to be riskier than others and will be allocated comparatively lower scorecard points.
Number of Credit Enquiries	The number of enquiries made by the principal business owner in the last 12 months	Applicants who have made a large number of enquiries in a short period are considered risky and will be allocated low scorecard points.
Time since payment profile	The time that has elapsed (in days) since the principal business owner opened a payment profile	Applicants who acquired their latest payment profile further in the past tend to be less risk than those who have acquired it more recently.
Worst Arrears Recent	The Worst Arrears in the Last 6 Months by the principal business owner	The worse the arrears level in the past 6 months the greater the risk of default and the lower the scorecard points to be allocated.
Worst Arrears Ever	This variable reflects the worst arrears level in the entire credit history of the principal business owner.	The worse the arrears levels in the principal business owner's credit history, the greater the risk of default and thus, the lower scorecard points to be allocated.
Guinness Rating	This variable is based on a set of matrices including Turnover, Time with the lender and time in business.	Applicants who have a low Guinness Rating tend to have a high default risk and will be allocated low scorecard points.
Time with Lender	The period of time (in months) an applicant has been a client of the lender	Applicants who have been the lender's clients for a longer time period are perceived to have a low default risk.
Excess	Business Entity Excess Indicator	Business entities which have never been in excess are perceived to have a low default risk and will be awarded high scorecard points.
Worst Excess	Principal business owner Worst Excess	Principal business owners who have never been in excess are perceived to have a low default risk and will be awarded high scorecard points
Worst Report	This variable represents the principal business owner's worst credit bureau report	The worse the principal's credit bureau report, the higher the risk of default and thus, the lower the scorecard points to be allocated.

Source: Authors' creation

4.3.3. Logistic Regression

The lending criteria in OD largely depend on customer attributes such as the business sector in which the customer operates, its size, growth stage, affordability and creditworthiness as defined by the scorecard rating. The variables consist of demographics, customer relationships with the bank and external information such as credit bureau data. A performance period was driven by the data used for the determination of the default event which will be modelled using logistic regression as follows:

The above logistic regression model was fitted with an expectation to produce:

- the main drivers of default,
- a probability model to be applied at the point of application, and
- a tool used to translate into scorecard points, depending on probability level.

4.3.4. Reject Inference

To achieve stability and robustness of the estimates, as well as to avoid bias in the scorecard, the application scorecard should take into account all applications received by the lender within the outcome period, regardless of whether the application was accepted or rejected. However, the performance information is only available for the accepted and taken-up applications only. If the objective is to measure the impact of the scorecard on all applications, there is need to assign an 'inferred' performance to the rejected applications in the training sample. Reject inference is thus, a process whereby the performance of the rejected applications is inferred or estimated. Furthermore, it is important to note that, not all approved loans get taken-up. To mitigate this complication, the reject inference is applied in two stages:

1. Assign each reject a probability of taken-up if rendered accepted.
2. Assign each reject a probability of good if estimated to be a Taken Up (TU).

5. RESULTS: APPLICATION SCORECARD MODEL DEVELOPMENT

The application scorecard can be used to decide whether to extend credit to applicants with an aim to reduce defaults and serious delinquencies on new applicants. In addition, the model can also be used to allocate capital, determine risk tendency and monitor the performance of the portfolio in scope.

Risk characteristics under the Basel Accords should be calculated and used in conjunction with the scorecard for risk management purposes (BCBS, 2017).

5.1. Reject Inference

– *Stage 1: Infer Non-Taken Up Applications*

The first stage in reject inference is to assign the Non-Taken-Up (NTU) records within the rejected population. All the accepted applications from the training sample were used to infer the TU and NTU probabilities to the rejected records. The application score of the principal business owner was used as a proxy for the credit performance of each SME across its loans. This score was used to fit a relationship between TU and NTU applications. As shown in Figure 2, an inverse relationship between the application score and the TU rate is observed, that is, a higher score results in a lower TU rate. This relationship is intuitive because if offered a loan, the worse performing applicants (lower score) are likely to take up the offer compared to the low risk (high score) applicants.

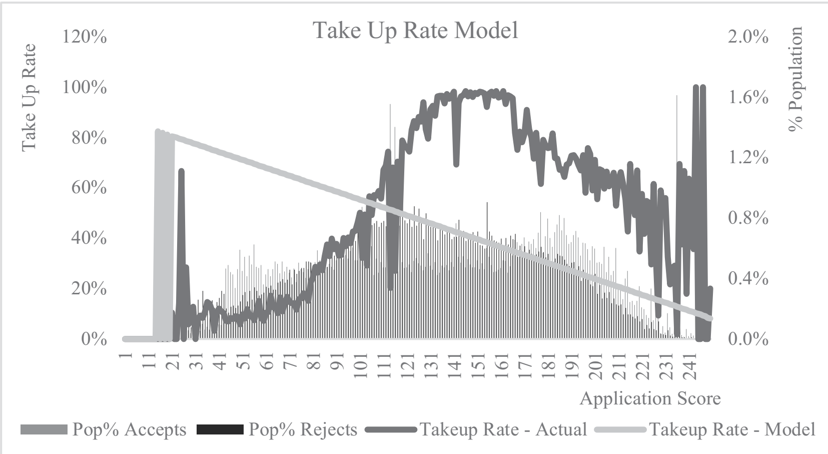

Figure 2: Take up Rate

Source: Authors' calculation

– *Stage 2: Infer Reject Good and Bad Applications*

The second stage in reject inference is to assign the probability of good to rejects if they are estimated to have taken up the loan. This is based on the Known Good Bad (KGB) model (Zeng & Zhao, 2014) built on the subpopulation with known performance (accepted and taken up). The model is based on fitting the application score at outcome point to obtain a relationship between good and bad accounts. The purpose is to predict the probability of TU account being good or bad at outcome. Logistic regression was applied to develop the KGB model. With the reject inference completed, the application scorecard can be developed on a full spectrum of applications received.

5.2. Reject Inference Validity

In application scorecard development, the Good: Bad Odds ratio for accepts and inferred rejects usually fall in the range from 2 to 6. The known to inferred ratio of 2.03 given in Table 3 is an indication that the reject inference results are satisfactory.

Table 3: Ratio of Known Odds to Inferred Odds

Sample		Number of Observations	Badrate	Odds Ratio
Accepts	Goods	31389	3.85%	24.95151033
	Bads	1258		
	Total	32647		
Rejects	Goods	53889	7.53%	12.28662295
	Bads	4386		
	Total	58275		
Development (Accepts + Rejects)	Goods	85278	6.21%	15.10952864
	Bads	5644		
	Total	90922		
Known to Inferred Odds Ratio			2.03078669	

Source: Authors' calculation

5.3. Model Fitting: Internal and Bureau Variables

Significant parameter estimates (at 5%) were selected and shown in Table 4. The Wald test was used to select significant variables by picking only variables with a *p-values* less than . The rest of the variables were dropped as they had no predictive power. The selected variables were consistent with the findings of Valluri, Raju & Patil (2021), Njoku & Odii (1991), Marimo and Chimedza (2017) and Pasha & Negese (2014), in cases where similar variables were studied.

The global tests show that the model with significant covariates is significantly different from a null model. Variables selected are significant drivers of the default rate in the SME credit industry.

Table 4: Model 1 - Internal and Bureau Variables

Parameter	Estimate	Standard Error	Wald χ^2	Odds Ratio	Z- Statistics	P-Value
Intercept	2.7155	0.0146	34676.980	15.112	186.2176	<.0001
Excess	-0.5027	0.0438	131.8428	1.653	11.48228	<.0001
Credit Enquiries	-0.9213	0.0377	596.2210	2.513	24.41764	<.0001
Time since Last Transaction	-0.7320	0.0583	157.54960	2.079	12.55188	<.0001
Time with Lender	-0.6970	0.0310	505.7309	2.008	22.48846	<.0001
Sector	-0.7426	0.0521	202.7845	2.101	14.24024	<.0001
Worst Bureau Report	-0.6899	0.0497	192.3296	1.994	13.86829	<.0001
Likelihood Ratio			2611.2488			<.0001
Score			2650.6101			<.0001
Wald			2495.5706			<.0001
N	90922					

Source: Authors' calculation

5.4. Model Fitting: Internal Variables Only

Table 5 shows the model built on internal variables only is significant, both at global level and at individual parameters. Thus, the model is significantly different from a null model.

Table 5: Model 2 - Internal Variables Only

Parameter	Estimate	Standard Error	Wald χ^2	Odds Ratio	Z- Statistics	P-Value
Intercept	2.717	0.0144	35810.95	15.135	189.2378	<.0001
Excess	-0.5425	0.0434	156.2004	1.720	12.49802	<.0001
Time since Last Transaction	-0.6847	0.0579	140.0839	1.983	11.8357	<.0001
Time with Lender	-0.8261	0.0305	735.7736	2.284	27.12515	<.0001
Sector	-0.8053	0.0518	241.2156	2.237	15.53112	<.0001
Likelihood Ratio			1856.5968			<.0001
Score			1864.4482			<.0001
Wald			1786.9986			<.0001
N			90922			

Source: Authors' calculation

Model 1 and Model 2 were compared to determine the benefit of inclusion/exclusion of bureau information in the SME Application Scorecard.

5.5. Final Model Selection

The two models described in the preceding sections were compared using various statistical measures. Both models were applied to the sample for validation to determine the suitability of the model. Satisfactory results were observed in both cases. In hindsight, the benefit of including/excluding rejected applications in the models was determined.

Table 6 provides measures of the discriminatory power of the models. An 18 percent increase in the Gini Statistic (GS) is realised when the Bureau information is added as part of the covariates. A benefit of 2.07 percent in discriminatory power is realised if the scorecard model development includes the reject inference process.

Table 6: Final Model Selection Criteria

AUC	Development Sample		Validation Sample		% Increase in AUC Development	% Increase in AUC Validation
By Scorecard	Accepts & Rejects	Accepts Only	Accepts & Rejects	Accepts Only		
Bureau & Internal Fields-Model 1	38.4	37.6	38.6	37.5	2.07%	2.96%
Internal Fields Only-Model 2	32.6	30.8	32.3	29.7	5.76%	8.69%
% increase in Gini	18.0%	22.2%	19.3%	25.9%		

Source: Authors' calculation

Similar trends were observed in the validation sample. Furthermore, the Receiver Operating Curve (ROC) confirms that Model 1 exhibits a better discriminatory power than Model 2 as it lies closer to the top left quadrant of the plot as shown in Figure 3.

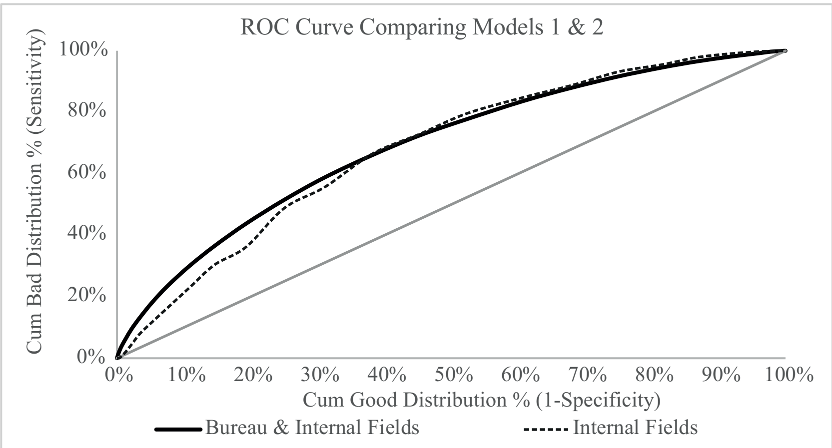

Figure 3: ROC Curves

Source: Authors' calculation

5.6. Scorecard Points

Model 1 (Internal + Bureau Fields) was finally selected as the best model for application in the development of the SME scorecard. The model was fitted to the training data set to obtain probabilities of default. These probabilities were then converted into scorecard points per variable per category within each variable. Scorecard points were linked to the probabilities returned by the model in each case. The intuitiveness of scorecard points, badrate and Weight of Evidence (WoE) for every variable in scope is provided below.

5.7. Final Variable Statistics

1. *Credit Enquiries*

The variable Credit Enquiries is a bureau field detailing the number of enquiries made by the applicant in the past twelve months, and these are shown in Table 7.

Table 7: Credit Enquiries

Credit Enquiries	Scorecard Points	WoE	Goods	Bads	Badrate
01 : Low to <= 1	41	0.6118	13365.72404	479.7830401	3.47%
02 : > 1 to <= 3	18	0.2644	28089.19044	1427.123515	4.84%
05 : > 3 to <= 9	-1	-0.0162	25733.06009	1730.913592	6.30%
11 : > 9 to <= 12	-29	-0.4354	5921.000083	605.6545366	9.28%
12 : > 12 to High	-37	-0.5533	12168.92992	1400.507112	10.32%

Source: Authors' calculation

It satisfied the univariate analysis criteria as shown in Figure 4. The population in each group exceeded five percent. The badrate, WoE and the scorecard point curves are intuitive and monotonic. The larger the number of enquiries, the more uncertain and riskier the applicant is. The badrate increases with an increase in the number of enquiries. Risky applicants have been allocated the lowest scorecard points. This analysis was conducted for the final six variables and the results were satisfactory.

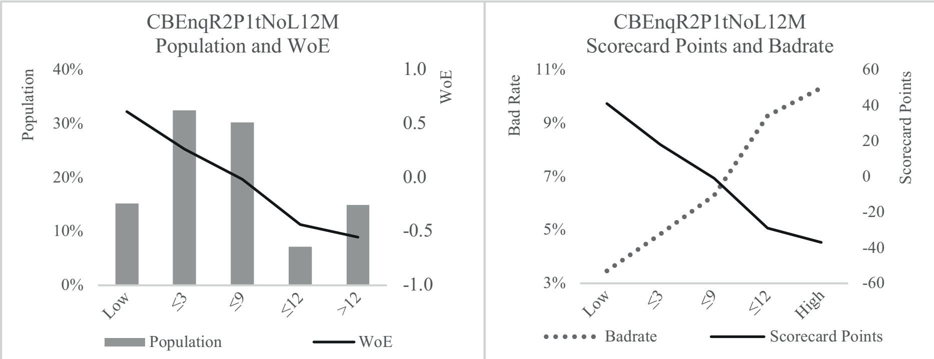

Figure 4: Credit Enquiries

Source: Authors' calculation

2. *Time since Last Transaction*

From Table 8, the higher the number of months since the applicant's last credit transaction is, the greater the likelihood of the applicant not having sufficient funds to meet debt obligations is and thus, the higher the risk of default is. Therefore, the worst scorecard points allocation falls in the highest bracket of this variable.

Table 8: Time since Last Transaction

Time since Last Transaction (months)	Scorecard Points	WoE	Goods	Bads	Badrate
0	16	0.2957	14383.31677	708.2523992	4.69%
00: Missing	-3	-0.0543	13879.35187	969.8282712	6.53%
03 : > 0 to <= 5	7	0.136	38048.75622	2198.078426	5.46%
08 : > 5 to <= 10	-12	-0.2207	8968.773373	740.1437171	7.62%
10 : > 10 to <= 25	-20	-0.3771	7150.569308	690.0311119	8.80%
13 : > 25 to High	-31	-0.5833	2847.13702	337.6478697	10.60%

Source: Authors' calculation

3. *Time with Lender*

Applicants who have been clients of the lender for a longer time period are perceived to have a low default risk and have therefore been allocated with the highest scorecard points as shown in Table 9.

Table 9: Time with Lender

Time with Lender (months)	Scorecard Points	WoE	Goods	Bads	Badrate
1:00 New to Bank	6	0.117	11255.50475	662.6782406	5.56%
02 : > 0 to <= 12	-48	-0.9461	4417.26264	753.01663	14.56%
03 : > 12 to <= 18	-35	-0.6935	3383.392057	447.9991329	11.69%
04 : > 18 to <= 24	-25	-0.5003	3025.404126	330.2269838	9.84%
05 : > 24 to <= 33	-20	-0.4016	3836.680936	379.4216338	9.00%
06 : > 33 to <= 54	-10	-0.1891	7740.215539	618.9333206	7.40%
08 : > 54 to <= 63	-6	-0.1149	3114.742058	231.2523917	6.91%
09 : > 63 to <= 75	3	0.0687	4366.76174	269.8152901	5.82%
10 : > 75 to <= 84	8	0.1635	3381.998157	190.0766928	5.32%
11 : > 84 to <= 93	11	0.2145	3332.768537	177.9909928	5.07%
12 : > 93 to <= 138	14	0.2876	15351.55696	762.0391878	4.73%
16 : > 138 to <= 153	17	0.3315	4098.585437	194.7132633	4.54%
17 : > 153 to <= 270	29	0.5847	14499.19024	534.7366452	3.56%
21 : > 270 to High	47	0.9259	3473.841391	91.08138948	2.55%

Source: Authors' calculation

4. *Excess Levels*

At the point of application, customers are allocated excess levels as seen in Table 10. Business entities which have never been in excess are perceived to have a low default risk and have been awarded the highest scorecard points.

Table 10: Excess Levels

Excess	Scorecard Points	WoE	Goods	Bads	Badrate
01: High	-18	-0.4858	5946.34370	639.720802	9.71%
06: Medium	-9	-0.2454	39080.1601	3305.75989	7.80%
03: Low	16	0.4439	37921.0940	1610.06386	4.07%
05: Never	20	0.5561	2330.30662	88.43723469	3.66%

Source: Authors' calculation

5. *Sector*

Of the non-missing categories, the services sector has been observed to be the best performing with the least bad rate as shown Table 11. The construction industry has been the riskiest and therefore allocated comparatively the lowest scorecard points.

Table 11: Sector

Sector	Scorecard Points	WoE	Goods	Bads	Badrate
01: Missing	113	2.1039	5.765305757	0.046544243	0.80%
02: Retail	6	0.1142	39248.31299	2317.178511	5.57%
03: Construction	-17	-0.3248	19488.07644	1784.704503	8.39%
04: Transport	-14	-0.2592	5922.884616	507.9624441	7.90%
05: Trade	-5	-0.0922	5562.503692	403.698418	6.77%
06: Services	56	1.0488	3535.99568	81.99309005	2.27%
07: Manufacturing	18	0.329	11514.36585	548.3982848	4.55%

Source: Authors' calculation

6. *Worst Bureau Report*

In Table 12 it is observed that the worse the principal's credit bureau report is, the higher the risk of default is and thus, the lower the scorecard points allocated are.

Table 12: Worst Bureau Report

WrstCBReport	Scorecard Points	WoE	Goods	Bads	Badrate
01: C (Worst)	-51	-1.0295	31.84074861	5.899801386	15.63%
02: D	-3	-0.0694	3976.712801	282.1089394	6.62%
03: F	36	0.7191	727.4598513	23.45613871	3.12%
04: N	6	0.1158	66224.36464	3903.819676	5.57%
05: O	-27	-0.5364	8032.360169	909.009951	10.17%
06: S	-19	-0.373	5232.890076	502.9072244	8.77%
07: X (Best)	71	1.4232	1052.276276	16.78006413	1.57%

Source: Authors' calculation

5.8. Scoring Alignment Parameters

The scorecard is aligned to:

- A score of 500 has Good: Bad odds of 5:1
- 50 points double the odds

These parameters were chosen in order to reflect the portfolio bad rate at the reference score. In this case, the development/training sample odds ratio is 15.11 as shown in Table 13. In order for a score of 500 to represent this and a bad rate of 6.21%, the following function is used to determine the Reference Odds (RO):

$$Reference\ Odds = \frac{1}{Odds\ Ratio_{Development}} - 1 = \frac{1}{0.1511} - 1 \approx 5$$

Table 13: Scorecard Alignment Parameters

Alignment Parameter	Value
Bad Rate (Accepts + Rejects)	6.21%
Reference odds	15
Reference Score	500
Points to double odds	50

Source: Authors' calculation

Table 14 demonstrates the relationship between the theoretical Odds, Log (Odds) and bad rate for the alignment parameters.

Table 14: The SME Application Scorecard

Scaling: 500 = 5:1 with 50 points to double the odds

Score	Goods (G)	Bads (B)	Odds (G/B)	Log(Odds)	Bad Rate (B/G+B)
250	0.5	1	0.46875	-0.758	68.09%
300	0.9	1	0.9375	-0.065	51.61%
350	1.9	1	1.875	0.629	34.78%
400	3.8	1	3.75	1.322	21.05%
450	7.5	1	7.5	2.015	11.76%
500	15.0	1	15	2.708	6.21%
550	30.0	1	30	3.401	3.23%
600	60.0	1	60	4.094	1.64%
650	120.0	1	120	4.787	0.83%
700	240.0	1	240	5.481	0.41%
750	480.0	1	480	6.174	0.21%
800	960.0	1	960	6.867	0.10%
850	1920.0	1	1920	7.560	0.05%

Source: Authors' calculation

5.9. SME Scorecard Implementation

At the point of loan application, the client profile gets scored according to the respective scorecard points allocation of the six variables above. A constant of 500 discussed in the preceding section gets added to the total score of applicants obtained from each of the six drivers of risk. Table 13 shows the alignment parameters linking the total score of individual applications to the scorecard. The scorecard rejects any applications with scores less than 500 and accepts applications scoring 500 points or more.

6. DISCUSSIONS AND CONCLUSIONS

Drawing on knowledge from the developed world, this study adopted a similar approach by developing an application scorecard tailored to SMEs in an emerging market context, but from a single money lending financial institution. The application scorecard developed in this study is set to enable the lenders to quantify the risk associated with SME loan applicants and offer improvements on objective decision-making processes and reduce transaction costs as seen in the developed world, as highlighted in literature (Wendel & Harvey, 2006).

Given that the application scorecard is developed for use on through the door applicants, it was imperative to design a model that reflects the riskiness of SME borrowers. Therefore, to achieve stability and robustness of the estimates and to avoid bias in the scorecard, the development of the application scorecard considered all applications received by the lender, regardless of whether the application was accepted or rejected. Of the accepted applications, some were not taken up due to the issues of cold scoring. The taken up/non taken up model was developed to determine the likelihood of the rejected applications to take up the loan should it have been accepted. Furthermore, a Known Good Bad (KGB) model was developed to assign inferred performance to the rejected applications through the reject inference process. The KGB model tracked the performance of the accepted and taken up population from the point of application to at least eighteen months in performance. This was done to generate the target variable and to assign weights to all the data including accepts and rejects used for the development of the scorecard. The SME application scorecard developed herein can be used to decide whether to extend credit to SMEs with an aim to reduce defaults and serious delinquencies on new applicants. In addition, the model can also be used to allocate capital, determine risk tendency and monitor the performance of SME credit portfolios.

For future research, it is worthwhile incorporating various dimensions of product offerings, secured and unsecured lending, amortising and revolving products, to obtain a more holistic view of the behaviour of SME customers within the bank. Credit bureau institutions such as TransUnion and Experian have access to credit information from various banking and non-banking financial service institutions.

It would be valuable for these bureau to develop application scorecard tailored to the SME credit market in emerging and frontier markets by consolidating this information to improve debt management, risk control and cost effectiveness.

Conflict of interests

The authors declare there is no conflict of interest.

REFERENCES

Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, *23*(4), 589-609. DOI: 10.2307/2978933

Bank for International Settlements. (2012). Basel Committee on Banking Supervision. The internal audit function in banks. *Bank for International Settlements*, *June, 29.*

BCBS. (2017). Basel III: Finalising post-crisis reforms. *Bank for International Settlements*. Retrieved from: https://www.bis.org/bcbs/publ/d424.pdf

Beck, T. (2013). Bank financing for SMEs–lessons from the literature. *National institute economic review, 225*, R23-R38. DOI: 10.1177/002795011322500105

Berger, A., & Udell, G. (1998). The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of Banking and Finance*, *22*(6-8), 613-673. DOI: 10.1016/S0378-4266(98)00038-7

Chimucheka, T., & Rungani, E. C. (2013). Obstacles to accessing finance by small business operators in the Buffalo City Metropolitan Municipality. *The East Asian Journal of Business Management*, *3*(2), 23-29. DOI: 10.13106/eajbm.2013.vol3.no2.23.CITATION

De Noni, I., Lorenzon, A., & Orsi, L. (2007). Measuring and Managing Credit Risk for SMEs: a Quantitative and Qualitative Rating Model. In *Workshop on Visualing, Measuring and Managing Intangibles and Intellectual Capital*. Retrieved from: https://air.unimi.it/bitstream/2434/33546/1/Measuing%20Credit%20Risk%20in%20SMEs_DeNoni_Lorenzon_Orsi.pdf.

Dvorský, J., Schönfeld, J., Kotásková, A., & Petráková, Z. (2018). Evaluation of important credit risk factors in the SME segment. *Journal of International Studies, 11*(3), 204-216. DOI:10.14254/2071-8330.2018/11-3/17

Fantazzini, D., & Figini, S. (2009). *Random Survival Forests Models for SME Credit Risk Measurement.* DOI: 10.1007/s11009-008-9078-2

Laurens, F. (2012). Basel III and prudent risk management in banking: Continuing the cycle of fixing past crises. *Risk Governance & Control: Financial Markets & Institutions*, *2*(3), 17-22. Retrieved from: https://virtusinterpress.org/IMG/pdf/10-22495_rgcv2i3 art1.pdf

Makina, D., Fanta, A. B., Mutsonziwa, K., Khumalo, J., & Maposa, O. (2015). Financial access and SME size in South Africa. *Occasional Paper*, *Dec*, 001-2015. Retrieved from: https://www.researchgate.net/profile/Ashenafi-Fanta/publication/303683234_Financial_access_and_SME_size_in_South_Africa/

links/574d236f08ae82d2c6bc8 e37/Financial-access-and-SME-size-in-South-Africa.pdf

Marimo, M., & Chimedza, C. (2017). Modeling competing risks in the presence of long-term survivors. *South African Statistical Journal*, *51*(1), 199-216. Retrieved from: https://hdl.handle.net/10520/EJC-67c55f18e

Mutezo, A. T. (2015). *Small and medium enterprise financing and credit rationing: the role of banks in South Africa*. Unpublished Doctoral Dissertation. University of South Africa. Retrieved from: http://uir.unisa.ac.za/handle/10500/23479

Njoku, J. E., & Odii, M. A. C. A. (1991). Determinants of loan repayment under the Special Emergency Loan Scheme (SEALS) in Nigeria: A case study of Imo State. *African review of money finance and banking*, 39-52. Retrieved from: https://www.jstor.org/stable /23027508

Pasha, S. A. M., & Negese, T. (2014). Performance of loan repayment determinants in Ethiopian micro finance-An analysis. *Eurasian Journal of Business and Economics*, *7*(13), 29-49. Retrieved from: https://ejbe.org/EJBE2014Vol07No13p029PASHA-NEGESE.pdf

Protection of Personal Information Act No 4 of 2013 (RSA). [POPI]. (2013). Retrieved from: https://www.gov.za/sites/default/files/gcis_document/201409/3706726-11act4of2013protectionofpersonalinforcorrect.pdf

President's Office. (1996). *Act No. 102 of 1996: National Small Business Act*. Retrieved from: https://www.thedti.gov.za/sme_development/docs/act.pdf

SEDA. (2016). The small, medium and micro enterprise sector of South Africa. *Bureau for Economic Research*, *1*, 1-35. Retrieved from: http://www.seda.org.za/Publications/Publications/The%20Small,%20Medium%20and%20Micro%20Enterprise%20Sector%20of%20South%20Africa%20Commissioned%20by%20Seda.pdf

SEDA. (2019). *SMME Quarterly Update 3rd Quarter 2018.* Pretoria: SEDA. Retrieved from: http://www.seda.org.za/Publications/Publications/SEDA%20Annual%20Report%202019%20Final%20Document%20Sent%20to%20parliament.pdf

Valluri, C., Raju, S., & Patil, V. H. (2021). Customer determinants of used auto loan churn: comparing predictive performance using machine learning techniques. *Journal of Marketing Analytics*, 1-18. DOI: 10.1057/s41270-021-00135-6

Visser, K., Hassink, E., Bonsel, G., & Kalkman, C. (2000). Does Anxiety Predict Postoperative Nausea and Vomiting? *ANESTHESIOLOGY-PHILADELPHIA THEN HAGERSTOWN*, *93*(3), SUPP/2), A31-A31. DOI: 10.1097/00000542-200009001-00031

Vrigazova, B. (2021). The proportion for splitting data into training and test set for the bootstrap in classification problems. *Business Systems Research: International Journal of the Society for Advancing Innovation and Research in Economy*, *12*(1), 228-242. DOI: 10.2478/bsrj-2021-0015

Wendel, C. B., & Harvey, M. (2006). *SME Credit scoring: key initiatives, opportunities, and issues* (No. 38955, pp. 1-6). The World Bank. Retrieved from: https://documents1.worldbank.org/curated/en/221231468339556964/pdf/389550AF101Wendel1article01PUBLIC1.pdf

Zeng, G., & Zhao, Q. (2014). A rule of thumb for reject inference in credit scoring. *Math. Finance Lett.,* 2014. Retrieved from: http://www.scik.org/index.php/mfl/article/download/1477/872

МОДЕЛИРАЊЕ КРЕДИТНОГ РИЗИКА МАЛИХ И СРЕДЊИХ ПРЕДУЗЕЋА У ЈУЖНОЈ АФРИЦИ: СТУДИЈА СЛУЧАЈА

1 Мерси Маримо, Афричка експорт-импорт банка, Каиро, Египат
2 Чарлс Кајмедза, Универзитет у Витвотерсрану, Јужноафричка Република

САЖЕТАК

Овај рад је испитао загонетку финансирања процјеном стопе успјешности захтјева малих и средњих предузећа (МСП) за финансирањем код комерцијалних банака. Урађена је квантитативна кохортна анализа о прекорачењу, добијена од једне од водећих финансијских институција у Јужној Африци како би се утврдили покретачи неплаћања. Картица резултата за МСП је развијена коришћењем логистичке регресије на кредитним захтјевима током седмогодишњег периода посматрања у циљу анализирања неиспуњења обавеза као дио управљања кредитним ризиком. Робусност, стабилност и релевантност картице са резултатима апликације је побољшана процесом закључивања одбијања и укључивањем информација бироа. Мала предузећа са пословањем у сектору услуга која имају дугогодишње односе са банком, могу лако приступити изворима финансирања комерцијалних банака. Мала и средња предузећа у грађевинској индустрији са великим бројем кредитних упита показују малу вјероватноћу издржавања строгих услова и критеријума банкарских кредита. Прерогатив главног власника предузећа јесте да поштује своје финансијске обавезе у кредитној индустрији ако жели да се финансира кредитима комерцијалне банке. Њихов кредитни квалитет, што се огледа у информацијама бироа, чини упориште апликација за МСП. Модел развијен у овој студији може се користити као алат за смањење неизвршених обавеза и озбиљних деликвенција приликом прихватања нових кандидата. Штавише, модел се може примијенити за одређивање тенденције ризика и праћење перформанси кредитног портфолија МСП.

Кључне ријечи: *картица резултата МСП, потрошачки кредит, закључак о одбијању, бодовање кредита.*