



**A SYSTEM DYNAMICS MODEL FOR CREDIT RISK
MODELLING AND SIMULATION: THE CASE OF LICENSED
CREDIT REFERENCE BUREAUS IN KENYA**

BY

KANYAMBU MBATHA FLORENCE

15/01360

**A RESEARCH DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN DATA
ANALYTICS AT THE FACULTY OF COMPUTING AND INFORMATION
MANAGEMENT AT KCA UNIVERSITY**

OCTOBER 2021

DECLARATION

I hereby declare that this research project is my original work and has not been presented for academic award in any other institution of higher learning. I hereby also declare that this study contains no material written by others except where reference is made and duly acknowledged.

Student Name: Kanyambu Mbatha Florence

Reg. No: 15/01360

Signed: _____



Date 15/10/2021

I hereby declare that this research project has been submitted for examination with my approval as the assigned University supervisor.

Supervisor Name: Dr. Lucy Mburu

Signature: _____



Date: 15/10/2021

ABSTRACT

Credit risk modelling and analysis are needed in finance and have over many years become active parts of research, motivating statistical modelling. Because of the high numbers of borrowers who fail to fulfil their loan repayments, credit reference bureaus (CRBs) have existed for quite some time especially in countries that have long histories of hosting multinational companies. However, the current standards for regulating how credit risk is quantified have often used assumptions that don't match the reality. This study deals with modelling and estimating through simulating the risk from borrowing activities as it relates to CRBs. Data was collected from annual default reports from the CBK, CRBs and major financial institutions over three years (2018, 2019, and 2020). The study also used focus group discussions to establish baseline levels of default factors. A sample of 12 participants was drawn from the total population of CRB staff members performing the core functions of credit risk determination. The study data was collected using document analysis and Focus Group Discussions (FGDs) to gather historical and current data both qualitative and quantitative. Using the advanced system dynamics approach, the study conducted simulations with starting values from real world scenarios to produce actual measurements of defaulting risk. The model involves analysing the dynamics of common factors which influence the borrower's ability to repay loans. Descriptive analysis was through tabled summaries and bar charts, and explorative analysis applied Causal loop diagrams (CLDs). The simulation was conducted with the aid of graphical output generated from calibration of stock-and-flow diagrams. Through simulations, the study demonstrated how influential parameters of the model are estimated and provided statistical evidence that the model fits the Kenyan CRBs situation better than other often used techniques. The information gained from this study will benefit the government, the Central bank of Kenya (CBK), research scholars and other major financial institutions around the country.

ACKNOWLEDGEMENT

This research proposal could not have been concluded without the help, guidance and contribution of my supervisor Dr. Lucy Mburu, lecturer in charge who owing to her long-standing experience as a scholar, guided me through the entire research process.

Special thanks go to my current employer. I cannot forget to acknowledge the KCA University fraternity which has been of great benefit to my research work. I finally acknowledge my lecturers, especially Mr. Okola, and Dr. Mwendia for their guidance.

In addition, I acknowledge my late loving parents Mr. and Mrs. Kanyambu together with my Bishop, Nancy Mutemi, as well as my sisters and brother who dedicated their assistance through economic support, moral support and with their prayers towards this research.

TABLE OF CONTENTS

DECLARATION	II
ABSTRACT	III
ACKNOWLEDGEMENT	IV
TABLE OF CONTENTS	V
LIST OF FIGURES	VIII
LIST OF TABLES	IX
LIST OF ABBREVIATIONS	X
OPERATIONAL DEFINITION OF TERMS	XI
CHAPTER ONE: INTRODUCTION	1
1.1 BACKGROUND OF THE STUDY.....	1
1.2 STATEMENT OF THE PROBLEM.....	3
1.3 OBJECTIVES OF THE STUDY	5
1.3.1 <i>General Objective</i>	5
1.3.2 <i>Specific Objectives</i>	5
1.4 RESEARCH QUESTIONS.....	6
1.5 MOTIVATION OF THE STUDY	6
1.6 SIGNIFICANCE OF THE STUDY	7
1.7 SCOPE OF THE STUDY	7
1.8 LIMITATIONS AND DELIMITATIONS OF THE STUDY.....	8
1.9 ORGANIZATION OF THE STUDY	8
CHAPTER TWO: LITERATURE REVIEW	9
2.1 INTRODUCTION.....	9
2.2 THEORETICAL FRAMEWORK.....	12
2.1.1 <i>Credit Risk Theory</i>	12
2.1.2 <i>Portfolio Theory of Investment</i>	13
2.1.2 <i>Finance Distress Theory</i>	13
2.3 APPROACHES FOR CREDIT RISK MODELING	14
2.3.1 <i>Descriptive techniques</i>	14
2.3.2 <i>Predictive Data Mining Techniques</i>	14
2.3.3 <i>System Dynamics Modeling (SDM)</i>	17
2.3.4 <i>Related Works</i>	20
2.4 FACTORS THAT INFLUENCE CREDIT RISK	21
2.5 KNOWLEDGE GAPS	23
2.6 CONCEPTUAL MODEL.....	24

2.7 CHAPTER SUMMARY.....	24
CHAPTER THREE: RESEARCH METHODOLOGY	25
3.1 RESEARCH DESIGN	25
3.2 TARGET POPULATION	27
3.3 SAMPLING	28
3.5 RESEARCH INSTRUMENTS.....	30
3.6 VALIDITY AND RELIABILITY OF THE INSTRUMENTS.....	31
3.7 DATA PROCESSING AND ANALYSIS	31
3.9 ETHICAL CONSIDERATIONS	32
CHAPTER FOUR: DATA ANALYSIS FINDINGS AND DISCUSSION	33
4.1 INTRODUCTION	33
4.2 RESPONSE RATE.....	33
4.3 DEMOGRAPHIC INFORMATION	34
4.4 DESCRIPTIVE RESULTS FROM STUDY VARIABLES.....	35
4.4.1 <i>Customer Personal Characteristics</i>	36
4.4.2 <i>Customer Income</i>	36
4.4.3 <i>Creditor Characteristics</i>	37
4.4.4 <i>Issued Loans</i>	37
4.4.5 <i>Performing Loans Non-performing Loans</i>	38
4.4.6 <i>Application Scoring</i>	38
4.4.7 <i>Inquired Funds</i>	39
4.4.8 <i>Collateral Amount</i>	39
4.5 SYSTEM DYNAMICS MODEL OF STUDY VARIABLES	40
4.6 SIMULATION OF CREDIT RISK MANAGEMENT USING STUDY VARIABLES	41
4.7 SENSITIVITY ANALYSIS AND MODEL VALIDATION.....	42
4.7.1 <i>Adjustment of Borrower's Expenditure</i>	43
4.7.2 <i>Adjustment of Borrower's Monthly Remittance</i>	44
4.7.3 <i>Adjustment of Borrower's Income</i>	44
4.8 DISCUSSION OF RESULTS	45
4.9 CONCLUSION	47
CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS	49
5.1 INTRODUCTION	49
5.2 KEY FINDINGS	49
5.3 STUDY LIMITATIONS	51
5.4 RECOMMENDATIONS FOR FUTURE RESEARCH.....	51
5.5 CONTRIBUTIONS OF THE STUDY.....	52
REFERENCES.....	54

APPENDIX 1: PROJECT SCHEDULE.....	61
APPENDIX 2: BUDGET PLAN.....	62
APPENDIX 3: STRUCTURED INTERVIEW GUIDE	63
APPENDIX 4: KEY INFORMANTS INTERVIEW.....	67
APPENDIX 5: MAP OF STUDY AREA.....	69
APPENDIX 6: AUTHORIZATION LETTER FOR DATA COLLECTION.....	70

LIST OF FIGURES

Figure 2.1: Examples of causal loop diagram (a) and stock-and-flow diagram (b)	19
Figure 2.2: Conceptual model.....	24
Figure 3.1: Stages of the dynamic synthesis methodology. Source: rwashana et al., 2009.....	26
Figure 4.1: Focus group responses	34
Figure 4.2: Distribution of profession for the focus group participants	35
Figure 4.3: A causal loop diagram of credit risk for banks and microfinance institutions.....	41
Figure 4.4: Quantitative stock-and-flow model of credit risk.....	42
Figure 4.5: Sensitivity analysis - adjusting borrower's expenditure	43
Figure 4.6: Sensitivity analysis – adjustment of monthly remittance by -35%	44
Figure 4.7: Sensitivity analysis – adjustment of borrower's net income by 25%	45

LIST OF TABLES

Table 2.1: Comparison of simulation-based credit risk modeling techniques.....	18
Table 2.2: Proposed credit risk modeling techniques	21
Table 3.1: The study population: licensed crbs, commercial banks and licensed saccoS.....	28
Table 3.2: Sample size selected from crbs and financial institutions according to neymann allocation sampling technique	29
Table 4.1: Focus groups’ response rate.....	33
Table 4.2: Demographic information for participants	35

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
CAMPARI	Character, Ability, Means, Purpose, Amount, Repayment, Interest
CBK	Central Bank of Kenya
CRB	Credit Reference Bureau
EDA	Exploratory Data Analysis
KBA	Kenya Bankers Association
MIT	Massachusetts Institute of Technology
ML	Machine Learning
NPLs	Non-performing Loans
SPSS	Statistical Package for Social Sciences
UK	United Kingdom
US	United States

OPERATIONAL DEFINITION OF TERMS

Credit

This refers to the level of trust which permits a party to deliver money or resources to another party where the second party does not repay the first party immediately, therefore generating debt.

Credit Reference Bureau

This corresponds to a firm or organization that accumulates information from different sources and provides consumer credit information on entities. Credit reference bureau is a platform that enables employers and loan or credit providers to check one's payment status.

Systems thinking

This refers to an analytical approach that deals with how a systems' constituent parts inter-relate, and how systems work across time within the context of other greater systems.

Artificial Intelligence

This is taken as a field of computing and computer science that highlights the development of intelligent machines that work and reacts like human beings.

Behaviour over time

A reference mode of key variables machine particularly using the system dynamics approach.

Machine Learning

This is a field of artificial intelligence described above, which offers systems the ability to automatically learn and enhance from knowledge without the necessity of getting programmed.

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

The Kenya financial sector has players from banking institutions, microfinance institutions, insurance companies, capital markets, mutual funds, and development finance institutions (Kipngetich, 2019). These institutions play the important role of providing financial and credit services with deep and extensive outreach. Since 2010, strong competition within the lending institutions has preceded a liberalization of interest rates posing serious challenges to formal and informal sectors in terms of high transaction costs, mobilization of funds, grants and the managing of risks (Chelogoi, 2013). Providing credit facilities is one of the core functions of the Kenyan financial sector. However, one of the crucial elements of organizing lending is the risk management problem. The credit function must be effectively managed to ensure equitable distribution of funds while also encouraging plans for liquidity.

Kenya's financial institutions are highly exposed to credit risk since they grant credit facilities as they accept deposits. According to the Basel Committee of Banking Supervision credit risk is the probability of losing an outstanding loaned amount either partially or in full because of a default risk (BCBS, 2001). Financial institutions are exposed to credit risk when borrowers default in honoring their debt obligations on the due date or upon maturity (Muriithi et al., 2016). Since according to Kargi (2011) credit creation is a major income-generating activity for many financial institutions, effective modeling of credit risk during loan processing is crucial for the growth of the financial institutions. Otherwise the credit-availing activity may inevitably result in financial distress.

The supervision annual report showed that the ratio of non-performing loans (NPLs) to gross loans increased from 4.7 percent in December 2012 to 5.2 in December 2013, and further to

5.6 in December 2014 (CBK, 2015). By December 2018 the NPLs ratio had grown to 12.7 percent (CBK, 2018). These supervision reports also indicate a decreased capital adequacy within financial sector, a factor that is calculated as the ratio of total capital to total risk-weighted assets within a specific financial year. The increase in levels of non-performing loans in statements of financial institutions, inefficient loan processing, unwarranted interference with the loan granting process, insufficient and/or absent loan collaterals are among the factors that have been linked to poor and unproductive managing of credit risk and the resulting negative impact on performance of banks and other financial institutions. The deteriorating economic situation that has resulted from the Coronavirus disease 2019 (COVID-19) pandemic also mean that many can only meet their basic needs are unable to pay existing debts (Kansiime et al., 2021). It is therefore important to analyze which are the credit risk factors influencing the financial performance of the institutions in the attempt to modestly contribute knowledge to the existing literature on credit risk.

The credit risk for financial institutions has been studied from many different angles. Song et al. (2016) investigated that whether SMEs that were strongly coupled in the supply chain had a wide availability of equity and debt capital through information sharing. Peura et al. (2016) came up with a new viewpoint by assessing whether trade credit generates benefits to suppliers through a horizontal channel. Su and Zhong (2017) established a credit evaluation model using the principle of variable weight with dynamic data in order to improve the accurate modeling and assessment of credit risk. From the existing literature, researches on financial credit risk mainly focus on credit rating and risk aversion of SMEs. However, the problem of credit risk scoring is not a linear problem, but a highly complex and dynamic problem (Baradaran & Keshavarz, 2017). The country's economy, different personal and business behavior mechanism, claims for compensation, change of interest rates and many other dynamics add to the complexity of the credit risk determination problem.

System dynamics offers a potential solution to modeling of this complex problem. System dynamics is a mathematical and modelling methodology to frame, understand, and discuss complex issues and problems (Rahim, Hawari & Abidin, 2017). Recently the application of system dynamics has been gradually extended to the modelling of credit risk among financial institutions in order to design more effective policies. Studies such as Jiang and Su (2016) and Dong, Hou and Zhao (2017) have explored application of system dynamics to modeling credit risk in the supply chain of small and medium enterprises (SMEs) in China, but these studies were very fundamental and did not explore many factors. In a later attempt, Kaminskyi (2018) modelled the credit risk management of consumer lending institutions in Ukraine with the system dynamics approach and found out that total income was a great factor. This research was also not extensive. Other models such as Discreet Event Simulation, Monte Carlo, Mathematical and Agent Based Modelling which have been used in credit modeling systems have their own limitations starting with the earliest attempts (Blumstein and Larson, 1969; Blumstein, Belkin and Glass, 1971) and lately (Evans, Biles & Bae, 2019).

1.2 Statement of the Problem

The provision of loans is a main concern for banks and other financial institutions given that it carries the risk of non-payment. However, it is impossible for all financial institutions to generate their own techniques for establishing credit risk and they must rely on credit reference bureaus (CRBs) for this. Banks in Kenya have lost a high amount of money due to non-performing loans because the models used by the CRBs have many times failed to accurately predict a potential defaulting customer (CBK, 2018). According to a recent study by the Central Bank of Kenya the gross loans went up by 1.4% from 2,999.47 billion Kenya shillings in December 2020 to 3,040.45 billion Kenya shillings in March 2021 (CBK, 2021).

At the same time, non-performing loans increased in five sectors namely Personal and Household, Transport, Real Estate, Trade, Communication and Tourism. The same report indicated that increased non-performing loans were mainly caused by the challenging operating environment in the face of the COVID-19 pandemic. More than half of the survey respondents also stated that they expected the number of non-performing loans to increase during the second quarter of 2021 because of the containment measures that the Kenyan government put in place in March 2021.

Traditionally, such CRBs have applied static models with demographic factors or factors that do not capture dynamics of the real-world credit risk patterns. Economic factors which heavily affect the ability or inability of customers to repay their loans do not exist independently of other factors, such as political fluctuations, environmental changes and behavioral changes. Models that accommodate these factors and their changes are very limited, with most of them having been developed for first-world countries with stable socio-economic and political environments (Moradi & Rafiei, 2019). The modeling of credit risk within developing economies with non-static factors is rare.

According to Zhu et al. (2019), most models of financial credit risk have been designed and developed through trial and error and lack a proper theoretical framework. Because of this, the models are not tried and tested, and are unable to perform efficiently in times of economic or political crisis. According to Csikosova et al. (2019) many prediction models show high prediction capability, but only in perfect conditions thus they cannot be effectively used in developing countries that face a lot of socioeconomic uncertainties. Additionally ARIMA models (autoregressive integrated moving average models) which are commonly used in financial situations assume normal distribution, but large gains and losses always exhibit higher probabilities than the normal distribution can model (Petrică et al., 2016). Therefore, such models might be able to function well under optimum financial conditions, but the lack

of responsiveness to the evolving environment renders these models inefficient, particularly in the event of concept drifts. An example of such a concept drift is where a certain proportion of customers with no default history fall into default due to unexpected circumstances. Dynamic models that can easily accommodate new factors, changing factors and uncertain factors have not been developed for most CRBs in Kenya.

In order for models to be flexible to a variety of socio-political and economic factors and to generate results which are compatible with real-life situations, such models must be able to model the delays within feedback structures that sometimes do not manifest themselves until policies have changes or intervention measures have been revised (Langarudi, 2016). System dynamics models that take into account the feedback from multiple factors and the delays in these feedbacks have been used for modeling various types of risk, but their use in modeling credit risk for CRBs is limited. This is why the modeling of credit risk management that takes account of dynamics in risk management and the delays in feedback structure is central to this study.

1.3 Objectives of the Study

1.3.1 General Objective

The general objective of this study was to apply the system dynamics approach to model and simulate credit risk for licensed credit reference bureaus in Kenya.

1.3.2 Specific Objectives

The main objective for the study has been achieved through the following specific objectives;

- i) To determine the factors that influence credit risk by individuals and businesses in Kenya.

- ii) To design and develop a system dynamics model for credit risk determination
- iii) To test and validate the credit risk determination model developed above.

1.4 Research Questions

- i) Which are the factors that influence credit risk by individuals and businesses in Kenya?
- ii) How can these factors be modeled using the system dynamics approach?
- iii) Is the developed credit risk determination model acceptable to the stakeholders?

1.5 Motivation of the Study

This study was mainly motivated by the constantly increasing rates of non-performing loans and the frequent failure of CRB models to accurately model the non-payment probability of borrowers. Factors which did not exist before the COVID-19 pandemic such as lockdowns, curfews, closure of businesses, etc. have completely changed the dynamics of any debt repayments (Wanjala, 2020). The failure of Kenyan CRBs to predict non-payment has far-reaching consequences on sustainability of financial institutions and generally on economic growth and provision of loan facilities to future borrowers. Therefore, this study adds to the missing literature in the dynamics that drives the ability of borrowers to repay their loans. It is also hoped that the findings in this study will provide useful insights and foresight into policy matters concerning credit risk determination for individual borrowers as well as businesses.

1.6 Significance of the Study

Credit scores are used by lenders in order to know who to lend to and how much credit they should be given. The information contained in credit scores is important to governments and banking associations which play a role in the formulation of monetary and fiscal policies.

This study is important to financial institutions and CRBs to their monetary management strategies and credit allocation. This is important to investors, bank managers and investors in general.

Key policymakers can also make use of this study to make them understand how credit is important to economies and institutions and what levels of credit are permissible. ML has also made scholars understand analytics better and come up with more efficient business models. This study could thus be used by other academicians as a point of reference.

1.7 Scope of the Study

This research was based on selected Commercial banks in Kenya, under the guidance of Kenya Bankers Association (KBA), selected credit reference bureaus, and the policy guidance of the Central Bank of Kenya. In addition, this study established the role of machine learning model on credit risk prediction among licensed credit reference bureaus in Kenya within the time scope of June 2019 up to May 2020.

The research was also limited to deducing how the three CRBs mitigate efficiently against loan defaults among all the forty-two banks in Kenya, with their main headquarter in the County of Nairobi city. The researcher believes that these participants provide an adequate population and sample for the study and therefore give reliable results and findings.

1.8 Limitations and Delimitations of the Study

It is known and accepted that some respondents can withhold information because of the confidential nature of the subject matter. This setback can be effectively mitigated by the researcher offering assurances during the research process stating that the information collected was purely for academic purpose and thus not to be offered to unauthorized outside parties. The research may also look to address the limitation by looking into other empirical data collected from past research.

1.9 Organization of the Study

The research consists of five chapters. Chapter one offers a general introduction to the study, chapter two conducts a critical review of related studies, and also explores research gap estimates. The chapter three articulates the methodology used in the research. This consists of methods of investigating the research design, target populations, sample and sampling procedures, data collection as well as tools used in order to acquire all this information. Chapter four is about data analysis, how it is presented and the explanation behind the data collected together with the created model. Chapter five is focused on summarizing the data gathered into a broad spectrum of information which concludes by the research giving recommendations.

CHAPTER TWO

LITERATURE REVIEW

This chapter reviews the literature materials on credit risk prediction and provide a theoretical background. It provides a discussion of existing research on theories, models, approaches and variables. The chapter further identifies the gaps on emerging issues associated with the research that have not been exhaustively studied from the literature. It then conceptualizes a theoretical framework from which the methodology has been designed.

2.1 Introduction

It is unavoidable that banking and other financial institutions continue taking risks in issuing loans to customers because the availability of credit and the resultant interest rates both drive the economy (Moradi & Rafiei, 2019). Nevertheless, there exists a delicate balance because a bank's survival depends on how it takes appropriate risks. A no-risk financial institution is just as vulnerable as the one that takes heavy risks (Mandalaa et al., 2012). When financial institutions are faced with risks, effective risk management depends on how they identify, understand, measure and provide the appropriate risk management strategies (Bekhet & Eletter, 2014).

Within the finance sector, credit risk is a risk that banks face during resource allocation, defined as the probability of nonpayment or delayed payment by customers. Customers with an increased probability to repay a loan are classified as "good customers" while customers with an increased probability to default are classified as "bad customers" (Akkoc, 2012). Modeling credit risk assessment is an approach that is crucial for financial institutions because it allows these institutions to ensure that customers are able to pay their installments,

i.e. be classified as “good customers” before allocating a loan facility to them (Moradi & Rafiei, 2019).

When the Kenya banking industry was liberalized in 1992, this activated the beginning of highly intensive rivalry among commercial banks and other institutions that offer finance products that led to extensions of large amounts of credit to customers with the aim of increasing profitability (Kithinji, 2010). A good number of the awarded loans belonged to the category of “political loans” granted with very limited credit assessment. Many other loans were granted to insiders on a know basis. The poor loan issuing criteria led to accelerated levels on non-performing loans, subsequently eroding profitability of banks. It is during this crisis that many commercial banks started to adopt different credit risk management policies and strategies. Variations in these policies were based on criteria such as:

1. Ownership of the financial institution (e.g., private, foreign, government-influenced or locally owned)
2. Credit management policies of the financial institution
3. Available credit scoring systems
4. Caliber of management for the financial institution
5. The regulatory environment of the financial institution

It soon became evident however, that even though the financial institutions could have the best credit management policies, they did not necessarily record high profitability. Additionally, although there had emerged industry standards regarding what could be termed as a “good credit policy” or “bad credit policy”, there existed high variation in the detailed characteristics for different banks (Mercylynne & Omagwa, 2017). This meant that the market could regard the poor performance of individual’s bank more leniently when the entire financial sector was hit with adverse liquidity crisis. Additionally, banks could

sometimes be forced to adjust their various credit policies in line with other banks in the market in a “herd behavior” when this was totally uncalled for.

With the emergence of credit reference bureaus (CRBs) in Kenya, credit lending processes have been significantly transformed and improved the performance of many financial institutions (Munene et al., 2018). A CRB is an information broker that provides dependable, appropriate and comprehensive data on repayment habits and current credit levels of loan applicants (Brown et al., 2009). This data is consolidated from creditors of all forms and packaged into individual-report style for distribution to creditors at a fee. CRB services are particularly useful to any monetary system because of the information irregularity between creditors and their customers, particularly in the face of competition between financial institutions for customers and cases of multiple (sometimes synchronized) borrowing which can increase loan default probability in the absence of a central database that captures relevant aspects of customers’ borrowing behaviors. According to Nasieku and Ngugi (2016) CRBs can contribute substantially to a reduced cost of screening the many loan applications submitted to a financial institution by allowing the creditor to efficiently identify prospective borrowers that have default histories with other financial institutions.

While CRBs help to make quicker and more precise decisions, Munene et al. (2018) found out in a survey study in Kenya that some banks thought CRBs were not effective in reducing credit risk. The Annual Supervision report by the Central Bank of Kenya also confirms that non-performing loans (NPLs) are on the rise, despite the presence of CRBs. Upon examination of the emphasis that is placed on credit risk modeling and management by financial institutions and CRBs in Kenya, the contribution of this management to institutional profitability is yet to be achieved (CBK, 2018). This begs the question “Is the practice of credit risk management really beneficial to the Kenyan commercial banks and financial institutions?” If it is, then there should be a significant contribution to profitability and

enhancement of shareholder value. One of the key reasons for the relative ineffectiveness of CRBs is the uncertainties and dynamism of borrowers and their environment that cannot be captured using static data that is stored in the databases. Hard grounding on theoretical frameworks to understand credit risk and the availability of models which capture uncertainties and dynamism must come into play to increase the accuracy with which CRBs can identify credit risk.

2.2 Theoretical Framework

2.1.1 Credit Risk Theory

Credit risk theory puts forth that a default event or nonrepayment arises from the evolution of a firm's asset, which can be modeled as a diffusion process with constant parameters (Melton, 1974). This theory has its basis on the variables relating to a specific issuer of a credit facility. The development of this model can be seen by today's asset models where the loss that is conditional on a default is exogenously specific. In such models, the default event could occur throughout the lifetime of a corporate bond rather than simply upon maturity (Longstaff & Schwartz, 1995).

Before the formulation of Credit Risk Theory, earlier research on credit risk used the conventional actuarial methods for computing credit risk, with the main limitation being a complete dependence on historical data. Until the theory formulation there had emerged three quantitative approaches for modeling credit risk; the structural approach, reduced form appraisal approach, and incomplete information approach (Crosbie & Bohn 2003). The credit-risk theory heavily improved the three approaches and was a major milestone to the modeling of financial credit risk.

2.1.2 Portfolio Theory of Investment

This theory developed by Markowitz (1952) attempts to maximize portfolio expected returns for an allotted amount of portfolio risk, or equivalently, to minimize risk for a given amount of expected return through selection of available assets. Portfolio theory has advanced the mathematical modeling of finance since its conception through to the 1970s. When purporting this theory, Markowitz (1959) further considers that constraints and judgement need to play essential roles in portfolio choice. He adds that when there is no properly selected function used to conduct expected-utility assessment, the next best alternative to perform a risk-return analysis. Nevertheless, the theory was heavily criticized by later researchers as being unrealistic since financial returns rarely follow Gaussian distribution or any symmetric distribution like the function associated with portfolio theory (Sproul, 1998).

2.1.2 Finance Distress Theory

This theory developed by Baldwin and Mason (1983) purports that when an entity deteriorates to the point of not being able to meet the set financial obligations, the entity is said to have come into the state of financial difficulty. The theory further explains that early signs of financial distress include violation of debt repayment and the failure or reduction of dividends pay-outs by a business. The relevance of financial distress Theory can therefore be quantified by the liquidity and credit risk faced by a borrowing entity. In the case of banks and other financial institutions this is the inability to avail cash to depositors and loans to borrowers upon demand leading to a liquidity crisis. In the case of businesses, it is the inability to pay employees and meet liquidity needs. Managing the loan portfolio is critical for determining the liquidity of a firm or other business entity. In advancing this theory, Whitaker (1999) defined entry into financial distress as the initial year that cash flows become fewer than current maturities of any long-term debts. With this information, a

business has enough cashflow to repay its creditors for as long as the cashflows exceed the current debt obligations. Therefore, the primary indicator of an entity's financial distress is failure to meet contractual debt commitments. Nevertheless, substantial effects of financial distress can be present before defaulting starts, and an entity can enter into financial distress due to economic distress, performance decline or poor management particularly with respect to risk management (Wruck, 1990). Furthermore, Boritz (1991) depicted a course of financial distress which began with an incubation interval typified with a set of poor economic conditions and the costly mistakes of poor management.

2.3 Approaches for Credit Risk Modeling

Analysis of credit risk is a pattern-recognition problem which includes functions to predict the possibility that a customer will repay a loaned amount (Emel et al., 2003, Kruppa et al., 2013). Consequently, the key features of any credit risk models are resolution and accuracy. Several methods exist for credit risk modeling and analysis which are discussed below:

2.3.1 Descriptive techniques

Techniques for descriptive analysis provide summarization of data to lay the analytical foundation for data preprocessing. Such techniques can range from min-max, mean, standard deviation and variance calculations to the provision of graphical output in the forms of tables, charts and graphs (Yu et al., 2008). These techniques are very important for identifying the factors to place focus on, but they are very ineffective for detailed assessment and modeling of credit risk over short term, medium term and long term.

2.3.2 Predictive Data Mining Techniques

Recent advancements in information technology have greatly improved data collection and analysis procedures and the demand for credit risk classification using real data has increased.

It is known that machine learning and data mining are some of the most popular analytical techniques. Systems like artificial intelligence systems that reveal the patterns in complex data are known as data mining systems (Zanin et al., 2016). Data mining has become indispensable for decision support. Data mining system and play an essential part in segmenting the market, providing customer services, detecting fraud, and also in credit and behavior scoring, and benchmarking. In this information age, individuals and organizations are producing large volumes of data every day (Berry & Linoff, 2000). Such discovery of useful knowledge from databases and transformation of this information into actionable results has been major challenges for financial institutions and credit service providers.

Common examples of predictive data mining techniques used for credit risk scoring include naïve Bayesian classifiers, classification trees and regression.

The Naïve Bayesian classifier (NB)

This classifier originates from the Bayes theorem and believes that the impact of an attribute value on a given group exists independently of the other attributes, an assumption known as “class conditional independence” (Hand & Yu, 2001).

Bayesian classifiers are valuable because they offer theoretical support for other classifiers such as support vector machines (SVMs) which don't explicitly utilize Bayes theorem. However, NB has the major limitation that to simplify computation, the predictive accuracy correlates highly with highly with the notion of class conditional independence. In practical application, however, dependencies can be present between variables.

Classification trees (CTs)

CTs represent binary classifications of problems with the top-most node being the root node and leaf nodes representing classes. Every internal node represents a test case on an attribute and each branch is equivalent to a result of the test. CTs generate a classification of observations based on each dependent variable and supervised by the values of independent variable or variables (Addo et al., 2018). The process of splitting outcomes is usually conducted using a single dependent variable at a time. Classification trees function by minimizing the level of impurity, i.e., the amount of variability of the outcomes on the observations.

The key strength of CTs is their ability to yield simple classification rules and deal with the nonlinear and interactive impacts of variables. However, their limitation is the difficulty in adapting a tree-like structure designed for a specific context and some other generalized contexts. The sequential natures as well as the algorithmic complexity cause CTs to be highly dependent on the observations such that a very small change in the observations can alter the structure of the tree. This static nature of modeling is not effective for dynamic problems such as customer default rate.

Regression

Regression analyses present a powerful statistical approach for modeling the association between two or more variables. There exist different types of regression analyses, but fundamentally they all examine how a single or many explanatory variables influence the values of a dependent variable (Chen et al., 2012). Two of the most common regression techniques used for credit risk modeling are ordinary least squares (OLS) regression (Kargi, 2011) and logistic regression (Blanco et al., 2013). OLS regression treats the dependent variable as continuous. Logistic Regression which is more popular for modeling credit risk

considers a binary outcome of 0 or 1 (e.g. default/no default). The result is therefore interpreted as the possibility of defaulting, and the variable's coefficients are interpreted independently to assess the usefulness of every variable in the credit-decision model (Rudd et al., 2017). While regression can easily model dynamic situations in credit risk modeling, it is not very efficient for modeling uncertainty.

2.3.3 System Dynamics Modeling (SDM)

One of the ways to take care of dynamism and uncertainty in customer behavior while also modeling feedback structures and delays in these feedbacks which can cause repayment behavior to change during a loan's lifetime and after maturity is through SDM. SDM is an approach that uses stocks- and flows-diagrams, internal feedback loops, and time delays for acquiring knowledge and understanding the non-linear nature of complex systems over a period of time (Rahim et al., 2017). As per the terminology of system dynamics, a *system* is as a collection of interconnected parts that interact over time for a common purpose to form a unified whole. The "Systems" approach is useful for simplifying complex problems, identifying significant processes and establishing controls (Reuter, 2013). *Dynamics* refers to behavior over time. *System dynamics* is a computer aided approach for assessing and redesigning policies. It is used to understand how systems change over time.

System Dynamics models can be applied over what are thought to be "data-poor" problems. The information basis for the conceptualization and design of system dynamics models is much wider than the numerical database used in operations research and statistical modeling resulting to better performing models. Table 2.1 below shows the attributes that make systems dynamic the methodology of choice in this research. Elements of SDM comprise of causal loop diagrams (CLDs), stock and flow diagrams, and simulation graphs.

Table 2.1: Comparison of simulation-based credit risk modeling techniques

Attributes/ Model	Uncertain	Dynamic Phenomena	Feedback Loop	Non-linear	Source
Discrete-event simulation	✓	✗	✓	✗	Juneja (2017)
Agent based simulation	✗	✓	✗	✓	Teply & Klinger (2019)
System-dynamics simulation	✓	✓	✓	✓	Dong et al. (2017)

Causal loop diagrams are an important tool for representing the feedback structures of systems. They help to 1) construct and test hypotheses about the causes of dynamics; 2) elicit and describe mental models of individuals and teams; 3) communicate the important feedback processes believed to be responsible for a problem (Sterman, 2019). Causes, effects and feedbacks are depicted and visualized through CLDs. Therefore, CLD can be crucial for representing a socio-economic ecosystem, political system or mechanical system from a mental model. Sterman, (2019) noted that a CLD can present a simple visual representation of a system that corresponds to phenomena about complex issues in a system together with their components and interactions. It captures system structures, feedback loops, delays in feedback, causes and effects. The CLD thus provides a visual illustration that can be applied for communicating information about changing phenomena with clarity therefore promoting better understanding.

Stock-and-flow diagrams extend CLDs by modeling the rate of change among variables as how changes in one related variable affect change in another variable over time. This uses a type of variable called a *stock, level of accumulation* whose values flow into another variable

at a predetermined rate using another variable called a *flow* or *rate*. This allows the model to be calibrated by changing initial values for the stock or the amount of flow into or out of a stock and conducting a simulation. Results of a simulation can then be mapped using graphs (Reuter, 2013). Figure 2.1 shows the differences in credit risk modeling studies that have used a simulation approach.

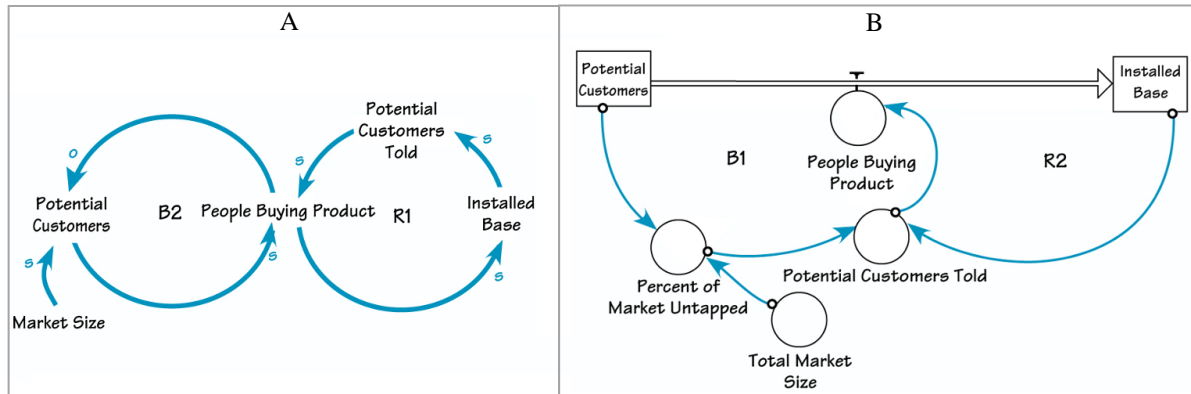


Figure 2.1: Examples of causal loop diagram (A) and stock-and-flow diagram (B)

Systems' thinking is understood to be a language for understanding and communicating the complex, interdependent issues that managers encounter on a day-to-day basis. Within the systems thinking framework, CLD can be seen as a group of sentences created through identifying the primary variables in a system, and showing the causal connections between them using links. By combining together several loops, one creates a concise story relating to a particular problem or issue (Lannon, 2018).

Littlejohns et al. (2018) indicated that complex systems have many nested and heterogeneous system components that demonstrate a lot of variation, and that element is a system in its own right. Many systems are complex because they are made up of multiple entities, organizations, groups and segments at local, regional, state, national and global levels, all of which differ structures, functions and interests. Every element in the system possesses a unique relationship to- and influence on the complete system. Self-organization is a notion

for describing how systems adapt to their environments, and for studying the way that systems organize, change and/or innovate. In general, feedback loops are the interconnections that demonstrate self-organization in complex systems. The behavior of complex systems is to a great extent the cumulative impact of positive and negative feedback mechanisms. The different approaches in SDM can be applied depending on each case and the type analysis that is appropriate under circumstances. The complexity of a problem can also determine the best approach to model that gives an accurate abstraction of reality that eventually giving better simulation results that can be acceptable to users. In this research we adopt system dynamics methodology since it is best suited to analyze qualitative data relating to dynamism, uncertainty and feedback delays in modeling a customer's loan repayment behavior over time.

2.3.4 Related Works

Many research studies have attempted to solve the problem of financial and credit risk assessment using different approaches and methods that have different levels of accuracy as shown in Table 2.2. From the table, most of the approaches that are traditionally used for modeling credit risk are either static in nature, i.e. they do not consider the variability of variable, or they function optimally only in the absence of uncertainty. Once uncertainty of prediction is introduced, these models can no longer produce dependable outcomes. Unfortunately, the credit environment is surrounded by a lot of uncertain scenarios which can lead to incorrect estimation of default risk. As shown in Table 2.2, very few studies exist that have incorporated both dynamic modeling and the consideration of uncertainty. One of these studies has however used a very small dataset (Kaminskyi, 2018) and the other (Dong et al., 2017) has focused on supply chains rather than financial institutions.

Table 2.2: Proposed credit risk modeling techniques

Data	Techniques	Constant / Dynamic	Type of predictor	Source
Kenyan banks	Descriptive analysis, analysis of variance (ANOVA)	Constant	Certainty	Kithinji (2010)
Nigerian banks	Descriptive analysis, correlation analysis, regression techniques	Constant	Certainty	Kargi (2011)
Ghana rural banks	Panel regression	Dynamic	Certainty	Afriyie et al. (2011)
Credit records for Germany	Support vector machine (SVM) classification, Fuzzy clustering	Constant	Certainty	Ghodselahe (2011)
Local bank in China	k-Means clustering, SVM classification	Constant	Certainty	Chen et al. (2012)
Nigerian commercial banks	Augmented Dickey Fuller (ADF), Pairwise Granger causality	Dynamic	Certainty	Onaolapo (2012)
Firms in Tehran stock exchange	Neural networks analysis, Self-Organizing Map	Constant	Certainty	Abdipoor et al. (2013)
Microfinance institutes in Peru	Multilayer perceptron approach (MLP), linear discriminant analysis, quadratic-discriminant analysis, logistic regression.	Dynamic	Certainty	Blanco et al. (2013)
China banks	Kernel, fuzzification, penalty factors	Constant	Certainty	Zhang et al. (2014)
Iran bank	Genetic algorithm (GA), SVM, Principal components analysis (PCA)	Constant	Certainty	Koutanaei et al. (2015)
India Banks	Panel data analysis	Dynamic	Certainty	Maji & Hazarika (2016)
Kenya microfinance banks	Multiple regression	Constant	Certainty	Otieno et al. (2016)
Simulation of Kenyan credit	Semi-Markov modeling	Constant	Uncertainty	Wagacha & Othieno (2016)
China supply chain	System dynamics modeling	Dynamic	Uncertainty	Dong et al. (2017)
USA default credit card clients	Naïve Bayesian classifier, Decision trees	Constant	Certainty	Niloy & Navid (2018)
Ukraine Banks	System dynamics modeling	Dynamic	Uncertainty	Kaminskyi (2018)

2.4 Factors that Influence Credit Risk

There exist a wide variety of factors that can be modeled when determining credit risk. The following are some of the most common factors according to research (Kithinji, 2010; Otieno et al., 2016; Juneja, 2017; Kaminskyi, 2018; Teply & Klinger, 2019)

Customer Personal Characteristics: this includes factors that are specific to the borrower, such as age, gender, educational background, lifestyle, etc.

Customer Income: this is a generalized quantification of the net total amount per month that the customer receives from multiple sources or the average income per month receive over a predefined period such as 3 or 6 months. This directly affects the ability of the customer to repay a loaned amount.

Creditor characteristics: This includes factors such as flexibility of the lender to non-payment behavior by the customer, the amount of cushioning that the lender can obtain from insurance against bad loans, the financial liquidity of the lender, etc.

Issued loans: this is the total number and amount of active loans issued to the customer by multiple financial institutions.

Performing loans: This is a factor of issued loans that only includes the number and total amount for loans that have no default

Non-performing loans: This is a factor of issued loans that only includes the number and total amount for loans that have default

Application scoring: This is a parameter calculated external to the model by a bank, financial institution or CRB that classifies the customer on a “good” or “bad” scale.

Inquired funds: This is the total amount requested by the customer.

Collateral Amount: this is the total value of collateral that has been pledged by the customer during the loan application.

External factors: This corresponds to factors that are external to both; the creditor and borrower, such as the nation’s gross domestic product (GDP), the inflation rate,

environmental calamities, political stability, etc.

2.5 Knowledge Gaps

As has been determined by Kruppa et al. (2013) the problem of determining whether a customer will repay a loaned amount or not is a very complex problem and must employ high resolution and accuracy of prediction. This includes taking into account the variable nature of observations and factors influencing these observations, as well as the dynamism of the environment in which the lender and the customer operate. As shown in Table 2.1, most of the existing approaches that are traditionally used for modeling credit risk are either static in nature, i.e. they do not consider the variability of variable, or they function optimally only in the absence of uncertainty. Once uncertainty of prediction is introduced, these models can no longer produce dependable outcomes. Unfortunately, the credit environment is surrounded by a lot of uncertain scenarios which can lead to incorrect estimation of default risk. To the knowledge of the author, very few studies exist that have incorporated both dynamic modeling and the consideration of uncertainty. One of these studies has however used a very small dataset (Kaminskyi, 2018) and the other (Dong et al., 2017) has focused on supply chains rather than financial institutions. Furthermore, all these studies have been carried out in developed economies which are very different in operations and categorization of risk from developing countries.

In trying to fill the identified gaps, i.e. the static nature of credit models and their inability to account for uncertainty, this study contributes knowledge that will increase the accuracy for modeling credit risk for financial institutions in Kenya and places with similar conditions.

2.6 Conceptual Model

Using knowledge advanced from the examination of previous studies, the study has adopted the following conceptual model.

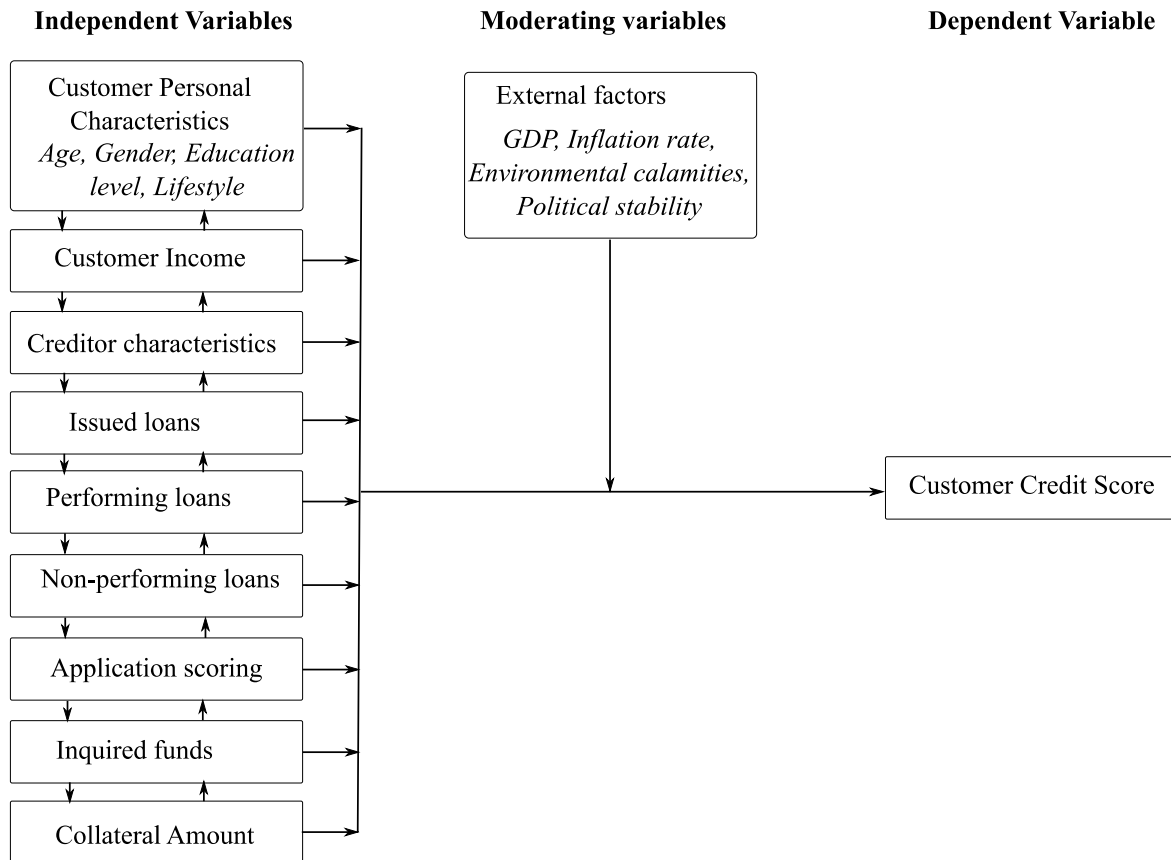


Figure 2.2: Conceptual Model

2.7 Chapter Summary

This chapter has presented a review of the related literature on credit risk modeling, information that can be used by various researchers, scholars, authors and analysts. After establishing the extent of practice and the gaps therein, the chapter has conceptualized a model of dependent and independent variables relying on knowledge from existing works. This conceptual framework guides the methodology chapter that follows to map out the analytical framework.

CHAPTER THREE

RESEARCH METHODOLOGY

This chapter outlines the methods and procedures that the research has applied. It describes the research design, target population and sample, study instruments, data collection procedures and data analysis.

3.1 Research Design

The study has adopted dynamic synthesis methodology (DSM) proposed by Rwashana et al. (2009) as the research design. DSM outlines 6 stages to be followed throughout the study process as shown in Figure 3.1. Stage 1 involves fact finding through review of literature and company records, identification of stakeholder involved in the credit risk determination process in Kenya and the explorations of the problems faced by the identified stakeholders. Stage 2 involves focus group discussions (FGDs) from a sample of identified stakeholders. Stage 3 involves data collection, model design and building. Causal loop diagrams and stock and flow diagrams have been constructed based on the identified variables. Stage 4 involves applying data from the field survey to conduct empirical assessment of the system dynamics model. Stage 5 involves testing and validating the system dynamics model using insight from the real-life context in stage 4. The final stage (6) involves analysis of the existing policies using knowledge obtained from the research experiments, and discussing the results obtained in terms of the policy implications for credit risk assessment in Kenya.

This study expects that the capabilities of a system dynamics model efficiently facilitate decision making processes. Applying a computer-based model in the resolution process alters the character of the adversarial exchange. Decisions made by experts compared to those arrived at by using system dynamics models can be as good as or even better than the experts.

However, the results can be different when human judgement and expertise is required to conceptualize and develop a model (Abotaleb & El-adaway, 2018).

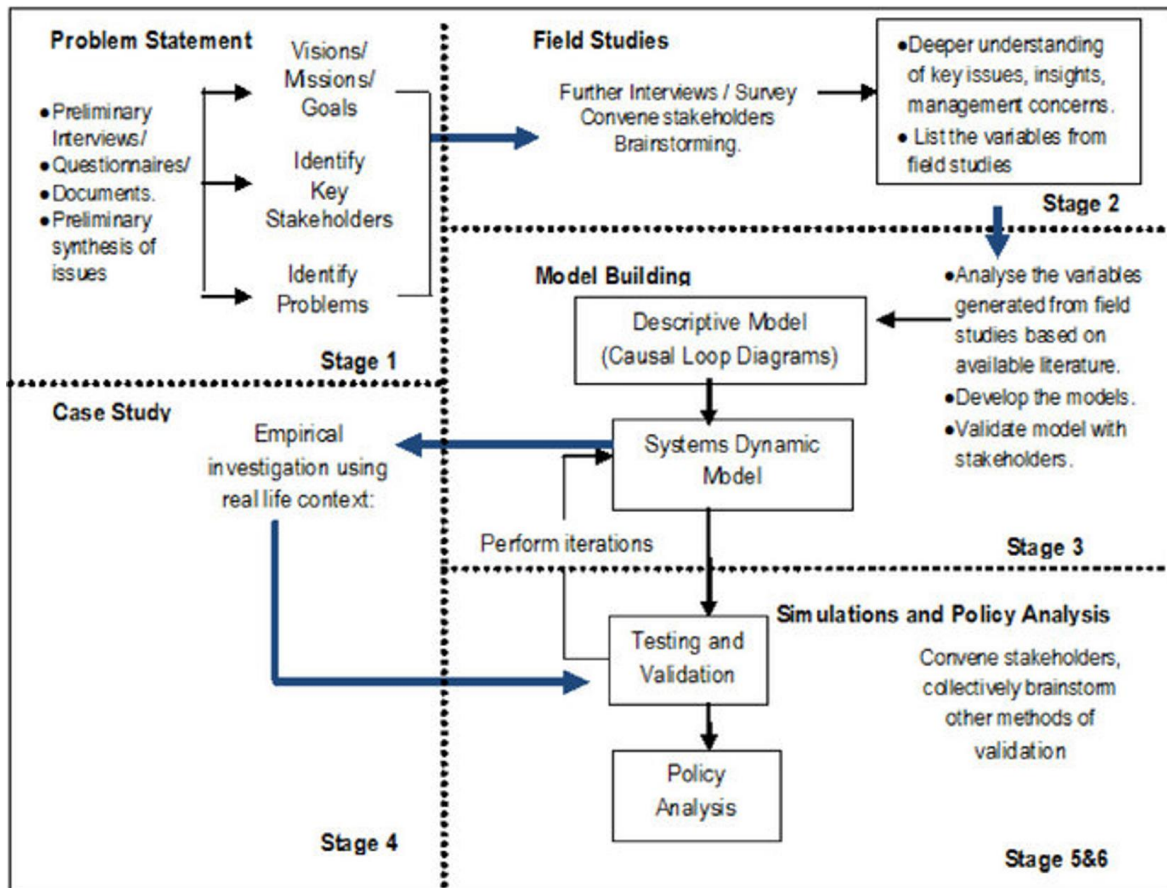


Figure 3.1: Stages of the Dynamic Synthesis Methodology. Source: Rwashana et al., 2009

Through field studies SD was used to provide a descriptive model arising from data collected in its natural environment giving the necessary fit for an abstraction from which the SD conceptual casual loop diagram was constructed. A combination two research methods; qualitative and quantitative have been applied to increase the accuracy and performance of the model and subsequently to strengthen results through cross validation with the case study results.

Causal loop Diagrams were used to develop feedback structure model. Nodes and edges diagrams were then transformed to stock-and-flow diagrams which represented the quantitative model of the system under development or review. Mathematical association

between and amongst the factors allowed the simulation of the prototype to be defined after which simulations of the key variables were run. Both the causal-loop diagrams and the stock and flow diagrams were created using Vensim PLE Student Version 8 which is freely available.

3.2 Target Population

The targeted number for the focus group discussions consisted of a total of **29** staff members that are involved in core functions from each of the five sections; CRBs, Tier 1, 2 and 3 commercial banks and SACCOs as outlined in Table 3.1. The **29** staff members were either working in the registries as analysts or decision makers which are the core functions of credit risk assessment.

A population refers to a complete set of people, cases or objects with some common perceptible characteristics (Mugenda & Mugenda, 2003). Target population in statistics is the specific population about which information is desired, thus the inclusion criteria of this study included all licensed credit reference bureaus and all the registered commercial banks in Kenya according to Central Bank of Kenya annual supervision report (CBK, 2018). There are currently 3 licensed CRBs, 36 commercial banks and 164 licensed SACCOS in the country as of January 2020 to study from as the diagram of Table 3.1 shows. The intention was to study all of them, but due to the limitations in resources for collecting primary data, a random sampling strategy was employed for the study.

Table 3.1: The study Population: Licensed CRBs, Tier 1, 2 and 3 commercial banks and licensed SACCOs

Entity	Population
Licensed CRBs	Metropol Credit Reference Bureau Limited, Credit Reference Bureau Africa Limited (TransUnion) and Creditinfo, Credit Reference Bureau Kenya Ltd.
Tier 1 Commercial banks	Kenya Commercial Bank, Equity Bank, Barclays Bank, Stanbic Bank and Cooperative Bank
Tier 2 Commercial banks	Citibank, National Bank(NBK), Bank of Africa, Family Bank, Diamond Trust Bank(DTB), Ecobank, I&M Bank, Housing Finance, CFC Bank, Bank of India, Amsa Bank, Prime Bank, Bank of Baroda(BoB), and Guaranty Trust Bank
Tier 3 Commercial banks	SBM Bank, Development Bank(DB), Paramount Universal Bank(PUB), ABC Bank, Jamii Bora Bank, Consolidated and Development Bank(CDB), Credit Bank, Spire Bank, Guardian Bank, Victoria Bank, Middle East Bank, Giro Bank, Oriental Commercial Bank(OCB), Habib Bank, First Community, UBA, Gulf Bank and Consolidated Bank
Licensed SACCOs	All 164 registered SACCOs

This study picked sample information personnel per category of financial institution resulting in **29** participants for the study as outlined in Table 3.2. The fraction of the targeted population that possesses the characteristics to be measured constitutes the information credit managers in charge who work in the headquarters of the selected banks.

3.3 Sampling

The sample size of **29** respondents from the 5 sub-categories was drawn from a population of 203 entities, i.e. 200 financial institutions and 3 CRBs. The respondents were staff members

involved in core functions by adopting Nasiurma’s Model (2000). According to Nasiurma’s Model: sample size n is calculated as:

$$n = (Ncv^2) / (cv^2 + (N-1) e^2)$$

Where; N is the population, cv is coefficient of variation taken to be (0.5) and e is tolerance level at 95% confidence level (0.05). This is 14% of the entire population of financial institutions and CRBs.

Table 3.2: Sample size selected from CRBs and financial institutions according to Neymann Allocation sampling technique

	Target Population	Sample size	Sample size as % of population
Licensed CRBs	3	1	33%
Tier 1 Commercial banks	5	1	20%
Tier 2 Commercial banks	14	3	21%
Tier 3 Commercial banks	17	4	24%
Licensed SACCOs	164	20	12%
Total	203	29	14%

According to Eeuwijk & Angehrn (2017) the average size of a Focus Group (FG) should range between 6 and 12 participants involved in a focus group discussion (FGD) at any one time. However, smaller groups are more informative because they give all participants enough time and opportunity to share. According to Mishra (2016) too little or too many contributors at one time may not generate valid or reliable data. Therefore, within the selected categories Neymann Allocation sampling technique was seen to be appropriate for this research because it allowed for selection of interviewees within each FG by considering diversity in seniority, experience and area of specialization thus reducing bias while maintaining an acceptable number of participants. The study also considered the knowledge level of the participants as well as the allocated budget

since a lot of travelling was critical to get to diverse audiences.

The study participants were brought together for an FGD through a scheduled online platform. The purpose of the FGD was to identify factors that are responsible for the high level of NPLs in Kenya. Results of the FGDs were also necessary for setting initial parameter for modeling and simulating credit risk using the system dynamics approach. Mini groups of between five (5) to six (6) people were used because the research required a deeper understanding of the factors influencing non-performance of loans, credit defaulting and current credit rating measures. The participants were grouped according to their specialty but when discussing capacity was required the discussion group comprised of up to a maximum of 12 members.

3.5 Research Instruments

Document analysis and focus-group discussions (FGDs) were the primary techniques of data gathering in this study. Focus group discussions were important in this research because of their strength in collecting qualitative data and the nature of the research through interaction with experts in the field to obtain their opinions. A Focus Group Discussion refers to a qualitative research method in social science which has been emphasized and applied in the developmental program evaluation spheres. This is in agreement with Mishra (2016) who stated that FGD is a group of co-operating persons having a common interest or who have been concerned by a moderator. The moderator engages the group and its interaction to acquire information relating to a specific or focused topic.

The use of FGD formed the part of the DSM methodology that requires both the qualitative and quantitative study approaches. The Dynamic Synthesis Methodology is underpinned by qualitative data collection methods which is significant in developing causal loop diagrams.

In this research, qualitative methods facilitated obtaining of perceptions, opinions and attitudes as well as generation of recommendations on how to deal with delays in feedback from the credit environment and the resulting unforeseen defaults. The ideas that were generated were useful for enriching the causal loop diagrams, stock-and-flow diagrams, and the simulation models.

In this research document analysis, which formed a critical examination of public or private recorded information relating to loan default rates were used as a data collection technique. Document analyses enabled the researchers to access data at their convenient time and also obtain nonobstructive information for the purpose of the research. Data was collected in order to measure the frequency of defaults, correlation between CRB-predictions and actual performance of customers, and factors influencing default probabilities by customers.

3.6 Validity and Reliability of the Instruments

For validity and reliability, ten semi-structured interview schedules were administered to staff performing core functions of credit risk analysis as part of pilot testing. The same interview schedules were re-administered to the same respondents after two weeks and the results obtained were verified for accuracy and correctness.

3.7 Data Processing and Analysis

In order to analyze the qualitative data collected, thematic approach was used. These themes included general knowledge and technical knowledge about causes of customer default considering each key variable as stated in the conceptual model, i.e., customer personal characteristics, customer income, creditor characteristics, issued loans, performing loans non-performing loans, application scoring, inquired funds and collateral amount. Both

quantitative and qualitative data collected from case and field studies were analyzed for descriptive statistics using SPSS and thematic analysis to obtain exploratory values. Thematic analysis was necessary in order to aggregate similar codes to form major concepts or themes for model building.

It was also important to separate group discussions for the separate nine key variables identified as well as one that incorporated all the nine variables. Since the groups were small, the researcher incorporated at least one participant from each category in each group. The first meeting was to create awareness of what SDM entails and let the group members understand what the research was all about. The second meeting was to discuss the factors influencing non-performance of loans; credit defaulting and current credit rating measures while taking notes and doing the recording at the same time. The third visit entailed a presentation using Vensim (software for developing causal loop diagrams and stock-and-flow diagrams) to confirm that everything discussed has been properly presented and accurately recorded.

3.9 Ethical Considerations

The researcher ensured that genuine and accurate results were recorded to the best of her capacity. This was done through constant communication with study participants and the supervisor and followed all the stated guidelines. Consent was sought from respondents before any information was collected from them and this was done both verbally and in writing. The interviewees and focus group discussion participants were well informed of their rights to pull out of the exercise at any moment that they choose. Confidentiality about the respondents' identities was protected and all data obtained was treated in aggregation.

CHAPTER FOUR
DATA ANALYSIS FINDINGS AND DISCUSSION

4.1 Introduction

This chapter presents results of the analysis of data collected during the study. It contains analyzed data arranged according to specific objectives through the use of computer aided data analysis software; SPSS and Vensim PLE. The main objective of the study was to develop a systems dynamics model for assessing credit risk of licensed credit reference bureaus in Kenya. Data was collected using FGDs (interview schedules) and document analysis.

4.2 Response rate

Loosely structured discussions of various topics of interest by the researcher were administered to respondents in 5 FGs during the second visit as outlined in the methodology chapter. All the questions were responded to and coded for analysis. Since FGDs were structured, directed and expressive, they yielded a lot of information in a relatively short time. A summary of number of questions given to each FG, which were 100% responded to, are summarized in Table 4.1 and represented in Figure 4.1 below;

Table 4.1: Focus Groups' Response Rate

<i>Focus Group</i>	<i>Items Discussed</i>	<i>Number of Questions as a %</i>	<i>Number of Participants</i>
1	Customer personal characteristics and customer income	25%	6
2	Creditor characteristics	12.5%	5
3	Issued loans, performing loans non-performing loans	25%	6
4	Application scoring	18.5%	6
5	Inquired funds and collateral amount	18.75%	6
		100%	29

Figure 4.1 above shows the return rates of interview schedules administered to respondents of five FGs in the chosen areas of interest: FG1 (Customer personal characteristics and customer income) 25%, FG2 (Creditor characteristics) 12.5%, FG3 (Issued loans, performing loans non-performing loans) 25%, FG4 (Application scoring) 18.5% and FG5 (Inquired funds and collateral amount) 18.75%.

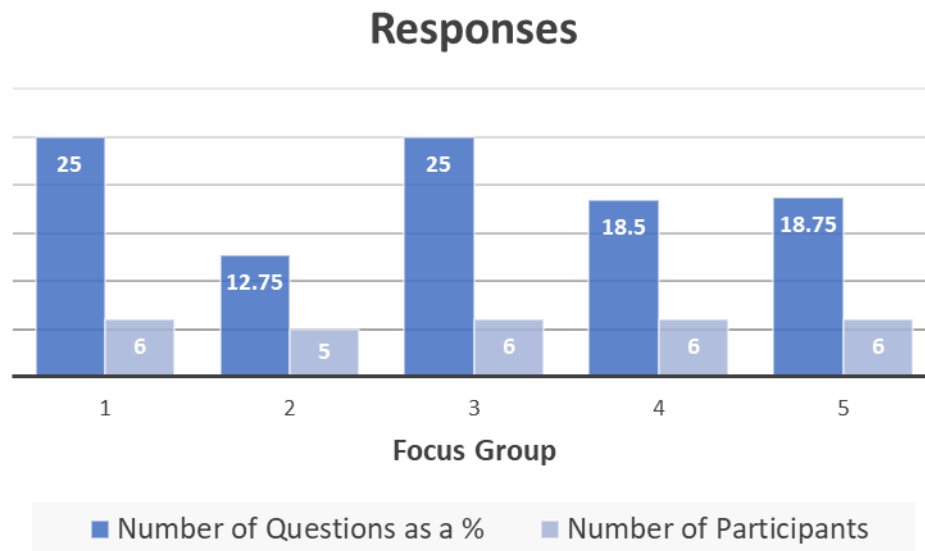


Figure 4.1: Focus Group Responses

4.3 Demographic Information

The demographic characteristics of a population included gender, profession, age, education, and experience. Out of the 29 participants, 18 of them (62%) were male while 11 (38%) were female. On the part of profession, the participants were made up of 6 asset managers, 6 financial risk managers, 1 business analyst, 2 internal auditors, 5 credit analysts and 9 loan officers. The percentage of this distribution is shown in Figure 4.2.

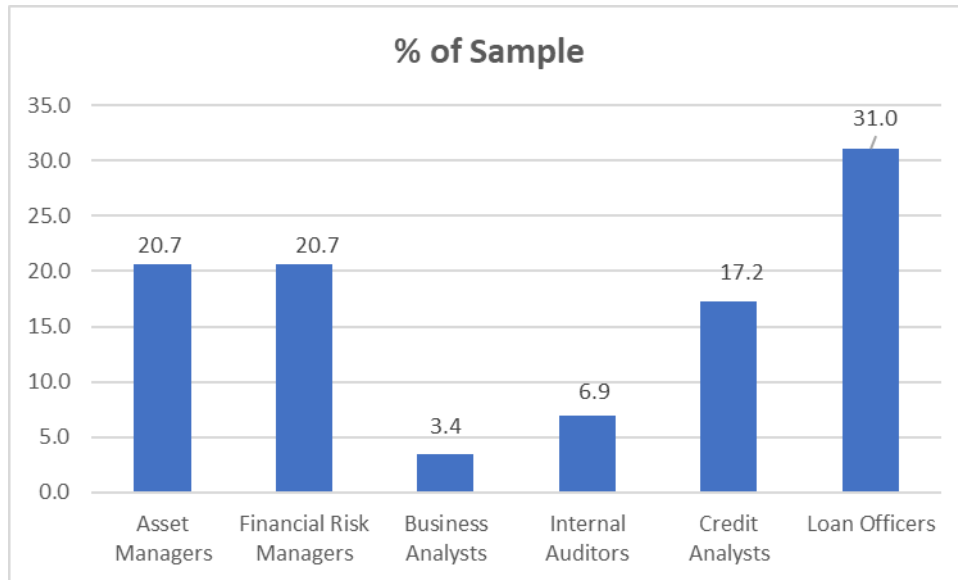


Figure 4.2: Distribution of Profession for the Focus Group Participants

The rest of the demographic information is summarized in Table 4.2 below. From the table results, participants had a lot of age and experience differences. The highest standard deviation (SD) was in age, with ± 21.26 years among participants. All participants had at least a bachelor's degree and years of technical training.

Table 4.2: Demographic Information for Participants

<i>Numerical Variable</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Age (years)	24	57	32.44	21.26
Academic education (years)	16	21	18.73	8.87
Technical education (years)	4	15	7.08	7.61
Work experience (years)	2	36	9.35	10.52

4.4 Descriptive Results from Study Variables

Five FGDs were held with the employees performing the core functions of credit risk management between July and August 2021. The objectives of the interviews and the FGDs were to explore and unveil the key factors that affect non-payment of loans once issued. During the first visit the researcher created awareness of what causal loop diagrams entail and

letting all the interviewees understand what the research was all about. The second visit discussed the causes and effects of credit risk using five GDs within each thematic area of interest. Each focus group was moderated by the researcher and lasted between 45 and 60 minutes.

4.4.1 Customer Personal Characteristics

Majority of the respondents had been involved in active processing of loan applications and surveying the loan applicants. One loan officer stated that “We have found out that older people of 45 years and above are more likely to repay their loans than younger people (18 – 25 years) who are carefree and with no family entanglements” All the respondents agreed that female applicants posed less credit risk than their male counterparts, and that loan applicants who are married have less capability for defaulting on their loans than the single people. However, one participant also noted that the expenses that come with raising children can also interfere with the ability to pay back a loaned amount. One credit analyst also mentioned that lifestyle of the loan applicant must be considered when issuing a loan, since most of the loan applicants, especially those living in the cities tend to live above their means.

4.4.2 Customer Income

The first focus group (FG1) also discussed the customer income variable influence on credit risk. General participants agreed that issues such as employment details (including contract expiry), gross and net monthly salary, mobile banking details and monthly financial commitments come into play when determining credit risk. One asset manager stated that “contract employees are more difficult to deal with because there is no guarantee of financial stability if their contract is not renewed”. A business analyst offered the information that loan applicants who strictly followed the 50-30-20 rule (50 percent expenditure into “needs,” 30

percent expenditure into “wants,” and 20 percent expenditure into financial goals) were more likely to repay their loans in full. Two participants contributed that some of the high-income earners are also heavy spenders and therefore risky borrowers. In general, it was agreed that monetary income is positively correlated with the ability to repay a loan.

4.4.3 Creditor Characteristics

Participants of the second Focus Group (FG2) were to discuss the effects of creditor characteristics on credit risk. Participants agreed that sometimes creditors can decide to forgo the investment of loan. For instance a mortgage applicant with a superior credit rating and a stable income is likely to be seen as posing a lower credit risk and also likely to get a low interest rate on mortgage. On the other hand a loan applicant with a poor credit history may be forced to get mortgage from a high-interest lender who can afford to offer loans with much higher interest to cater for the risk of defaulting. In the same manner bond-issuers and mobile money lenders who have less than perfect ratings usually offer much higher interest rates than those with almost-perfect credit ratings. The creditors who have low credit ratings use high returns to attract investors to assume risk associate with the monetary products that they are offering.

4.4.4 Issued Loans

Participants of the respective focus group agreed that loans make up the greatest and the most obvious source of credit risk for financial institutions. All FG3 interviewees thought that the amount of loan issues, as well as the number of active loans can largely determine a borrower’s ability to pay. A loan officer stated that “loan applicants must provide information regarding any previously issued loans from the current or other financial institutions”. They also agreed that sometimes it is difficult to track issued loans since some of these are not

from recognized financial institutions.

4.4.5 Performing Loans Non-performing Loans

The number and quantity of performing and non-performing loans by both the creditor and the borrower was brought up as a major risk determinant for banks and microfinance institutions. It was agreed that as the prevalence of non-performing loans increases, asset quality will inevitably deteriorate and the return on investment will decrease. One of the most effective ways to achieve a good level of performing loans, the members agreed, was through the formulation of a well-documented and properly managed credit risk management policy. Members thought that financial institutions which did not have a good credit risk management policy often failed to design effective strategies on which customers to grant credit, how and what to do at the branches and other corporate division levels while thoroughly assessing their potential borrowers. The resultant amount of non-performing loans affected the lending capacity of financial institutions heavily because of a diminished core capital. This meant that non-performing loans was mostly affecting the financial institutions negatively by denying these institutions access to capital market, both debt and equity. Increasing levels of non-performing loans had over the recent years led to undercapitalization of the financial institutions, and also resulted in negative effects on shareholders funds, sometimes leading to insolvency and collapse of some financial institutions.

4.4.6 Application Scoring

Participants generally thought that application scoring was a good way of credit risk management. Members agreed that the best scoring model must be highly discriminative, such that high scores should reflect almost no risk and low scores should indicate very high risk. “The most discriminating credit scoring system offers the best ranking of potential

customers from low-risk to high-risk”, a credit analyst stated. Nevertheless, participants also discussed the problem of lack of testing and validation of the emerging credit application scoring methods. One participant stated that “The adoption of application scoring techniques which have not been properly tested in new or innovative areas of the market has led to some serious problems in many financial institutions.” Participants suggested that this problem could be mitigated through constant calibration of the scoring application with localized data and analysis of past observed credit losses for every score. It was also necessary to segment credit scores into homogenous pools in order to take into account the relative diversity of potential borrowers.

4.4.7 Inquired Funds

Regarding this topic, the members of the concerned FG were of the collective agreement that all other factors held constant, the amount of funds inquired by an individual or business was positively associated with the borrower’s ability to repay and thus to the credit risk for the issuing bank or microfinance institution. One member contributed that particularly when operating within an imbalanced social lending environment, the inquired funds directly interplayed with factors such as interest rate, the loan terms, social network relationships and personal characteristics to determine the actual loan that a bank or other financial institution could safely issue to a borrower.

4.4.8 Collateral Amount

Participants of this Focus Group brought out the issue that a common and very crucial among troubled banks and microfinance institutions is the inability to monitor borrowers or collateral values. Many institutions usually fail to get periodic financial information from borrowers or real estate appraisals in order to evaluate the quality of loans on their books and the adequacy

of the collateral amount. As a consequence, many banks do not recognise the early signs that asset quality was deteriorating and missed opportunities to work with borrowers to identify and prevent their financial deterioration and to protect the bank's position. This inefficiency with monitoring has resulted in unexpected costs by senior management to determine the level and severeness of the problematic loans, and this has generated huge losses for many banks and microfinance institutions.

4.5 System Dynamics Model of Study Variables

The variables and sub-variables identified from the FDs as well as the baseline data that was provided by participants allowed the research to come up with the architecture of credit risk model presented in Figure 4.3. A positive link between from one variable to another indicates that the variable on the tail end of the link is positively influencing the variable on the head end. For example in Figure 4.3, an *increase* in the age of a borrower *increases* repayment ability, as was pointed out in the focus group discussions. Likewise, a negative polarity between two variables shows that the variable at the tail end is influencing the variable at the head end negatively. In the case of Figure 4.3, as repayment ability *increases*, it *reduces* credit risk. Several variables have positive influence, such as marital status, age, customer income, creditor ratings and collateral amount on repayment ability. A feedback loop forms when the feedback returns back to the originating variable, resulting in a closed loop. An example if the loop between the variables *performing loans* → *customer* → *income* → *expenditure* → *marital status* → *repayment ability* → *performing loans*. Such loops are important for determining the general behaviour of the system.

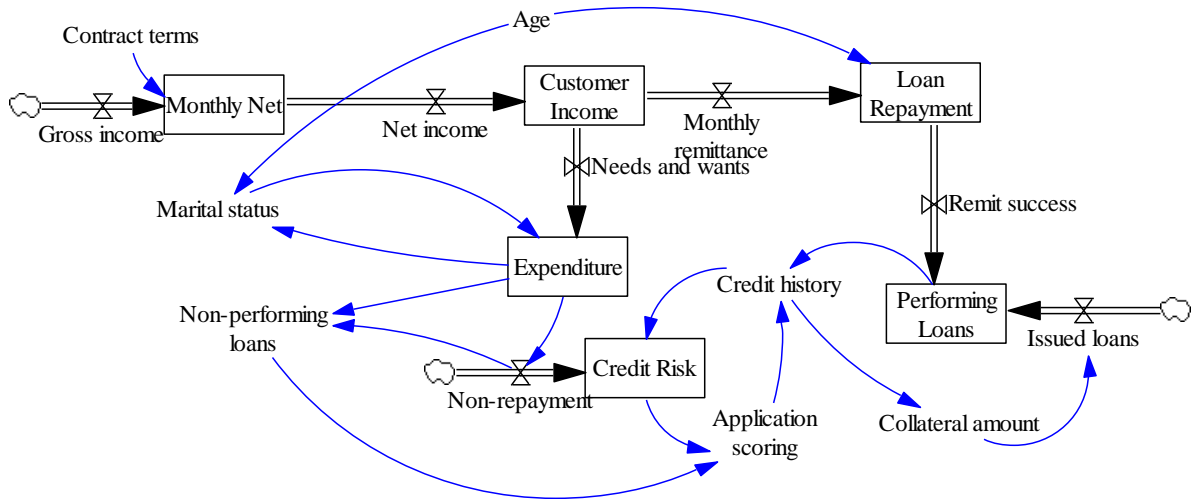


Figure 4.4: *Quantitative Stock-and-flow model of credit risk for banks and microfinance institutions in Kenya*

The variables **Monthly Net**, **Customer Income**, **Loan Repayment**, **Expenditure**, **Credit Risk** and **Performing Loans** are stock variables. This means that they contain a quantity which can increase or decrease depending on the amount flowing into or out of the variables. Other variables such as monthly remittance are flow variables. This means that monthly remittance controls the amount of units (Kenya Shillings) that flow from the variable Customer income to the variable Loan repayment. Other variables such as contract terms, credit history and collateral amount simply inform the model.

4.7 Sensitivity Analysis and Model Validation

Important variables that were identified for the model were Loan repayment, performing loans, credit risk. Whenever changes were made to other variables in the model, these 3 variables were heavily influenced. Therefore, the testing and validation of the model concentrated on sensitivity analysis while observing these variables over an observation period of 5 years (60 months). For the sensitivity analysis the values of all output variables (Loan repayment, Performing loans and Credit risk) were normalized to range from 0-10 to

increase consistency in plotting and visualization.

4.7.1 Adjustment of Borrower's Expenditure

Figure 4.5 shows the results of sensitivity analysis conducted on the model as part of the validation test by adjusting the borrower's expenditure from Kes 50,000.00 to 75,000.00 per month. The results show that with 50% increase in monthly expenditure, the performing loans dropped from an index of 8.9 out of 10 to 5 out of 10 by month 33 of observation, and further to 2.5 out of 10 by month 60 of observation. At the same time the loan repayment which had been stable over several months dropped drastically between month 30 and 60 by which time it had moved from 8.9 out of 10 to 1.7 out of 10. Consistent with the two observations, credit risk increased systematically with the increase in borrower's expenditure, following an S-shaped pattern of behaviour. Starting from month it rose sharply and continued to rise until the end of the observation period by which it was at 8.45 out of 10.

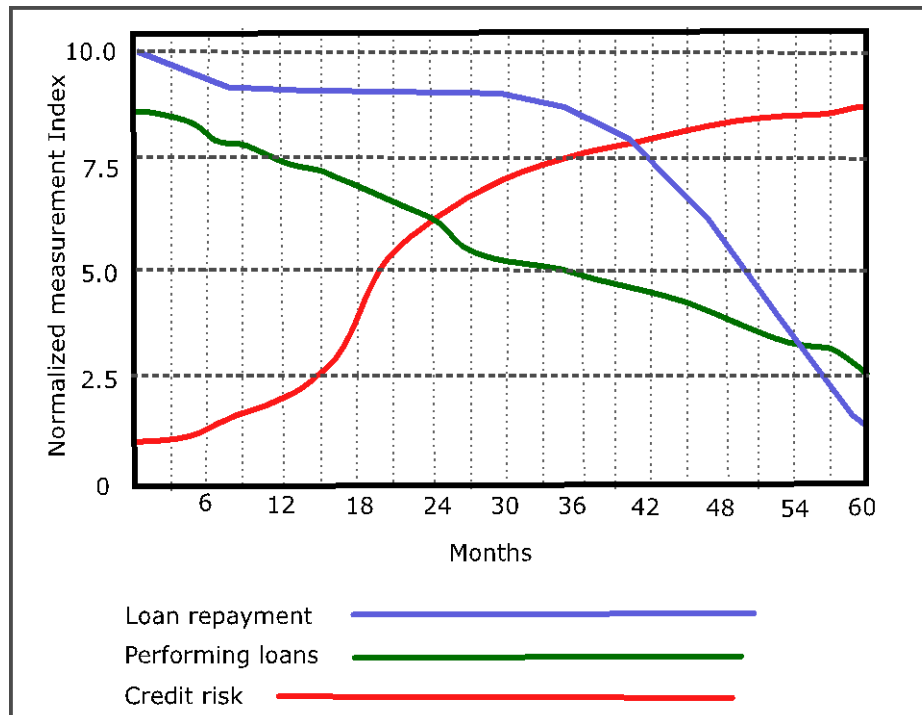


Figure 4.5: Sensitivity analysis - adjusting Borrower's Expenditure from Kes 50,000 to 200,000 pm

4.7.2 Adjustment of Borrower's Monthly Remittance

In the second scenario, other model factors were held constant and the borrower's monthly remittance amount was reduced while the repayment period was increased. Results shown in Figure 4.6 show that when the borrower's monthly remittance was reduced to 65% of the initial remittance amount, the loan started performing well around month 40 and performance went up consistently until the end of the observation period. Similarly, loan repayment continued to fluctuate but became stable around month 32 and continued an upward trend with increasing repayment ability until around 10/10 perfect score at month 60. Credit risk portrayed an oscillating pattern but it did not go beyond 4.5/10 throughout the observation period. From month 48 onwards there was a continuous reduction in credit risk until the end of the simulation period.

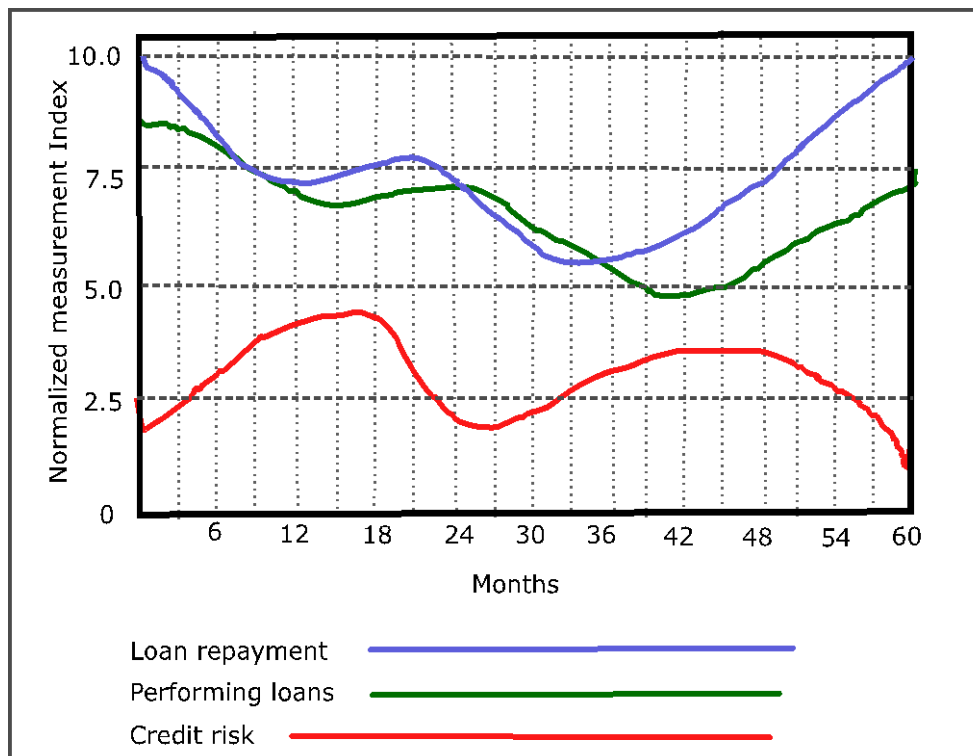


Figure 4.6: Sensitivity analysis – Adjustment of monthly remittance by -35%

4.7.3 Adjustment of Borrower's Income

The final sensitivity analysis was to increase the borrower's net income by 25%. A result of

this validation is shown in Figure 4.7. Loan repayment was improved significantly after this adjustment such that by month 40 it had stabilizing and went on an upward trend with consistent repayments until the end. Performing loans were also stable and did not go below the index of 7/10. By month 39 until the end there was no downward trend observed. While credit risk went down to some extent around the first observation months, the index did not go beyond 3.5/10 as was observed in the first sensitivity analysis. Around month 39 onward credit risk went down to below 1/10.

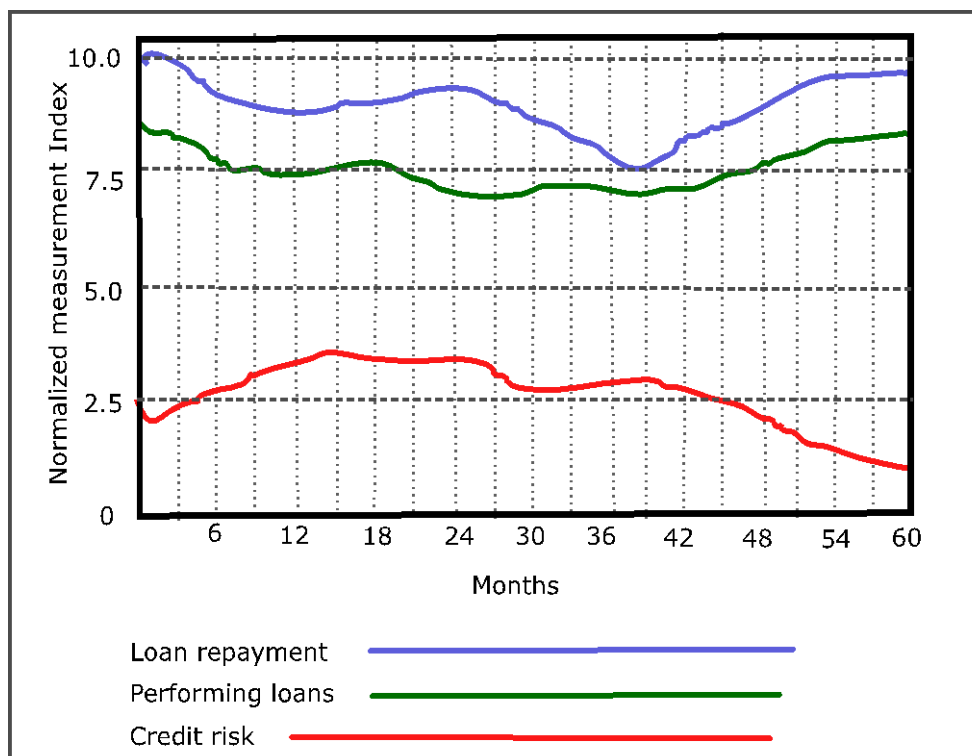


Figure 4.7: Sensitivity analysis – Adjustment of borrower’s net income by 25%

4.8 Discussion of Results

The objectives of this study were to determine the factors that influence credit risk, develop a system dynamics model for determining credit, and test and validate the credit risk model. Three main factors were prominent from the observations as well as the sensitivity analysis that was conducted. These were the loan repayment ability, the number of performing loans,

and credit risk. These three variables were mostly influenced by factors such as the borrower's income and expenditure and the loan remittance amount. The findings of this study are consistent with the findings of past studies such as Duong and Huong (2017), Belás et al. (2018), Frei (2020), Ngumo et al. (2020) and Misati et al. (2021) as discussed in this section.

Duong and Huong (2017) concluded after conducting a credit risk analysis of Vietnamese banks that borrowers' income significantly affected their ability to pay loans, which in turn contributed to credit risk as has also been seen in this study. Belás et al. (2018) and Frei (2020) found out that social and economic aspects of borrowers were major determinants of credit risk. While both studies concluded that there can be no certainty of optimal outcomes, an increase in the borrower's income and a decrease in other primary and secondary expenditures will both improve credit risk.

In a series of local studies done on credit risk in Kenya, Ngumo et al. (2020) and Misati et al. (2021) concluded that considering the high number of loan defaults per year, measures must be taken by financial institutions to rigorously assess their potential borrowers, but in the case of unexpected outcomes then the financial institutions should consider mitigations such as renegotiating the repayment amount. This is consistent with what the current study observed, because when the repayment amount was reduced, the credit risk reduced significantly. This means that there is a positive correlation between reduction in loan remittance amount and the ability for borrowers to repay their loans.

Although this study did not conduct sensitivity analysis on borrower's personal factors such as age and marital status, Belás et al. (2018) and Ntwiga (2020) found that factors such as age, gender, marital status and household size influenced the ability to repay loans, especially among low income earners. Ntwiga (2020) included in the analysis other Micro-credit loan products such as M-shwari. The study concluded that women were more stable in loan

repayment compared to men. They however advocated for diversification of credit facilities to cater for diverse peer groups particularly of low-income earners to increase financial inclusion and offer affordable options. Indeed while factors such as age and marital status were not highly significant in the current study, there was however a level of influence on credit risk that these factors contributed.

4.9 Conclusion

This study sought to investigate causes of failure in repayment of amounts issued to borrowers by banks and microfinance institutions and also to build a system dynamics model that would provide an in-depth understanding of the complex relationships that exists amongst the key variables in a credit risk management system. Data was collected through the use of FGD and the information obtained was used to improve the initial casual loop diagram which eventually was used to develop the model using Vensim PLE software. The model was verified, validated and results from simulation replicated the mental reference mode. The system dynamics model was validated, as part of diagnostic test, in order to determine the degree to which it was an accurate representation of the real-world from the perspective of the intended uses of the model. A test of suitability was performed through policy sensitivity and robustness test in order to validate policy implications. The sensitivity of a policy with respect to change in parameter values was used to judge whether the model-based policy recommendations change with reasonable changes in parameter values or reasonable alteration in equation formulations.

The findings in this study indicate that when the income of borrowers decreases or their expenditure increase, as has mostly been witnessed during the recent pandemic period this can significantly decrease their ability to repay loans, and in turn increase credit risk for banks and microfinance institutions. For effective economic assessment, a careful analysis is

needed to determine the extent to which this income-expenditure balance is affecting credit risk across all financial institutions countrywide. Optimal determination of credit risk also needs to consider other factors in the ecosystem such as socioeconomic, political and environmental influences.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter summarizes the key findings of the current study. A conclusion is made from the findings and recommendations for areas of potential improvements are identified. The chapter also makes suggestions for future studies to expand the scope of knowledge on the factors influencing credit risk for commercial banks and financial institutions.

5.2 Key Findings

The first objective of the study was *to determine the factors that influence credit risk by individuals and businesses in Kenya*. It became clear from the findings of the study that credit risk is mainly caused by customer-income-related factors (such as contract terms, gross and net monthly salary), creditor characteristics, the number and size of loans already issued, the number and amount of performing and non-performing loans, the scoring level of applicants, the amount of funds inquired, and the amount of collateral used to secure a loan. It was discovered that mobile banking details and monthly financial commitments come into play when determining credit risk. The lack of a permanent employment generally makes it difficult to assure of loan repayment capability given the lack of guarantee of financial stability. At the same time, restricted and planned expenditure increased the capacity to repay a loan and significantly reduced credit risk for individuals and businesses. It was further found out that loan applicants with superior credit ratings are more likely to pose low credit risk and can thus get loans with lower interest rates. Having a good scoring model, one that is highly discriminative can allow perfect ranking of customers according to levels of risk and

thus reduce credit risk. Finally, requiring sufficient collateral, and being able to track the value of secured assets can significantly reduce credit risk. On the other hand, the inability to monitor borrowers or collateral values and obtain periodic financial information from borrowers or real estate appraisals for evaluating the quality of loans on their books and the adequacy of the collateral amount can generate huge losses for banks and microfinance institutions.

The second objective of this study was *to design and develop a system dynamics model for credit risk determination*. The model was successfully built using the Vensim PLE software which conforms to the requirements of the dynamic hypothesis and graph behavior over time (mental models). First a causal loop diagram was generated from variables of the conceptual model, on the basis of analyzed data collected from focus group discussions (FGDs). This model represented the architecture of the credit risk ecosystem. The causal loop diagram was then converted into a stock and flow model which is highly quantitative and allows the running of simulations to test the model. The quantitative stock-and-flow model was developed integrating a graphical output that allows users to visualize the accumulation of values and also to change the initial model parameter values or revert to the initial values thus allowing animation of output during the simulation runs. In general, the model generated simulations that were consistent with the observations of variable behavior from the focus group discussions.

The final objective of this study was *to test and validate the credit risk determination model*. After developing the final model, an empirical investigation was conducted to collect data using both document analysis and FGDs and analyzed using statistical software and thematic analysis. Descriptive statistics arising from the analysis of the bank default statistics was used as the basic data to populate the model. This was followed by scenario building and testing of various credit risk outcomes as well as validation of the model. One of the validation methods

applied was the comparison between computational results from the simulation runs and the actual data from the system. These results were used to confirm that the developed model is consistent with the reality of credit risk management in terms of its structure and behaviour (real system).

5.3 Study Limitations

Although the current study had many important findings, it was also faced by a number of limitations. The first limitation is that the current research did not consider the influence of the country's variation in the gross domestic product (GDP) despite the fact that a country's economy largely affects how individuals and businesses are able to sustain their income or repay their loans. Additionally, the study did not take account of the direct factors resulting from measures imposed by the government to curb the COVID-19 pandemic, and how these measured affected business operations.

Regarding the stability of the system dynamics model, the limitation encountered was that the variables applied were constructed using data from limited sources, i.e., FGDs and credit default statistics. The research did not use raw data from financial institutions or conduct a countrywide survey of borrowers. The model was also restricted to a few variables, which limited its predictive capability.

5.4 Recommendations for Future Research

Given the identified limitations, the study recommends several areas of future research. To address the first limitation of limited factor evaluation, it is necessary for future study to consider all relevant variables that possibly affect the credit risk management of individuals and businesses. Further research should be undertaken in future to investigate the dynamics

and effects of a wide range of variables on credit risk assessment. All potential variables should be included in both the conceptual model and the quantitative system dynamics model to enable simulation runs using the final model to include diverse conditions for both individuals and businesses, given that risk analysis is a critical procedure in the processing of loans.

Finally, the study recommends as policy actions, that the existing models of credit risk management used in the financial sector should be scaled up to incorporate research outcomes. The upgrading of credit evaluation systems is necessary to address the numerous financial-related issues that crop up when borrowers fail to repay loans and lead to heavy losses for the issuing lending institutions, which in turn affects the country's economy. The incorporations of recent research outcomes into the revamping of current credit risk management models will go a long way to significantly reduce the possibility of credit risk and insolvency of commercial banks and microfinance institutions in Kenya.

5.5 Contributions of the Study

In this study, the descriptive and explorative research has shown that there are a number of factors which influence credit risk for commercial banks and microfinance institutions. The use of SD methodology facilitated an in-depth understanding of the aspects which fall under the ecosystem of credit risk management in Kenya. Usage of the system dynamics methodology provided new insight to the existing studies which have mainly used other predictive methods which perform well under stable conditions but fail to perform within a rapidly changing ecosystem. Thus, the study facilitated an in-depth understanding of the problems surrounding credit risk management and the mitigation factors. It also provided useful insights and foresight into policy matters concerning financial lending processes for

banks and microfinance institutions that can drastically improve credit rating practices. This study has increased the available literature on application of systems dynamics on financial systems while focusing on several components of banking applications that are capable of predicting the effects of the changes in policies or operations of the entire credit management ecosystem.

This study provides a usable reference that aids researchers, learners and academicians in readily selecting an appropriate SD approach for designing and developing their system models and prototypes, cognizant of both their expertise and those requirements of their application domain. Creating, discussing and sharing causal loop diagrams with the credit risk management personnel enabled visualization of the emergent properties of the financial system that need urgent attention.

REFERENCES

- Abdipoor, S., Nasser, A., Akbarpour, M., Parsian, D., & Zamani, S. (2013). Integrating neural network and colonial competitive algorithm: a new approach for predicting bankruptcy in Tehran security exchange. *Asian Economic and Financial Review*, 3(11), 1528-1539.
- Abotaleb, I. S., & El-adaway, I. H. (2018). Managing construction projects through dynamic modeling: Reviewing the existing body of knowledge and deriving future research directions. *Journal of Management in Engineering*, 34(6), 04018033.
- Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Risks*, 6(2), 38.
- Belkin, J., Blumstein, A., & Glass, W. (1973). Recidivism as a feedback process: An analytical model and empirical validation. *Journal of criminal Justice*, 1(1), 7-26.
- Berry, M. A., & Linoff, G. S. (2000). *Mastering data mining: The art and science of customer relationship management*. Industrial Management & Data Systems.
- Blumstein, A., & Larson, R. (1969). Models of a total criminal justice system. *Operations Research*, 17(2), 199-232.
- Akkoc S (2012) An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzy inference system (ANFIS) model for credit scoring analysis: the case of Turkish credit card data. *Eur J Operat Res* 222(1):168–178.
- Baradaran, V., & Keshavarz, M. (2017). System dynamics modelling of retailers' credit risk. *International Journal of Industrial and Systems Engineering*, 26(3), 380-396.
- Baldwin, C. Y., & Mason, S. P. (1983). The resolution of claims in financial distress the case of Massey Ferguson. *The Journal of Finance*, 38(2), 505-516.
- Basel Committee on Bank Supervision, BCBS (2001). *Overview of the new baselcapital accord*. Technical report, Basel Committee on Bank Supervision, CH-4002Basel, Switzerland. Consultative Document.
- BCBS (2001), *Principles for the Management of Credit Risk*. Consultative paper, The Basel Committee on Banking Supervision

- Blanco-Fernández, A., Colubi, A., & García-Bárzana, M. (2013). A set arithmetic-based linear regression model for modelling interval-valued responses through real-valued variables. *Information Sciences*, 247, 109-122.
- Brown, M., Jappelli, T., & Pagano, M. (2009). Information sharing and credit: Firm-level evidence from transition countries. *Journal of Financial Intermediation*, 18(2), 151-172.
- Bekhet HA, and Eletter SFK, (2014) Credit risk assessment model for Jordanian commercial banks: Neural scoring approach. *Review of Development Finance*. 4(1):20–28.
- Belás, J., Smrcka, L., Gavurova, B., & Dvorsky, J. (2018). The impact of social and economic factors in the credit risk management of SME. *Technological and Economic Development of Economy*, 24(3), 1215-1230.
- Berry, M.J.A. and Linoff, G.S. (2000), *Mastering Data Mining*, Wiley, New York, NY.
- Blanco, A., Pino-Mejías, R., Lara, J., & Rayo, S. (2013). Credit scoring models for the microfinance industry using neural networks: Evidence from Peru. *Expert Systems with applications*, 40(1), 356-364.
- Boritz, J.E. (1991) *The assumption of Business Continuity and its Repercussions on Accounting and Auditing*, Research Report, CICA Publication.
- CBK (2015). *Bank Supervision Annual Report*. Available at https://www.centralbank.go.ke/uploads/399346751_2015%20Annual%20Report.pdf
- CBK (2018). *Bank Supervision Annual Report*. Available at https://www.centralbank.go.ke/uploads/banking_sector_annual_reports/1174296311_2018%20Annual%20Report.pdf.
- CBK (2021) *Credit Officer Survey March 31 2021*. Available at https://www.centralbank.go.ke/uploads/banking_sector_reports/1506362303_Credit%20Survey%20Report%20for%20the%20Quarter%20ended%20March%202021.pdf
- Chen, W., Xiang, G., Liu, Y., & Wang, K. (2012). Credit risk Evaluation by hybrid data mining technique. *Systems Engineering Procedia*, 3, 194-200.
- Chelogoi, M. (2013). *"Micro credit and the Poorest of the Poor: Theory and Evidence*. Nairobi, Kenya
- Coyle, B. (2000). *Framework for Credit Risk Management*, London: Chartered Institute of Bankers,

- Crosbie, P., & Bohn, J. (2003). *Modelling Default Risk*. Retrieved 16 August 2009.
- Csikosova, A., Janoskova, M., & Culkova, K. (2019). Limitation of financial health prediction in companies from post-communist countries. *Journal of Risk and Financial Management*, 12(1), 15.
- Dong, Y., Hou, C., & Zhao, D. (2017). Research on System Dynamic Modeling and Simulation of Chinese Supply Chain Financial Credit Risk from the Perspective of Cooperation.
- Duong, N. T., & Huong, T. T. T. (2017). The analysis of major credit risk factors-The case of the Vietnamese commercial banks. *International Journal of Financial Research*, 8(1), 33-42.
- Emel A., Oral M., Reisman A. & Yolalan R. (2003). A credit scoring approach for the commercial banking sector. *Socio Econ Plan Sci* 37(2):103–123.
- Evans, G. W., Biles, W. E., & Bae, K. H. G. (Eds.). (2019). *Analytics, operations, and strategic decision making in the public sector*. IGI Global.
- Van Eeuwijk, P., & Angehrn, Z. (2017). *How to conduct a Focus Group Discussion (FGD). Methodological Manual*.
- Frei, C. (2020). A New Approach to Risk Attribution and Its Application in Credit Risk Analysis. *Risks*, 8(2), 65.
- Ghodselahe, A. (2011). A hybrid support vector machine ensemble model for credit scoring. *International Journal of Computer Applications*, 17(5), 1-5.
- Hand, D. J., & Yu, K. (2001). Idiot's Bayes—not so stupid after all?. *International statistical review*, 69(3), 385-398.
- Jiang, L., & Su, Y. (2016, November). Research on the Evaluation of Supply Chain Finance Credit Risk of Small and Medium-Sized Enterprise Based on System Dynamics. In *First International Conference Economic and Business Management 2016*. Atlantis Press.
- Juneja, S. (2017). Dynamic Portfolio Credit Risk and Large Deviations. In *Econophysics and Sociophysics: Recent Progress and Future Directions* (pp. 41-58). Springer, Cham.
- Kaminskyi, A. (2018). *System Dynamics Modelling of Credit Risk Management*.

- Kansiime, M. K., Tambo, J. A., Mugambi, I., Bundi, M., Kara, A., & Owuor, C. (2021). COVID-19 implications on household income and food security in Kenya and Uganda: Findings from a rapid assessment. *World development*, 137, 105199.
- Kargi, H. S. (2011). *Credit Risk and the Performance of Nigerian Banks*. Ahmadu Bello University, Zaria.
- Kithinji, A. M. (2010). *Credit risk management and profitability of commercial banks in Kenya*. School of Business, University of Nairobi, Nairobi.
- Kipngetch, R. M. (2019). *Effect of Liquidity Management on the Financial Performance of Insurance Companies in Kenya* (Doctoral dissertation, University of Nairobi).
- Koutanaei, F. N., Sajedi, H., & Khanbabaei, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. *Journal of Retailing and Consumer Services*, 27, 11-23.
- Kruppa J., Schwarz A.G. & Ziegler A. (2013). Customer credit risk: individual probability estimates using machine learning. *Expert Syst Appl* 40(13):5125–5131.
- Langarudi, S. P. (2016). *A System Dynamics Approach to the Political Economy of Resource-dependent Nations* (Doctoral dissertation, Worcester Polytechnic Institute).
- Lannon, C. (2018). *Causal Loop Construction: The Basics*, Retrieved on 24th September 2019 from <https://thesystemsthinker.com/causal-loop-construction-the-basics/>
- Littlejohns, L. B., Baum, F., Lawless, A., & Freeman, T. (2018). The value of a causal loop diagram in exploring the complex interplay of factors that influence health promotion in a multisectoral health system in Australia. *Health research policy and systems*, 16(1), 126.
- Longstaff, F. A., & Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3), 789-819.
- Maji, S. G., & Hazarika, P. (2016). Factors affecting credit risk of Indian banks: Application of dynamic panel data model. *Research Bulletin*, 42(1), 85-96.
- Mandala I.G.N.N., Nawangpalupi C. A., and Praktikto F. R (2012) *Assessing Credit Risk: An Application of Data Mining in a Rural Bank*. *Procedia Economics and Finance*. 4:406–412.
- Markowitz H.M (1952). *Portfolio Selection*. *Journal of Finance*, Vol.7 No.1.

- Markowitz H.M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Cambridge, MA: Basil Blackwell. 2nd ed.
- Melton, 1974. *Credit risk theory*
- Mercylynne, M. W., & Omagwa, J. (2017). Credit Risk Management and Financial Performance Of Selected Commercial Banks In Kenya. *Journal of Business and Management (IOSR-JBM)*, 19(11), 92-98.
- Misati, R. N., Kamau, A., Tiriongo, S., & Were, M. (2021). Credit risk and private sector loan growth under interest rate controls in Kenya. *African Review of Economics and Finance*, 13(1), 83-112.
- Mishra, L. (2016). Focus group discussion in qualitative research. *Techno Learn*, 6(1), 1.
- Moradi, S., & Rafiei, F. M. (2019). A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks. *Financial Innovation*, 5(1), 1-27.
- Mugenda, O. M. and Mugenda, A. G. (2003). *Research Methods: Quantitative and Qualitative Approaches*, Acts Press: Nairobi.
- Munene, R. N., Nyamache, T., & Nzioki, P. M. (2018). An Assessment of the Role of Credit Reference Bureau in Influencing Risk Identification in Mitigating Credit Default in Commercial Banks in Kenya.
- Muriithi, J. G., Waweru, K. M., & Muturi, W. M. (2016). Effect of credit risk on financial performance of commercial banks Kenya.
- Nasieku, T., & Ngugi, R. W. (2016). Credit Information Sharing and Credit Risk Reduction in Kenya Commercial Banks. *Journal of social science research*, 10(1), 1941-1949.
- Nasiurma, D. K. (2000). *Survey Sampling: Theory and methods*. Nairobi, Kenya.: University of Nairobi.
- Ngumo, K. O. S., Collins, K. W., & David, S. H. (2020). Determinants of financial performance of microfinance banks in Kenya. *arXiv preprint arXiv:2010.12569*.
- Ntwiga, D. B. (2020). Credit usage among the un (der) banked: consumer socio-economic characteristics and influence of financial technology. *International Journal of Financial Services Management*, 10(1), 38-54.

- Onaolapo, A. R. (2012). Analysis of credit risk management efficiency in Nigeria commercial banking sector,(2004-2009). *Far East Journal of Marketing and Management*, 2(4), 39-52.
- Otieno, S., Nyagol, M., & Onditi, A. (2016). Relationship between credit risk management and financial performance. Empirical evidence from microfinance banks in Kenya. *Research Journal of Finance and Accounting*, 7(6), 2222-2847.
- Petrică, A. C., Stancu, S., & Tindeche, A. (2016). Limitation of ARIMA models in financial and monetary economics. *Theoretical & Applied Economics*, 23(4).
- Peura, H., Yang, S. A., Lai, G. (2016). Trade Credit in Competition: A Horizontal Benefit. *Social Science Electronic Publishing*, 19(2), pp. 263-289.
- Rahim F. H. A., Hawari N. N., and Abidin N. Z., (2017) “Supply and demand of rice in Malaysia: A system dynamics approach”, *International Journal of Supply Chain and Management*, Vol.6, No.4, pp. 234-240, 2017
- Rwashana, A. S., Williams, D. W., & Neema, S. (2009). System dynamics approach to immunization healthcare issues in developing countries: a case study of Uganda. *Health Informatics Journal*, 15(2), 95-107.
- Reuter, J. (2013). Values and Worldviews. Diagnosing and Engaging with Complex Environmental Problems, 62-64.
- Rudd, M. P. H., GStat, J. M., & Priestley, J. L. (2017). A comparison of decision tree with logistic regression model for prediction of worst non-financial payment status in commercial credit.
- Song, H., Yu, K., Ganguly, A., Turson, R., 2016. Supply chain network, information sharing and SME credit quality. *Industrial Management & Data Systems*, 116(4), pp.740-758.
- Sproul, M. (1998). The quantity theory versus the real bills doctrine in colonial America. *Economics working papers*.
- Sterman, J. (2019), Fine-Tuning Your Causal Loop Diagrams-Part I Retrieved on 25th September 2019 from <https://thesystemsthinker.com/fine-tuning-your-causal-loop-diagrams-part-i/>
- Su, Y., Zhong, B., 2017. The Credit Risk Assessment Model of Internet Supply Chain Finance: Multi-Criteria Decision-Making Model with the Principle of Variable Weight. *Journal of Computer & Communications*, 04(16), pp.1-11.

- Teply, P., & Klinger, T. (2019). Agent-based modeling of systemic risk in the European banking sector. *Journal of Economic Interaction and Coordination*, 14(4), 811-833.
- Wagacha, A. & Othieno, F. (2016). Semi-Markovian credit risk modeling for consumer loans: Evidence from Kenya. *Journal of Economics and International Finance*, 8(7), 93.
- Wanjala, K. (2020). The Economic Impact Assessment of the Novel Coronavirus on Tourism and Trade in Kenya: Lessons from Preceding Epidemics. *Finance & Economics Review*, 2(1), 1-10.
- Whitaker, R. B. (1999). The early stages of financial distress. *Journal of economics and finance*, 23(2), 123-132.
- Wruck, K. H. (1990). Financial distress, reorganization, and organizational efficiency. *Journal of financial economics*, 27(2), 419-444.
- Yu, L., Wang, S., Lai, K. K., & Zhou, L. (2008). *BioInspired Credit Risk Analysis*. Berlin: Springer.
- Zanin, M., Papo, D., Sousa, P. A., Menasalvas, E., Nicchi, A., Kubik, E., & Boccaletti, S. (2016). Combining complex networks and data mining: why and how. *Physics Reports*, 635, 1-44.
- Zhang, Z., Gao, G., & Shi, Y. (2014). Credit risk evaluation using multi-criteria optimization classifier with kernel, fuzzification and penalty factors. *European Journal of Operational Research*, 237(1), 335-348.
- Zhu, Y., Zhou, L., Xie, C., Wang, G. J., & Nguyen, T. V. (2019). Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. *International Journal of Production Economics*, 211, 22-33.

Appendix 1: Project Schedule

Task	Start Date	End Date	Duration (Days)
Research concept, supervisor assignment and first meeting with supervisor	13/02/2020	22/02/2020	9
Prepare research proposal	25/10/2020	13/04/2021	18
Get final sign-off from supervisor and prepare and present proposal document to school of graduate studies and research	16/04/2021	19/04/2021	3
Research proposal presentation/approval	24/04/2021	24/04/2021	1
Search, capture and synthesise relevant literature	16/05/2021	30/06/2021	14
Prepare draft literature review	23/05/2021	06/07/2021	14
Finalise sampling plan	06/07/2021	08/08/2021	2
Develop data collection instrument	08/08/2021	10/08/2021	2
Pre-test/pilot data collection instrument	10/08/2021	11/08/2021	1
Carry out data collection	11/08/2021	22/08/2021	11
Prepare data for analysis	22/08/2021	23/08/2021	1
Analyze data	22/08/2021	27/08/2021	5
Draw conclusions/ recommendations	28/08/2021	07/09/2021	10
Prepare final draft of report	07/09/2021	14/09/2021	7
Review draft with supervisor	10/09/2021	17/09/2021	7
Get final sign-off from supervisor and prepare and present report document to school of graduate studies and research	18/09/2021	20/09/2021	2
Report Defence	27/09/2021	27/09/2021	1
Final editing	28/09/2021	05/10/2021	7
Printing, binding and final submission	05/10/2021	10/10/2021	5

Appendix 2: Budget Plan

Core Activities	Items/Participants	Total Cost
Equipment and Consolidating of literature	Secondary search, travel expenses	2,500
Design and development of research instruments	Typing and photocopying the research instruments	2,500
Research Instruction and training	Transport for researcher and research assistant	15,000
Finalizing the research instruments, stationery, typing and photocopying	Transport for researcher and research assistant	15,000
Main field/ data collection	Travel allowance for researcher and a research assistant	6,000
Data processing, encoding and cleaning	Researcher and a research assistant 10 x1000	8,000
Data Analyses and model building	Processing statistical data	20,000
Report writing and Typesetting	Prepared report drafts and final copies	10,000
Others	Books, Reams of paper and Telephone	5,000
10 % contingency and institutional cost		5,000
Miscellaneous		12,000.00
Grand Total		101,000.00

Appendix 3: Structured Interview Guide

Section I: Respondent Profile

1. Your department?

2. Gender of the respondent?
Male [] Female []

3. For how long have you been working with this bank?
Less than 2 years [] 2 to 5 years [] 6 to 10 years [] 11 to 15 years []
16 years and above []

4. Age in years?
30 years and below [] 31 to 40 years [] 41 to 50 years []
51 to 60 years [] 61 years and above []

5. Highest education level?
Certificate [] Diploma [] First degree []
Masters and above [] Others (specify)..... []

6. To which CRBs does your bank forward its customer's credit information?
.....
.....

7. What do you understand by Credit Referencing?
.....
.....

8. What do you understand by the term Machine Learning?
.....
.....

9. Does machine learning play a role on credit risk prediction? Yes [] No []
Why?.....
.....
.....

Section II: The Role of Machine Learning Model on Credit Risk Identification

10. How exactly does machine learning play a role on credit risk identification?

.....
.....
.....

11. How does machine learning identify credit risk factors?

.....
.....
.....

12. How would you say machine learning has helped in reducing credit risk rates?

.....
.....
.....

13. What are the setbacks of machine learning play in credit risk prediction?

.....
.....
.....

14. Has the bank experienced a decline in non-performing loans since the establishment of machine learning in CRB?

Yes [] No [] To what extent?

.....
.....
.....

Section III: The Influence of Machine Learning Model on the Rate of Credit Repayment

15. Are your clients aware that they are being monitored by machine learning systems?

Yes [] No [] How do they respond?

.....
.....
.....

16. How precisely does machine learning model influence rate of credit repayment?

.....

.....
.....
17. How accurate is the use of machine learning model to establish customers' repayment behaviour?

.....
.....
18. What can be done to improve the effectiveness of Machine learning model on the rate of credit repayment?

.....
.....
19. Does the use of machine learning on CRB enable banks to lend at a reasonable interest rate?

Section IV: The Mechanisms of Machine Learning Appropriate for Predictive Model

20. In addition to the CRB report, which other methods does your financial institution use to determine customers' abilities to pay?

.....
.....
21. What are some of the mechanisms involved in machine learning appropriate for predictive model in the Kenyan perspective?

.....
.....
22. Can anyone look at the credit information evaluated by machine learning and clearly understand?

Yes [] No [] To what level?

.....
.....
.....

23. What type of credit evaluation information data is provided by commercial banks to CRB machine learning?

.....
.....
.....

24. What ensures that the data provide to CRB is completely flawed proof?

.....
.....
.....

Appendix 4: Key Informants Interview

The discussion and filling in instructions

The Key Informant Interview is designed to collect information on the role of machine learning model on credit risk prediction among licensed credit reference bureaus in Kenya. The earliest use of machine learning was credit risk modeling, whose objective was to use financial data to forecast default risk.

This research will be of value to many stakeholders in the financial industry, including the bank managers, credit reference bureaus management, investors, bank customers and the general public.

It is my request that you fill the following questionnaire with specific reference to the above enquiry, by giving a brief statement. Please indicate any further opinions in the spaces provided.

Credit Reference Bureau

1. Name of organizations?

.....
.....
.....

2. How long have you been in the business of credit referencing?

.....
.....
.....

3. Description of the Credit Reference Bureau

- a) A group of banks
- b) A group of other financial intermediaries
- c) Individual shareholders
- d) Foreign owned
- e) Others (specify).....

4. Who uses a credit reference bureau?
- a) Financial institutions
 - b) Insurance
 - c) Utility companies
 - d) Others (specify).....

5. What is the role of machine learning model on credit risk identification among licensed credit reference bureaus in Kenya?

.....

.....

.....

6. What is the influence of machine learning model on the rate of credit repayment among licensed credit reference bureaus in Kenya?

.....

.....

.....

7. What are the mechanisms of machine learning appropriate for predictive model?

.....

.....

.....

8. Samples of CRB machine learning generated reports?

.....

.....

.....

9. Other comments.

.....

.....

.....

Thank You for Your Positive Response.

Appendix 5: Map of Study Area



Source: Google (2019)

Appendix 6: Authorization letter for data collection:



Thika Road, Ruaraka
P.O. Box 56808-00200 Nairobi Kenya
Pilot Line: +254 20 8070408/9

Tel: +254 20 3537842
Fax: +254 20 8561077
Mobile: +254 734 888022, 710 888022
Email: kcal@kca.ac.ke
Website: www.kca.ac.ke

SCHOOL OF GRADUATE STUDIES AND RESEARCH

KCA/SGS/Sept. 21/1

17th September 2021

TO WHOM IT MAY CONCERN

Dear Sir/Madam,

RE: KANYAMBU MBATHA FLORENCE REG NO: 15/01360

It is my distinct pleasure to introduce to you Ms. Florence Kanyambu who is a student in our institution pursuing a Master of Science in Data Analytics at the College of Technology

Florence is conducting a research on a topic titled: "*A System Dynamics Model for Credit Risk Modelling and Simulation: The Case of Licensed Credit Reference Bureaus In Kenya*" which is part of the requirements of the program she is pursuing. The research as well as the data procured thereof shall be used for academic purposes only.

Any assistance accorded to her is highly appreciated.

In case of further inquiry, do not hesitate to contact the undersigned.

Yours faithfully,

Dr. Nyaribo Misuko

Dean, School of Graduate Studies & Research
