UNIVERSITY OF THE WITWATERSRAND

SCIENTIA ET LABORE

JOHANNESBURG

Country risk analysis: An application of logistic regression and neural networks

School of Statistics and Actuarial Science

By

Gugulethu Ncube

Student Number: 461227

Supervisor

R Krommenhoek

Submitted in partial fulfilment of the Master of Science in Mathematical Statistics degree

June 08, 2017

Abstract

Country risk evaluation is a crucial exercise when determining the ability of countries to repay their debts. The global environment is volatile and is filled with macro-economic, financial and political factors that may affect a country's commercial environment, resulting in its inability to service its debt. This research report compares the ability of conventional neural network models and traditional panel logistic regression models in assessing country risk. The models are developed using a set of economic, financial and political risk factors obtained from the World Bank for the years 1996 to 2013 for 214 economies. These variables are used to assess the debt servicing capacity of the economies as this has a direct impact on the return on investments for financial institutions, investors, policy makers as well as researchers. The models developed may act as early warning systems to reduce exposure to country risk.

Keywords: Country risk, Debt rescheduling, Panel logit model, Neural network models.

Plagiarism declaration

I, Gugulethu Ncube, declare that this MSc research report is my own unaided work. It is being submitted as a partial fulfilment for the degree of Master of Science. This work has not been submitted for any degree or examination in any other University, to the best of my knowledge.

Signature of Student

Date

Acknowledgements

I owe my deepest gratitude and am profoundly indebted to the following people for their invaluable roles throughout this research:

1. Mrs Krommenhoek, who encouraged, guided and supported me from the commencement of my research to the final level and provided the means for me to gain a better understanding of the subject under study;

2. My parents, for their undying support and love;

3. My friends, who supported me in many respects throughout my research.

Contents

List of Figures

List of Tables

List of Acroynms

Acroynm	Description
AIC	Akaike Information Criterion
AUC	Area Under Curve
BIC	Bayesian Information Criterion
EDA	Exploratory Data Analysis
EIU	Economist Intelligence Unit
FPF	False Positive Fraction
GCF	Gross Capital Formation
GDP	Gross Domestic Product
GEE	Generalized Estimating Equations
GNP	Gross National Product
IBCA	International Bank Credit Analysis Limited
ICRG	International Country Risk Guide
IMF	International Monetary Fund
LR	Logistic regression
MLE	Maximum Likelihood Estimation
NN	Neural network
NPV	Negative Predictive Value
PPV	Positive Predictive Value
PRS	Political Risk Services
ROC	Receiver Operating Characteristics
Rprop	Resilient Backpropagation
SAS	Statistical Analysis Systems
S&P	Standard and Poor's
TPF	True Positive Fraction
WGI	World Governance Indicators

Chapter 1

Introduction

1.1 Background of the Study

The last decades have seen a deterioration in the world economic environment which has led to an interest in the study of country risk. The 1960's and 1970's were filled with political crises, while the 1980's with debt crises which saw many underdeveloped countries with large debts from foreign banks failing to service them in the agreed time frame. Lastly, the 1990's had financial crises which were unique for different countries as they were triggered by different factors. These historical events are still relevant and are used as a guide in present day to help adopt a holistic view and create general guidelines to apply in the assessment of countries at risk of debt rescheduling (Bouchet, Clark and Groslambert, 2003).

There has recently been a European debt crisis which emerged after the recession in 2008 and has seen countries like Greece, Portugal, Ireland, Spain and Cyprus failing to repay their debts. The Greek debt crisis began in 2009 after it announced that it had been downplaying its yearly deficit. Greece's economy declined by 26% from 242 billion Euros in 2008 to about 179 billion Euros in 2014. In 2010 and 2012, Greece received bailouts valued at 240 billion Euros and 264 billion Euros from the International Monetary Fund (IMF) and European states respectively (ABC, 2015). Greece was the first developed nation to default on an IMF loan valued at 1.5 billion Euros in June 2015, leading to a third bailout worth 88 billion Euros (ABC, 2015). An assessment of country risk which is the risk related with lending or investing in a country is therefore relevant as countries classified as being risky are faced with tax increases, budget cuts, difficulties in securing loans and lack of confidence among investors (ABC, 2015).

The financial institutions that give loans to risky countries face potential losses as many may need to reduce the value of the debts they are owed or as some countries default completely on their debts. In order to minimise the potential effects of country risk, international banks and risk rating agencies have been trying to identify and explain the general factors that lead to these difficult situations (Bouchet *et al.*, 2003). According to Kosmidou, Doumpos and Zopounidis (2008), the various dimensions found in country risk are a result of the different causes of the economic and financial situations which may stem from political and civil unrest and external factors such as droughts or floods as well as other factors.

1.2 Rationale for the Study

The results obtained from country risk analysis are of importance to banks, multinational corporations, governments, policy makers and investors operating in different markets worldwide as pre-decision and post-decision making tools. According to Henisz and Zelner (2010) these organisations will gain a competitive advantage only if they are able to anticipate and manage risk. Experts use different macro-economic and financial indicators such as gross domestic product (GDP) and inflation to determine risk ratings for the classification of countries. These ratings give investors and lenders the platform to decide whether to take a "gamble" in an investment or lending opportunity (Becerra-Fernandez, Zanakis and Walczak, 2002). A good "gamble" would yield profits for investors and financial institutions especially if they are the first to enter the markets.

Although some organisations offer country risk assessment services, the ratings they provide are general and may not be specific to the actual project being considered by the multinational corporation. According to Asiri and Hubail (2014) risk rating agencies take time in lowering credit risk ratings in the face of crisis. It is therefore necessary to predict how the macro-economic conditions will change. The quantitative methods that have been used for country risk assessment by banking institutions and risk rating agencies include both parametric and non-parametric methods, where the latter involves fewer model assumptions. The statistical techniques commonly used include the use of logit and probit analysis, principal component analysis, factor analysis, cluster analysis, discriminant analysis and regression analysis. The non-parametric methods include neural networks (NN), rule induction, fuzzy set theory and rough sets (Kosmidou *et al.*, 2008). The diversity of the empirical literature and research aimed at developing the best country risk models indicates the importance of this area.

Furthermore, country risk is widespread throughout the global markets. Figure 1.1 is a heat map indicating the levels of risk for different countries. The colour scale ranges from green, yellow, orange and red with green representing the least risky countries and red showing the riskiest countries. The map shows that most developed nations like Norway, Switzerland, Singapore, Luxembourg, Sweden, Denmark, Finland, Netherlands, Canada and Australia have lower levels of country risk while most African countries are flagged as being risky as well as countries like Iraq, Iran, Afghanistan and their neighbours. An understanding of the factors that make the red and orange zones risky climates is critical to loan and investment decisions.

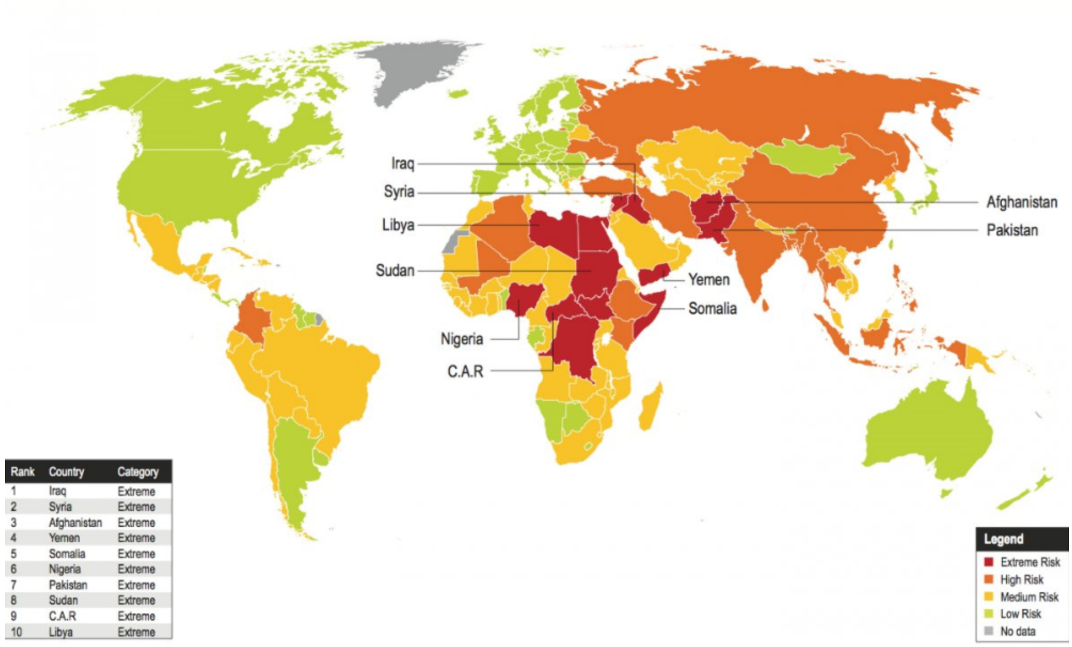


Figure 1.1: The distribution of country risk (Verisk Maplecroft, 2015)

1.3 Aim and Objectives

The study seeks to classify 214 countries into "risky" and "stable" economies by identifying countries that will fail to repay their external debt probably due to political, social and economic factors. The results can be used for the early identification of risky countries and will assist in the reduction of possible financial losses associated with incorrect and inefficient decisions. The objectives of the research are:

1. To emphasise the importance of country risk evaluation and to analyse the causes and factors of country risk assessment identified by risk analysts and researchers;

2. To identify the important political, macro-economic and financial factors to aid in investment and loan decisions involving countries;

3. To clarify the merits and demerits of the quantitative country risk assessment methods that will be used;

4. To develop a logistic regression (LR) model and artificial NN model for modelling country risk that is capable of evaluating country risk taking into account the relationships between the predictor variables to be used;
5. To compare the best results obtained from the developed models using real world historic data and propose a model for classifying countries at risk.

1.4 Research Hypotheses

1. The adoption of country risk analysis using economic, political and financial risk factors strengthens the ability to identify countries that will default on their debt.
2. Financial, economic and political factors are all relevant in classifying high risk countries.

1.5 Research Questions

1. Which variables are of importance in the development of a country risk model?
2. What is the best NN architecture to be used in the classification of countries?
3. What is the predictive power of the NN model to be developed as compared to LR?
4. What is the relationship between the predictor variables and the relationship between each predictor variable and the dependent variable?

1.6 Limitations of the Study

The country information available is for those countries that have statistical systems including official sources and the World Bank country management units. This means that countries that do not have statistical sources are excluded from the research. In addition, some countries do not want to share their debt information as they may consider it as being of sensitive nature. Country risk is also a diverse and broad subject with many elements and varying contributing factors so this study will be limited to modelling the risk of debt rescheduling using variables available from the World Bank that are commonly used in literature. Natural occurrences such as floods, earthquakes and drought that may escalate country risk are not modelled directly but these are factored indirectly through the macro-economic variables that are affected by these conditions.

1.7 Structure of the Research Report

The research is composed of a literature review (Chapter 2) describing country risk and its aspects. The methodologies of country risk used by other authors are also presented to give a better understanding of the subject. This chapter ends with model building aspects detailing the methods that are used to develop the models. Chapter 3 presents the data used as well as a step by step methodology of the country risk assessment model building process while Chapter 4 summarises the model development process. Chapter 5 presents an analysis and assessment of the models and lastly, Chapter 6 will discuss the general conclusions obtained from the research.

Chapter 2

Literature Review

2.1 Introduction

With the subject of country risk becoming more important, it is necessary to develop an understanding of the country risk elements in order to be able to identify the attributes that could interfere with a country's capability to pay back its external debt. Definitions of the different types and aspects of risk including macro-economic, financial and political risk provide a good comprehension of the key factors involved in country risk modelling. It is also necessary to understand the parametric and non-parametric approaches that have been commonly used in literature to model country risk with a strong emphasis on the advantages, disadvantages as well as limitations of the different methods. The aspects involved in the NN and LR approaches are also detailed within the literature review to give insights and identify the methods to apply in the country risk model building process.

2.2 Overview of Country Risk

There has been a growth of interest among private and official lending institutions as well as investors on the subject of country risk. Researchers have given various definitions of country risk. Meldrum (2000) defines risk as the uncertainty that occurs when an event lacks some pre-specified requirements.

Nath (2009) differentiates between the risk related to obtaining a loan from a sovereign government and the risk related to lending or investing in a country and defines country risk as the risk associated with factors that deter sovereign states or organisations from these countries to fulfil their obligations to foreign investors or lenders. This is because business transactions become risky when they occur across international borders as there may be additional risks present. The nations may have differences in terms of policies, geography, currencies or economic structures.

Country risk also exists despite the country's economic level. The 2016 USA elections resulted in highly volatile financial markets across the world due to uncertainty among investors on the impact on the global economy and trade caused by Mr Trump's victory. Analysts at Goldman Sachs believe that Mr Trump will pursue a looser fiscal policy that may result in higher interest rates from banks leading to inflation (The Telegraph, 2016). The British exit from the European Union (Brexit) also occurred in June 2016 and resulted in volatile global markets with the British pound falling to its lowest level in decades. According to The New York Times (2016), the British nationals voted to exit the Union to regain their identity, culture, place in the world and to minimise immigration which creates a sense of abandonment among the poor and working class Britons. It is projected that this decision will result in loss of trade and investment for Britain as it will no longer have access to the Unions open trade.

Calverley (1985) argues that country risk involves undesirable macroeconomic or political climates which result in economic and financial losses. Cosset, Siskos and Zopounidis (1992) describe country risk as the likelihood that a nation will back out from paying its external loans to creditors due to its inability to generate adequate foreign exchange. Meldrum (2000) discusses how researchers have separated country risk into six sections that are interrelated and very often overlap. These country risk categories are economic, transfer, exchange rate, neighbourhood, political and sovereign risk. Heinrichs and Stanoeva (2012) stated that most researchers fail to differentiate country risk from sovereign risk and use the latter as a proxy of country risk causing difficulties in corporate modelling.

2.2.1 Sovereign risk

Sovereign risk may refer to a country as a whole or to the government itself and is described by Ghose (1988) as the risk that arises when a sovereign Government renounces its external obligations and hinders local individuals from fulfilling their obligations. This is different from country risk, which, according to Heinrichs and Stanoeva (2012), is the negative side of a country's business and legal environment, corruption levels and other socio-economic factors. In financial lending, it is fundamentally important to have an understanding of the factors that prevent loan defaulting and repudiations (Eaton and Gersovitz, 1981). Although it is often difficult to ascertain if a borrowing nation will comply with the loan agreements, it is possible to analyse whether a country is vulnerable and if it has weak fundamentals that may trigger a crisis if there are sudden shifts (Christl and Spanel, 2001). This vulnerability is the one that impacts on a country's ability to service its debt and is dependant on a country's wealth, the strength and stability of its economy, internal and external stability, income derived from exports, the size of its debt burden and the available liquidity dedicated to the servicing of its debt.

In discussing the issues that make a country to either service or default its debt payment, Mellios and Paget-Blanc (2006) argued that a country is more likely to pay its debt to ensure that it has access to future loans. Moreover, these countries are driven by the need to prevent financial institutions which have assets in the country from seizing them and by so doing, maintaining both the country's reputation and its positive impact on international trade. Bulow and Rogoff (1989) stated that a country has a higher probability of defaulting if cash in advance commitments allow it to evade future stochastic payments and lending as the amount a country is supposed to pay may fluctuating at different times.

In this research, the definition of default will be in line with the study by Feder and Just (1977) and will be described as any instance in which public or publicly guaranteed loan payments to lending institutions are delayed or rescheduled without the consent of the financial institution. Owing to the close relationship between country risk and sovereign risk, the political, economic and financial variables that have been used to model them are similar. In some cases, sovereign risk is used to model country risk as they have similar influential factors.

2.2.2 Political risk

Political risk has been a critical attribute to consider in business judgements that involve foreign markets. Simon (1992) describes political risk as social and political developments that influence the value or withdrawal of foreign investments. Hoti and McAleer (2002) state that political risk is caused by political forces in combat, domestic and external strife, land disputes, rebel attacks and revolutions. Moreover, they discuss how social forces such as religion, income and ideological differences may ferment political unrest.

Governments may also cause political risk through capital controls, hiking taxes, expropriation of private property and the freezing of assets. This may result in delays in fund transfers which, in turn, will affect the investment profits. Changes in politics due to social factors, non economic factors and the reorganisation of government reign are said to be the causes of political risk (Meldrum, 2000). He adds traditional political analysis as a factor causing this type of risk. According to Kosmidou *et al.* (2008), political risk also encompasses the readiness of foreign debtors to service their loans as the decision to reschedule is political in nature. Simon (1992) highlights that there has never been a reliable method of predicting future social and political crises as there are too many situations that can unravel in the host countries.

2.2.3 Neighbourhood risk

Neighbourhood risk materialises when a country experiences an influx of politically motivated migrations from citizens from the neighbouring war-torn countries. This results in spillover effects as the countries surrounding the politically unstable country also become risky investment areas (Meldrum, 2000).

2.2.4 Economic and financial risk

Economic and financial risk refer to fluctuations in the economic conditions of a country such as prices, interest, foreign exchange rates and decline in the terms of trade (Bouchet *et al.*, 2003). This risk also involves poor investment of foreign funds as well as unwise lending by banks (Nagy, 1988). Meldrum (2000) highlights that economic risk may be due to depletion of resources, demographic changes and the distribution and creation of wealth.

2.2.5 Transfer and exchange rate risk

Transfer and exchange rate risk involves the placing of capital controls by foreign governments which make it difficult to return profits, capital and dividends to the countries that have foreign investments. This risk is difficult to quantify as governments normally place capital controls in response to crises. Exchange rate risk is defined as the devaluation of currency or sudden changes in the foreign currency exchange rate (Meldrum, 2000).

The definitions of country risk, fail to highlight the differences between natural risk which includes the weather and seismic occurrences that can impact the business and economy and the man-made sources of risk such as socio-political risks and economic risks.

2.3 Sovereign Rating Agencies

Risk assessment organisations such as Fitch, Standard and Poor's (S&P) and Moody assess the likelihood that a borrower will fail to repay their debt (Cantor and Packer, 1996). These three agencies are considered to be the best and have been providing ratings since the 1980's debt crisis in third world countries. They combine economic, financial and political measures into a risk rating (Hoti and McAleer, 2002). These agencies also run stress tests to assess the ability of the countries' economies to overcome crises. The sovereign credit ratings they developed are alphabetical indicators where higher grades represent a lower probability of risk.

2.4 Macro-economic Indicators of Risk

Substantial judgement on country risk is based on knowledge from experts. Political, economic and financial risk indicators have been adopted for use in present day by risk analysis agencies. McAleer, da Veiga and Hoti (2011) discuss how various risk analysis agencies namely Moody, S&P, International Country Risk Guide (ICRG) and Political Risk Services (PRS) have been formed as a result of the importance of country risk analysis. Other agencies are Euromoney, Institutional investor and the Economist Intelligence Unit (EIU). These organisations combine qualitative and quantitative information to establish a nation's risk index. Nath (2009) emphasises the need to make use of political, economical and financial variables as indicators of profitability and stability of a borrowing country or foreign direct investment host country to ensure returns on investments and prevent rescheduling of loans. However, he indicates that there is need for continued research in country risk analysis due to the failure by these agencies to forecast a number of financial catastrophes.

The importance of country risk analysis came into play after the first and second oil price increase in 1973 to 1974 and 1979 to 1980 respectively in which most countries failed to service their debts (Nath, 2009). Prior to the crisis, Avramovic (1968) developed indicators for assessing the risk in countries so as to reduce undesirable financial outcomes. These indicators are:

1. Growth rate of export volume;
2. Ratio of debt service payments to exports;
3. Ratio of foreign exchange reserves to imports;
4. growth rate of GDP;
5. the ratio of investment to GDP;
6. the ratio of exports to GDP; and
7. the rate of price increases.

Christl and Spanel (2001) described the indicators as showing different aspects of a country's condition. They identified that the wealth and strength of a country's economy are best explained by GDP per capita, GDP growth, internal economic stability, the ratio of budget deficit to GDP as well as the inflation rate. On the other hand, they stated that the external equilibrium of a country is characterised by the current account balance as a ratio of GDP, its foreign debt to GDP ratio and its debt servicing capacity which is obtained by measuring the export product concentration ratio.

Some analysts build models to replicate the results from the risk agencies while others use an indicator of debt rescheduling as the dependent variable. They use political and economic indicators to assess the important factors that determine the risk ratings or risk of debt rescheduling such as those listed above. In their research, Asiri and Hubail (2014) analysed a sample of 70 countries from 2006 to 2011. They used country credit risk ratings from Euromoney and EIU as their dependent variables in two separate analysis. They concluded that the political risk indicator is the most significant while GDP and gross capital formation (GCF) were the important economic variables. They also discovered that the export growth and foreign debt to export ratio were significant in determining the ratings from the EIU. Other variables they used in their initial analysis were international reserves to imports ratio and current account balance on GDP.

Canuto, Dos Santos, and de Sa Porto (2012) surveyed the macroeconomic determinants of sovereign risk commonly used by rating agencies. They identified income per capita, public debt to GDP ratio, external debt, inflation (which they considered as one of the best barometer of risk), credit risk to GDP ratio and the trade to financial openness ratio as being common in literature. The countries that were classified as being less risky in their research had a high income per capita, low public debt to GDP ratio, low inflation and low external debt.

Aguiar, Aguiar-Conraria and Gulamhussen (2006) used data from more than a hundred countries to assess the factors that affect foreign direct investment decisions in Brazil and they also calculated the probability of non investors investing in the future. They used Euromoney's country risk rating which is a combination of credit ratings, economic stability and political risk of a country to determine which variables to use for their analysis. They discovered that there was 90% correlation between the three factors used in the rating and hence chose to use one variable to avoid over-fitting. They chose a credit ranking variable which is considered by Moody, S&P and International bank credit analysis limited (IBCA) as a good measure of economic and political stability. This credit ranking index is higher for countries with a lower risk. They concluded that a riskier climate in Brazil resulted in a lower number of investors using the information from the hundred countries used in the study.

2.4.1 The relationship between macro-economic variables and risk

Cantor and Packer (1996) gave detailed descriptions of the relationship between the variables they identified as being repeatedly used by rating agencies to show a country's willingness to service its debt. They argued that:

1. Per capita income indicates a level of political stability and that a borrowing country is more likely to service its debt if it has a larger tax base;

2. GDP growth indicates economic growth and a higher value shows that the nation can manage to repay its debt over time;

3. Inflation is an indicator of the financial position of a government with a high value showing financial problems;

4. External debt is indicated by a low balance in the current account which shows that both the private and public sector of a country rely on financial loans from external sources;

5. The economic development of a country indicates whether it will default or not. Those that reach a certain economic threshold being considered to have a minimal chance of defaulting;

6. Default history of countries is highly predictive of future defaults with those that have failed to repay their foreign debt being considered as highly risky.

Using the variables described above, Cantor and Packer (1996) built a multiple linear regression model to determine the importance of each of the variables. Their study showed that external debt balance (4) explained 10% while the remainder of the variables explained 90% of variation in the model.

2.5 Political Indicators of Risk

Debt rescheduling may be influenced by the political climate of a country. A country may postpone debt payment to reduce government budgetary constraints during times of political instability (Citron and Nickelsburg, 1987). Balkan (1992) developed a probit model to determine debt rescheduling based on two political variables which are democracy and political instability. Their study concluded that there is an inverse relationship between rescheduling and level of democracy and a direct relationship between rescheduling and political stability. Simon (1992) discusses how identifying the key political and social variables is as important as selecting a methodology for the analysis. Political risk variables are grouped into macro and micro risk variables. Macro political risk variables are aimed towards foreign businesses which are affected by the actions and policies of a host country. Micro political risk variables affect only selected businesses or sectors of a country (Robock, 1971). The political risk variables for both international banks and multinational organisations are summarised in Table 2.1.

Table 2.1: Political risk variables (Simon, 1992, p. 126)

Source	International bank	Multinational enterprise (MNE)
Macro-government	Repudiation of all foreign debt, suspension or cancellation of interest payments and demands for debt rescheduling.	Expropriation or nationalization of all foreign business, remittance restrictions for all foreign business and foreign exchange controls.
Micro-government	Repudiation of selected debt, suspension or cancellation of selected interest payments and demands for selected debt rescheduling.	Selective expropriation of foreign business assets, discriminatory taxes, local content laws and breach of contract.
Macro-societal	Interest group or public demands for repudiating debt, general instability due to revolution, civil war, etc.	Terrorism or violence directed at all foreign business, general instability due to revolution, civil war, etc.
Micro-societal	Societal protests against a particular bank, terrorist attacks against a particular bank.	Strikes, boycotts, etc. against a particular MNE and terrorist attacks against a particular MNE.

2.6 Methods used for Evaluating Country Risk

The impact of globalisation and constant changes in the fiscal, economic and political climates of countries have necessitated the use of improved country risk evaluation techniques. Quantitative methods are useful in establishing the relationship between country risk variables (which have been discussed earlier) and in modelling the risk exposure which is normally binary in form eg. (rescheduling debt or not) or (defaulting in debt or not). The major quantitative techniques applied in assessing country risk include discriminant analysis, principal component analysis (PCA), logit models, probit models, tobit models, classification and regression trees and artificial NN (Nath, 2009). These techniques allow the evaluation of the useful measures used in predicting country risk. The techniques that are going to be employed in this research are LR and artificial NN.

2.7 Logistic Regression

The LR model belongs to a class of generalized linear models and has over the years been integrated into studies describing the relationship between a binary response variable and one or more predictor variables that may be continuous or categorical in nature. It was developed in the 1970's as an alternative option to the ordinary least squares (OLS) regression method. The latter method has many limitations due to strict statistical assumptions that surround it. Some of the assumptions given by Hosmer, Lemeshow and Sturdivant (2000), include:

1. The conditional distribution of the errors has a mean of zero;
2. The observations X and Y are independently and identically distributed;
3. Outliers must be removed from the model;
4. There should be no multicollinearity.

LR is considered as a superior modelling technique to other parametric methods as it allows the modelling of modern day binary decisions for example presence or absence of a disease, defaulting or non-defaulting loan applicants, passing or failing students and in our case debt rescheduling and non-debt rescheduling countries. The aim of the model development process is to obtain the best fitting realistic and parsimonious model to explain the relationship between a binary outcome and its explanatory variables (Hosmer *et al.*, 2000).

2.7.1 Univariate logistic regression model

The simplest LR model involves one predictor variable, X and one binary response variable, Y. The central concept that underlies this model is the use of the logit transformation which is the natural logarithm of an odds ratio. The odds ratio measures the odds of an outcome occurring given exposure to a variable of interest. The LR model thus predicts the logit of the response variable (Y) from the predictor variable (X). This gives a model of the form:

$$\text{logit(Y)}=\ln(\text{odds}) = \ln(\frac{\pi}{1-\pi}) = \beta_0 + \beta X \tag{2.1}$$

π is used to estimate the probability that an event will occur for our dichotomous dependent variable (Y).

$$\pi = \text{P(Y=y|X=x)} = \frac{\text{exp}^{\beta_0+\beta x}}{1+exp^{\beta_0+\beta x}} \tag{2.2}$$

where β_0 is the intercept, β is the parameter or coefficient of the LR model's predictor variable and π is the probability of the outcome of interest occurring (Park, 2013).

The estimates will always fall between 0 and 1 due to the logistic transformation. If the probability is above the cut off of 0.5 then it is concluded that the event will occur and if it is less than 0.5 then it is inferred that the event will not occur (Park, 2013).

2.7.2 Multiple logistic regression model

The simple univariate LR model can be extended into a multiple LR model with many predictors ($X_1, X_2, ..., X_p$). The model is given by the following equations according to Park (2013):

$$\text{logit(Y)} \quad = \ln(\tfrac{\pi}{1-\pi}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \tag{2.3}$$

$$\pi \quad = \text{P(Y =y}|X_1 = x_1, X_2 = x_2, ..., X_p = x_p) \tag{2.4}$$

$$\pi = \frac{\exp^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots.+\beta_p X_p}}{1+\exp^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots.+\beta_p X_p}} \tag{2.5}$$

According to Stoltzfus (2011), the multiple LR model has the following model assumptions:

1. The dependent variable should be measured on a dichotomous scale;
2. There should be one or more predictor variables which can either be continuous or categorical;
3. There should be a linear relationship between the continuous predictor variables and the logit transformation of the dependent variable;
4. The observations must be independent and should have mutually exclusive and exhaustive categories;
5. There should be little or no multicollinearity between the predictor variables;
6. Large sample sizes should be used. The general rule that is applied is that there should be 10 to 20 events per covariate;
7. There should be no outliers in the model.

2.7.3 Advantages of logistic regression

The advantages of LR according to Song *et al.* (2005) are that:

1. It yields weights showing the contribution of each variable to the model;
2. The odds ratio can be calculated for each variable.

Fensterstock (2005) gives the following advantages:

1. The LR model removes redundant variables when there is high correlation between them to avoid over-fitting of the model;
2. It is possible for an individual to check the sources of error to optimise the model.

2.7.4 Disadvantages of logistic regression

1. It is difficult to implement some complex problems as compared to NN since the data are analysed using a formulated function (Ayer *et al.*, 2010);

2. Fensterstock (2005) states that variable preparation takes time and sometimes it is necessary to preselect variables using a separate analysis as it is necessary to make sure that the data used for the analysis is clean.

2.7.5 Dealing with panel data for the logit model

Country risk data is normally measured repeatedly over time therefore the between-subject heterogeneity and the within-subject correlation need to be taken into account. The generalized estimating equations (GEE) are an analysis method that is used to fit models to correlated repeated categorical responses through a mean function that relates the mean response to the regression equation for example using the 'logit' function for a LR model. The GEE procedure has the following advantages:

1. Binomial, gamma, inverse Gaussian, negative binomial, normal, Poisson and multinomial response variable distributions are supported;

2. Various link functions including the probit and the logit (to be used in this research)are supported;

3. A range of correlation structures including the first order autoregressive, exchangeable, independent, m-dependent and unstructured correlation structures are supported;

4. Allows for missing data;

5. Can perform alternating LR for binary and ordinal data.

The form of the GEE is like a generalized linear model but it is not necessary to specify the joint distribution and so there is no likelihood function (Penn

State Eberly College of Science, 2016). The variables in the GEE model are also described as follows:

1. Y is a categorical response variable where $Y = Y_{iz}$ is the response for each subject i, measured at varying time points, $(z = 1, 2, ..., n_i)$. Each y_i is a binary outcome variable and n_i is the number of time points;

2. $X = (X_1, X_2, ..., X_k)$ is a set of predictor variables which can be discrete or continuous in nature and $X_i^T = n_i \times k$ transposed matrix of covariates for each country i which will be used in the model and for k explanatory variables;

3. β is the regression coefficient.

The GEE model is given by:

$$g(\mu_i) = x_i^T \beta \tag{2.6}$$

where:

1. Random component is any distribution of the response that we can use for GLM, e.g. binomial, multinomial or normal;

2. Systematic component is a linear predictor of any combination of continuous and discrete variables;

3. μ_i is a link function;

4. The link function can be any $g(.)$ e.g. identity, log or logit (Penn State Eberly College of Science, 2016).

2.7.6 Assumptions of the model

The following model assumptions have to be satisfied for a panel logit model:

1. The responses $Y_1, Y_2.....Y_N$ are correlated or clustered and N represents the number of response variables;

2. The homogeneity of variance that assumes that all groups have similar variance does not need to be satisfied(Penn State Eberly College of Science, 2016).

2.8 The Artificial Neural Network Model

An artificial NN is a non-parametric technique used in forecasting, classification, multi-factorial analysis and pattern recognition. The network's structure resembles the biological neural system and adapts to different environments by learning from experience. This complex system may be decomposed into simpler elements for understanding. In his paper, Tucker (1996) describes the structure of NN as consisting of processing units known as neurones that have an input layer, hidden layer and output layer which are interconnected. In a basic network there is a set of nodes that operate as non-linear summing devices and connections between these nodes. Figure 2.1 is an illustration of an artificial NN with four inputs X_1, X_2, X_3 and X_4. The neural system's hidden layer consist of three hidden units H_1, H_2, H_3 and its output layer has two units Z_1, Z_2 and the two outputs are Y_1 and Y_2.

2.8.1 The processes in developing neural network solutions

The nodes resemble computational units which receive inputs and process them into outputs. The variables are presented in the input layer to the network. Major processing takes place in the hidden layer by taking the average of the weighted connections. The final results are then shown in the output layer (Gouvea and Gonçalves, 2007). The output is produced when weights of the corresponding link are multiplied by all inputs and the values summed up to determine the strength of each input connected to a neurone. These weights are adjusted when the data are presented to a network during training. An ac-

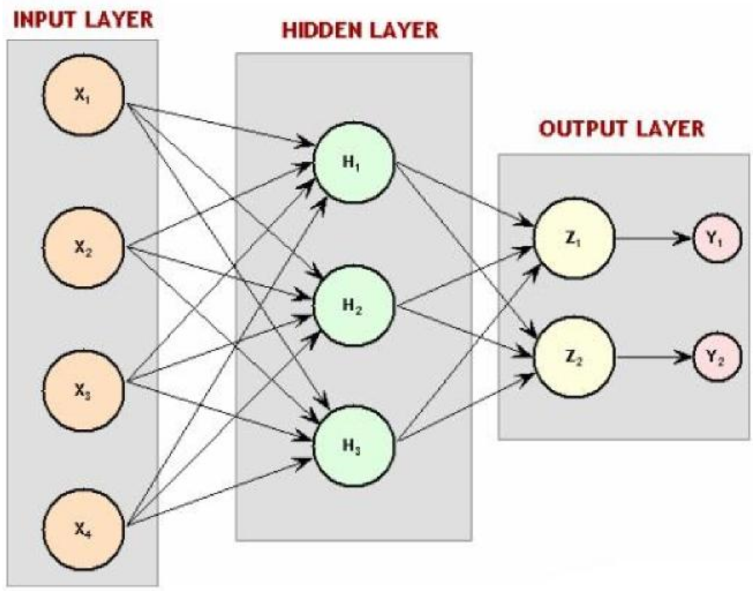


Figure 2.1: A two layer feed-forward network with four inputs and two outputs (Jones, 2004, p. 9)

tivation function such as linear, step or sigmoid is then applied to the weighted sum of inputs (Angelini, di Tollo and Roli, 2008). The most commonly used activation functions are the linear and sigmoid functions. Examples of sigmoid functions include the logistic function, hyperbolic-tangent, arc-tangent and the squash activation function (Jones, 2004).

$$\text{The sigmoid (logistic) function:} f(x) = (1 + (e^{-x}))^{-1} \tag{2.7}$$

$$\text{The hyperbolic tangent (tanh) function:} f(x) = \frac{(e^{x} - e^{-x})}{(e^{x} + e^{-x})} \tag{2.8}$$

$$\text{The arc-tangent function:} f(x) = \frac{2}{\pi}\text{arctan}\frac{\pi \times x}{2} \tag{2.9}$$

$$\text{The squash activation function:} f(x) = \frac{x}{1 + |x|} \tag{2.10}$$

NN may therefore be used for different models such as regression models and binary probit models by changing the activation function. In the training process the weights are adjusted until a pre-specified accuracy or threshold is achieved. The accuracy is calculated by comparing the NN results and the actual results obtained from the data to obtain the bias. The error calculations are important in the training and development of NN. These calculations differ depending on the NN's application. The error, E may be given by:

$$E = \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{C}(t_{ik} - \hat{t}_{ik})^2 \tag{2.11}$$

where N is the total number of training cases, C is the number of network outputs, t_{ik} is the observed output of the ith training case and the kth network output and $\hat{t}_{ik}$ is the networks prediction for that case.

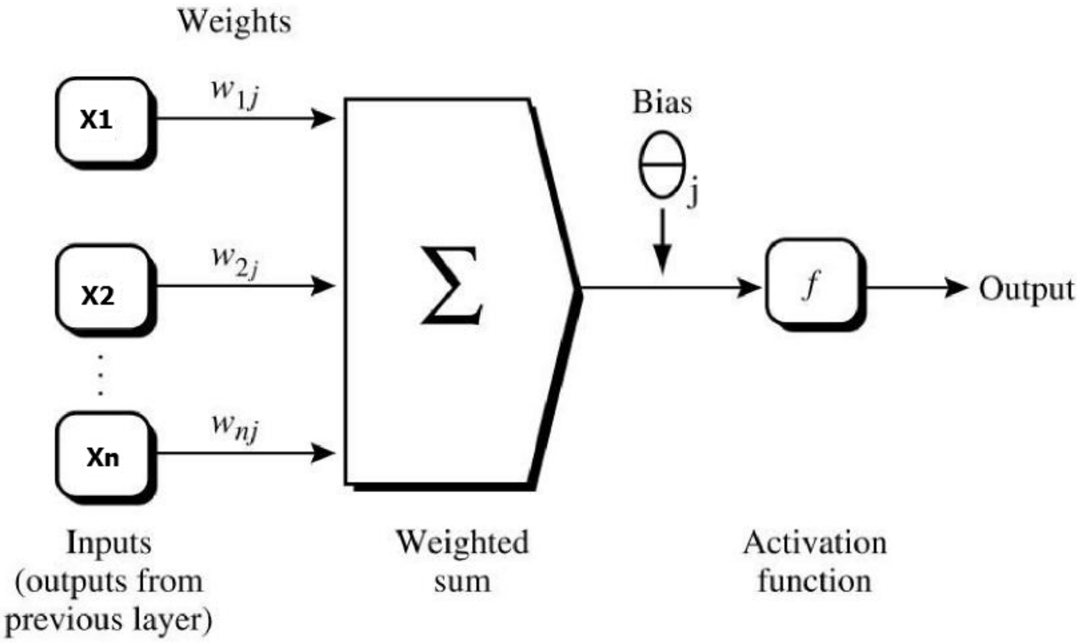


Figure 2.2: Basic structure of neural networks (Han and Kamber, 2006, p. 331)

Figure 2.2 illustrates how the NN works to reach a solution. The variables X_1, X_2 up to X_n represent the inputs. The corresponding weights are given by w_{ij} for $i = 1$ to n. The terms i and j represent the weight connections from unit i in the previous layer to unit j. The bias obtained in the network is denoted by θ_j and lastly, f represents the activation function. Successful training of the NN results in networks that can predict an output value, classify and estimate functions as well as recognise and complete patterns.

2.8.2 Neural network architecture

NN have different architectures and learning mechanisms. According to Angelini *et al.* (2008), most networks are either layered or completely connected. In completely connected networks all neurones are connected to one another. The layered networks may comprise of feed-forward network in which the connections from the input layer to the output layer are in one direction such as in Figure 2.1 or recurrent network which incorporates 'if else' loops into the network (Figure 2.3). According to Angelini *et al.* (2008), the learning mech-

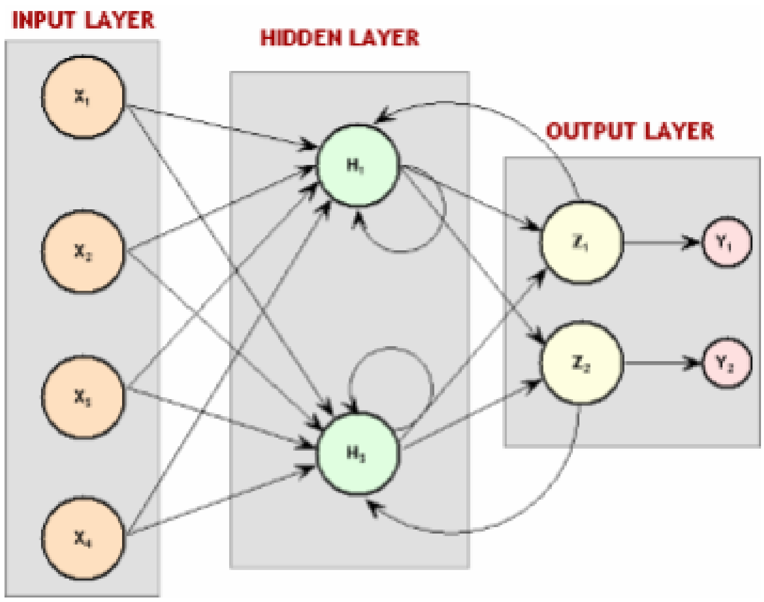


Figure 2.3: Recurrent neural networks (Jones, 2004, p. 10)

anisms that are used in NN may be supervised, unsupervised or reinforced. In supervised learning the NN are provided with a training set of correct variable problems (inputs) and the respective desirable output upon which it trains. The error produced by the network is used to adjust the weights. The network is provided with a set of inputs and no desirable output in unsupervised learning and therefore the algorithm has to guide the network towards a good output. In reinforced learning, penalties and prizes are used to modify the weights and these are a function of the network response. The network is supervised by an external reviewer (Angelini *et al.*, 2008).

The supervised learning method normally used to train networks is the backpropagation method (Tucker, 1996). This method allows for computation of the correct combination of the NN weights for complicated models that have a large number of parameters by calculating the gradient of the networks by obtaining the first derivatives of the weights in the network (Rojas, 1996). The aim of the backpropagation algorithm is to develop and adjust the weights so that the NN can learn how to accurately obtain outputs from the input variables. The purpose of the algorithm is to estimate an appropriate set of network weights based upon a training set. It achieves this through processing a training dataset iteratively and comparing the results from the training with the actual target values. It is necessary to provide a training set with some inputs as well as known correct outputs to show the network what kind of behaviour is expected (Tucker, 1996).

The backpropagation method allows the network to adapt. For each classification problem, the mean square error between the network results and the actual results is then minimised by adjusting the weights by strengthening connections which give accurate results and by weakening connections which give incorrect results. This is achieved by use of the method of gradient descent through the calculation of the gradient at each iteration for each error. It is therefore a requirement that the error function be continuous and differentiable (Rojas, 1996). Other optimisation methods that can be used to minimise the error term include the steepest descent, quasi-Newton as well as the conjugate gradient. The name backpropagation emanates from the fact that adjustments are made to the error value in a backward direction beginning from the output layer and to the hidden layers and to the first hidden layer and minor changes are made to the weights in each layer. These changes in weights are calculated to reduce the error in each case. According to (Rojas, 1996) the network weights are selected at random for initialisation and then the gradient of the error function is used to correct these initial weights.

The backpropagation is repeated until the network has learned correctly the relationship between the inputs and outputs and the overall error drops below some predetermined threshold (Tucker, 1996; Han and Kamber, 2006). According to the research by Jones (2004), the backpropagation method is the most efficient method for estimating the gradient and contributed largely by reducing the network training time as well as making it possible to train networks with a large number of inputs.

2.8.3 Advantages of neural networks

The advantages of modelling using artificial NN according to Tu (1996) are that:

1. There is no need to meet the assumptions of linearity since NN assume a non-linear relationship among predictor attributes;

2. Artificial NN can be adapted to sparse and noisy data and model selection can be applied;

3. The networks are flexible in solving complex problems;

4. NN are easily available in various software packages;

5. The networks have parallel processing ability and can perform many functions at once since computations at each node are independent of others (Topping *et al.*, 1998).

2.8.4 Limitations of neural networks

Bouchet *et al.* (2003) outlined some of the difficulties associated with NN.

1. It is difficult to perform any significance tests based on the results from NN;

2. The selection of the NN's architecture is not governed by any rules of selection. Determining which activation function to use or the number of hidden nodes to use is done on a trial and error basis and so can affect the accuracy of the training set;

3. The networks may have a large number of local minima which refers to neighbourhoods or intervals in the graph where the height of a function a is lower than or equal to height anywhere else in the interval. These minimum points occur if the objective function is not globally convex. Therefore the network may not find the best solution or it may converge towards a sub-optimal solution. This can be rectified by repeating the training with random starting weights. Rojas (1996) proposed the use of the backpropagation with momentum as a way of reducing the risk of getting a local minimum. This is done by introducing a momentum term in which instead of following the negative gradient direction, a weighted average of the current gradient and the previous correction direction is computed at each step. Making use of adaptive learning rates as well as statistically pre-processing the data to de-correlate the input patterns is useful in avoiding the effects of large eigenvalues of the correlation matrix. This is also known to help the algorithm to converge quickly.

Other algorithms discussed in Rojas (1996) included Silva and Almeida's algorithm and a variant of this algorithm known as resilient backpropagation (Rprop) which was developed by Martin Riedmiller and Heinrich Braun in 1993. The Rprop algorithm works through the adjustment of weights in the network that is being developed. It makes use of the combination of the speed with which a network undergoes training as well as the sign of the partial derivative of the error function with respect to each weight. This speeds up the learning in the regions that have local minimum values that prevent the network from converging towards a maximum solution of the error function. The flat regions or local minimums refer to areas with large errors on the error surface plane. Minimum and maximum learning rates are enforced to avoid accelerating and decelerating too quickly when the algorithm reaches these points. The dynamic adaptation algorithm is based on a global learning rate. This method develops two new points instead of one commonly used point by making use of the negative gradient direction. The point amongst the two with the lowest error is used in the next iteration. The algorithm accelerates if the point is far away by enlarging the constant. If it is the closest one, the learning constant is minimised (Rojas, 1996);

4. The networks are difficult to comprehend since they do not explain their decision making procedure (Worrell, Brady and Bala, 2012). They do not give the weighting of the variables to the user and so it is difficult to tell which variables are significant. West (2000) describes them as a black-box technology that does not make use of logic or rules;

5. The artificial NN models take a long time to train so that they give accurate results (Bouchet *et al.*, 2003).

2.9 Application of Neural Networks and Logistic Regression

Basu, Deepthi and Reddy (2011) group country risk methods used in literature as being either qualitative, quantitative or check-list. According to them, the methods applied in most research are quantitative methods and artificial NN are described as the most commonly used quantitative technique.

Artificial NN and LR have been used extensively both at firm and industrial level to solve a wide range of real world classification, forecasting and dimensionality reduction problems. The aim of these quantitative methods is to establish a clear relationship between an indicator of risk and the financial, economic and political factors. Furthermore, these methods help to identify patterns which allow the classification of countries into different risk categories.

The indicator of risk is usually in a binary form for example rescheduling debt or not, defaulting or not. These statistical techniques have also been used in evaluating the prediction strength of country risk measures published by banks and rating agencies as well as identifying the importance of different factors in the ratings of various agencies (Nath, 2009). This study will focus on the comparison of NN and LR which are shown in the literature to be superior modelling techniques with minimum error rates. Furthermore, LR models popularity may be attributed to the fact that it does not make any assumptions of normality, linearity, and homogeneity of variance for the independent variables as well as its interpretation of parameters which is simple.

Paliwal and Kumar (2009) reviewed 96 studies that compared multi-layered feed forward NN to one of the three traditional methods of regression analysis, LR and discriminant analysis. The NN outperformed the traditional methods in 58% of the cases, had equivalent performance in 18% of the cases and were outperformed in 18% of the cases. They stated how NN are preferred to the traditional statistical methods but highlighted some disadvantages of NN which were discussed in section 2.8.4.

Feder and Just (1977) applied the logit model to investigate the significance of the debt service ratio, export fluctuations, compressibility of imports, imports to gross national product (GNP) ratio, imports to reserves ratio, amortisation to debt ratio, per capita GNP and export growth on the binary dependent variable of debt servicing capacity. They used a sample of 30 countries for the period of 1962 to 1975. They found all variables except export fluctuations and compressibility of imports to be significant and their research gave the lowest error rate out of all statistical methods applied in country risk.

Saini and Bates (1984) applied logit and discriminant analysis on a sample of 25 countries containing 298 observations for the period 1960 to 1977. Their dependent variable was a combination of postponing the repayment of debt and remaining amount of foreign loans due for payment without which the countries would have rescheduled. They used a combination of different macro-economic variables such as the ratio of imports to GDP, the ratio of imports to reserves, GDP per capita, the difference in the consumer price, the rate of increase of money circulating within a country, the increase or decrease in the amount of exported goods and services, averaged over three years, adjusted for change in reserves to exports ratio, cumulative current account balance over five year period adjusted for change in reserves to exports in the latest year ratio, the ratio of the remaining foreign assets of the banking system to money supply and the increase or decrease in international reserves. They showed that the errors and coefficients generated by both models had no significant differences. The variables which they found to explain the dependant variables the most, the difference in the consumer price, the rate of increase of money circulating within a country, the ratio of the adjusted cumulative current account balance to exports, and the increase or decrease of reserves.

Cooper (2000) used a multi-layer NN trained by the backpropagation algorithm to identify countries that were most likely to reschedule their debt payments in 1983. He used a sample of 70 borrowers, 22 who rescheduled and 48 who made their loan payments. The data he used for the analysis was obtained from the Morgan Guaranty Trust Company and included four world bank economic indicators. The economic growth indicator was described as an average of the increase in GNP per capita from 1960 to 1982 while the import cover was obtained by dividing international reserves by imports in 1982. The remaining two indicators were the ratio of short term debt to exports and the debt service ratio which was described as an interest on debt from external sources and amortisation as a percentage of exports of goods and services in 1982. Their results showed that NN were able to classify the countries with a higher degree of accuracy as compared to discriminant analysis, the logit and probit methods.

Leon-Soriano and Munˇoz-Torres (2012) developed a three layer feed-forward NN using publicly available economic data for European countries for the years 1980 to 2011. They wanted to assess if the results from their model would be consistent with those from Moody and Fitch's credit sovereign ratings for the respective years. They used GDP per capita, gross debt to exports ratio, external balance relative to GDP, fiscal balance relative to GDP, inflation and unemployment rate. They used the rating time line to create a dataset and the results showed that NN should be used to model sovereign risk. They concluded that NN can accurately predict sovereign risk provided the input variables are accurately selected. Yim and Mitchell (2005) developed a hybrid NN to assess country risk. They wanted to assess whether their model could outperform traditional statistical models and the ordinary NN method. The results showed that country risk can be predicted accurately using a combination of discriminant analysis and artificial NN and still produce good results.

Bennell, Crabbe, Thomas and Ap Gwilym (2006) compared the predictive ability of a NN classification model, a NN regression model and a probit model in their research. The NN models they applied were generalised feed-forward networks. The NN classification models were used to categorize data while the NN regression model was used to predict continuous values after supervised learning, They obtained a dataset from 11 international risk rating agencies with 1383 annual observations of foreign currency sovereign ratings for the period 1989 to 1999. They used external debt to export ratio, inflation, GDP per capita, fiscal balance, external balance, GDP growth, indicator for development as well as the results from the agencies. Training and test sets were obtained from the 1999 ratings which were augmented using the 1998 ratings since they were fewer high quality ratings in comparison to the full data used. The new data then resembled the distribution found in the complete dataset. The training and test set comprised of 69.1% and 17.7% respectively, while the validation set consisted of the 1996 ratings which contributed 13.2% of the full dataset. The NN were trained and tested five times. The results that were obtained are summarised in Table 2.2. The best models for the Classification and Regression neural models and the Probit model were 42.4%, 33.9% and 31.8% respectively. The results from these models indicate that the predictive ability of the models is very low therefore, these may not be the best models to apply. The model that outperformed all others was the classification NN

Table 2.2: Predictive ability of different models

Results	Classification neural model	Regression neural model	Probit model
Best Model	42.4%	33,9%	31.8%
Model Averages	40.4%	34.6%	No average

followed by the regression NN and lastly the probit model. The best models chosen were the ones that had the lowest minimum error. As a result, in the regression based NN the best model had a lower percentage of correctly classified results but had the minimum mean absolute error.

Somerville and Taffler (1995) assessed the differences between the use of a banker's judgement and multivariate statistical techniques in predicting the creditworthiness of less developed countries over a one year time horizon. The data used for the analysis covered the period from 1979 to 1989. The banker judgement was represented by the Institutional investor's credit ratings. The dependent variable was creditworthiness which was a binary variable with the value one representing a country with a year case of arrears from financial institutions and a value zero representing those countries with no arrears in that year. Their analysis showed that bankers where biased when predicting the creditworthiness of less developed countries and classified most as having arrears. There was no type 1 error but a type 2 error of 62%. The multivariate techniques where shown to have high miscalculation costs. Discriminant analysis had a type 1 and type 2 error of 11% and 17% respectively while LR had a type 1 and type 2 error of 8% and 22% respectively. The error rates for the banker judgement, LR and discriminant analysis were 24%, 14% and 14%

respectively.

2.10 Model Building Aspects

Model building refers to the methods used in developing a model that best describes the relationship between the dependent and independent variables and selecting which independent variables to include. The model building steps that will be used in the development of the LR model and NN models as explained in Han and Kamber (2006) are listed below.

1. Data pre-processing;
2. Univariate exploratory data analysis (EDA);
3. Multivariate EDA;
4. Covariate selection methods;
5. Variable selection methods;
6. Cross validation methods;
7. Logistic regression;
8. Neural networks;
9. Model performance comparison.

2.10.1 Data pre-processing

Data pre-processing is a technique that involves transforming raw data into an usable format. Data pre-processing techniques need to be applied as most real world data consists of noise, missing values and inconsistencies. These techniques include cleaning, variable selection, transformation, sampling and normalisation (Han and Kamber, 2006).

2.10.2 Exploratory data analysis

EDA is an approach used to analyse data sets to summarise their main characteristics, often using visual methods. Primarily, EDA is a useful approach in deducing what the data can tell us beyond the formal modelling. This includes checking of model assumptions, detecting mistakes, determining the relationships between variables and preliminary selection of the models. EDA is classified as either graphical or non-graphical and as univariate or multivariate. Univariate methods look at one variable (data column) at a time while multivariate looks at two or more variables at a time (usually bivariate) to explore the relationships between these variables (Behrens, 1997).

2.10.3 Univariate exploratory data analysis

EDA of categorical variables involves the use of frequency tables. The range of these categorical values is assessed. Univariate EDA for the continuous independent variables is done to make preliminary assessments about the population distribution of the variable using the data of the observed sample. Descriptive statistics are used to learn about the data characteristics in terms of central tendency measures such as the mean, mode and median. The degree to which the numerical variables are dispersed is normally shown using the variance. These measures are used to understand the distribution of the data. The most basic graph that is used in the univariate graphical EDA is the histogram which is a bar plot in which each bar represents the frequency or proportion out of the total population. A substitute that can be used in place of a histogram is a stem and leaf plot which shows all the data points and the shape of the distribution. Boxplots are another form of graphical EDA used to visually represent the measures of central tendency. Outlying data points may be identified using this method (Seltman, 2015).

2.10.4 Multivariate exploratory data analysis

Scatter plots are often used as visual aids to determine the relationship patterns or trends between two numerical variables. This allows correlated variables and outliers to be identified within the data. Some of the methods used for detecting multicollinearity include checking for large correlations among a pair of predictor variables and the use of the variance inflation factor (VIF). Collinearity of predictor variables in regression refers to a situation where the explanatory variables are related to each other. The VIF quantifies how much the variance of the estimated regression coefficients is inflated by the existence of correlation among the predictor variables. The VIF is calculated as follows: An ordinary least squares regression is run that has predictor variable X_i as a function of all the other predictor variables in the equation. The equation is given below:

$$X_1 = \alpha_2 X_2 + \alpha_3 X_3 + \ldots\ldots + \alpha_k X_k + c_0 + e \tag{2.12}$$

where, c_o is a constant and e is the error term.

$$VIF_i = \frac{1}{1 - R_i^2} \tag{2.13}$$

where R_i^2 is the coefficient of determination of the regression equation, with X_i on the left hand side and all other predictor variables on the right side. The variable k represents the number of VIF values calculated, one for each X_i (O'Brien, 2007). If there is no correlation between the predictor variable and the remaining predictor variables the VIF value will be one. Any VIF's exceeding four warrant further investigation and those above ten are signs of major multicollinearity. The VIF can be calculated for variables in a LR model since its calculation is based on the predictor variables. Therefore, the method for VIF calculation described above is applied to the LR predictor variables.

It is important to check for multicollinearity as it may cause the estimates of the coefficients to be unstable, resulting in small changes in the dependent and independent variables producing large changes in the estimate values.

This is seen in the large variances (O'Brien, 2007). Another effect of multicollinearity is that R^2 may be large while individual slope estimates are not significant or they may have signs opposite to what is expected (Liao and Valliant, 2012). Collinearity can increase estimates of parameter variance; yield models in which no variable is statistically significant even though R^2 is large and, in truly extreme cases, prevent the numerical solution of a model (Belsley, Kuh and Welsch, 1980; Greene, 1993).

2.10.5 Variable selection methods

Variable pre-selection helps in identifying the most relevant input variables or subset of predictors. This process is necessary so as to explain the data in a simple way. Hosmer *et al.* (2000) indicated that the goal of any LR analysis is to select predictor variables that result in the best model. This is difficult because most real world applications of LR involve complex datasets that have many independent variables. They highlighted that the effect of having many independent variables in the model is that it results in over-fitting since more variables may result in larger standard errors. The LR model also becomes more dependent on the actual data leading to numerically unstable elements.

2.10.5.1 Stepwise methods

The stepwise method is a variable selection method in which variables are included or excluded based solely on statistical criteria. There are three common approaches used in developing a LR model. Variables may be added in a stepwise manner through forward selection in which the original model is empty and the most significant variable is added at each step. The stepwise backward selection method involves starting with a full model and removing the variables that are not significant. The third approach is a combination of the forward selection and backward selection and it is known as stepwise selection. In this method variables are both added and removed (Tucker, 1996; Dreiseitl and Ohno-Machado, 2002). The backward selection method is favoured as it allows the removal of variables based on their statistical significance as compared to other variables and prevents omission of variables, and it also deals with confounding variables (Edwards, 2012).

The non-linear nature of artificial NN hinders the use of significance tests therefore separate networks can be used to train each input variable and variables are added to the network one at a time to the best network to assess the effect of adding them. The advantage of variable pre-selection for NN are that the processing time will be reduced (Zhang, 2000).

2.10.5.2 Principal component analysis/ Factor analysis

PCA is a dimension reduction technique that is used to project data onto a smaller surface. It differs from attribute selection methods which select variables from their initial number and combines the essence of different attributes by creating a smaller set of variables. The basic PCA procedure involves normalising the variables so that they fall into the same range, computation of the principal components which are orthonormal vectors, ordering of the principal components in order of decreasing significance and lastly elimination of the weaker principal components (Han and Kamber, 2006). PCA is similar to factor analysis but the two methods differ in that PCA is much quicker computationally and researchers gain more information from PCA such as individual scores on certain components which can not be obtained from factor analysis. However the aim of conducting a factor analysis is to determine the factors accounting for the structure of the correlations between measured variables and therefore information relating to factor scores is not required (Fabrigar, Wegener, MacCallum and Strahan, 1999).

2.10.6 Cross validation methods

Cross validation was originally used in evaluating the predictive strength of linear regression methods. The three main types of cross validation methods are the holdout method, k-fold cross validation and leave one out cross validation. The holdout method involves splitting the data into a training and test set. The model is developed using the training set and it predicts the output values for data in a testing set. The k-fold cross validation is an improvement of the holdout method by dividing the data k times and repeating the hold-out method k times. One of the k sets are used for testing while the $k-1$ remaining variables form the training set. The leave-one out method involves training the model on all data points except one, and a prediction of that point is done. The disadvantage of this method is that it is computationally expensive. Three way data splits are also commonly used in literature. The training set is used to learn the patterns of the data and is usually the largest set. The validation set is used to confirm the results of the trained network while the test set is used to assess if the results from the trained NN or LR model may be generalised to different datasets (Kaastra and Boyd, 1996).

2.10.6.1 Bootstrapping

The bootstrapping technique involves re-sampling data with replacement several times in order to obtain an estimate of the entire sampling distribution statistic with few or no assumptions about the distribution of the underlying population. The original sample is duplicated repeatedly to obtain an extended sample that is treated as a virtual population. If the sample size being analysed is small, bootstrapping will be used. The advantages of bootstrapping are:

1. A simpler mathematical way to calculate the asymptotic distribution of a statistic is provided if computation is difficult;
2. In cases were finite samples exhibit a large sample bias, it provides better accuracy;
3. Bootstrapping proves useful in relation to partially observed dependent variables such as binary, discrete or censored variables which are small with unknown sampling distributions. Bootstrapping can be applied to identify the distributions for these types of data (DiCiccio and Efron, 1996).

The disadvantages of bootstrapping are discussed by Baser, Crown and Pollicino (2006) and include:

1. Misleading results may be obtained if the model is incorrectly chosen and since bootstrapping samples from the model it may lead to incorrect results;

2. They cause bias in the estimates since the confidence intervals and rejection probabilities are estimates.

2.10.7 Model goodness of fit measures

2.10.7.1 Criterion based measures

The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are the two commonly used methods of selecting the best models. They measure the quality of statistical models for a given data set. The AIC and BIC estimate the information lost and represents a trade-off between the complex nature of the model and its goodness of fit. The AIC and BIC resolve the problem of model over-fitting by introducing a penalty term based on the number of variables in the model.

$$AIC = -2log - likelihood + 2p \tag{2.14}$$

and

$$BIC = -2log - likelihood + plogn \tag{2.15}$$

The aim for the LR method is to minimise the AIC or BIC. Large models have smaller residual sum of squares and fit better but they generally use more predictor variables. Therefore, these two methods select the best models by balancing the model size and model fit. However, the BIC tends to penalise models with more predictor variables more heavily than the AIC making it prefer smaller models to the latter (Aho, Derryberry and Peterson, 2014).

2.10.7.2 The Wald test

The Wald test (also known as the Wald Chi-Squared test) is a parametric statistical test that is used when testing the ‘significance’ of explanatory variables in a statistical model. The term ‘significant’ means that a variable adds value to the model. The variables that are not significant are removed from the model. The Wald test can be used to test the true value of a variable based on the sample estimate. If the test fails to reject H_0 it means that removing the variables will not substantially affect the model fit. The Wald statistic, W_T, is used to test

$$\begin{aligned} &H_0 : \theta = \theta_0. \text{ vs.} \\ &H_A : \theta \neq \theta_0. \end{aligned} \tag{2.16}$$

and is calculated as :

$$W_T = \frac{[\hat{\theta} - \theta_0]^2}{1/I_n(\theta)} = I_n(\hat{\theta})[\hat{\theta} - \theta_0]^2 \tag{2.17}$$

where $\hat{\theta}$ is the maximum likelihood estimate (MLE) , and $I_n(\hat{\theta})$ is the expected Fisher information evaluated at the MLE. The variables with p-values < 0.05 are normally considered as not being significant in the model. However, for multiple LR models variables with p-values < 0.25 in its uni-variable analysis are considered as candidates for a multi-variable analysis. This is because some variables may become important predictor variables when taken together than when taken individually so using the traditional p-value of 0.05 may exclude important variables from the analysis (Agresti, 1996).

2.10.7.3 The score test

The score statistic indicates how sensitive a likelihood function is to its parameter, that is the score for θ is the gradient of the log-likelihood with respect to θ. The score tests are used in estimating the model improvement if more variables are added to the model. This test is sometimes referred to as a test for omitted variables (Weesie, 2001).

2.10.7.4 The likelihood ratio test

This is a test used to compare the goodness of fit of two models where one model is simpler or more parsimonious than the other. The test is based on the likelihood ratio which estimates how many times more likely the data are under one model than the other. It tests:

$$\begin{aligned} H_0 &: \text{The reduced model is true. vs.} \\ H_A &: \text{The current model is true.} \end{aligned} \tag{2.18}$$

where the reduced model refers to a model which omits m predictors and the current model is the one that includes them. The likelihood-ratio statistic is given by

$$\Delta G^2 = -2logL \text{ from reduced model } -(-2logL \text{ from the current model})$$

where m is the degrees of freedom and the p-value is $P(\chi^2_m >= \Delta G^2)$ (Penn State Eberly College of Science, 2016).

2.10.7.5 The Hosmer and Lemeshow goodness of fit test

The Hosmer and Lemeshow statistic is an alternative method of testing the model goodness of fit. The test is used to assess whether observed event rates match expected event rates. According to Hosmer *et. al* (2000), it tests the following hypotheses:

$$\begin{aligned} H_0 &: \text{The current model fits well. vs.} \\ H_A &: \text{The current model does not fit well.} \end{aligned} \tag{2.19}$$

The statistic is calculated by grouping the observations by the model predicted probabilities. The recommended number of groups is calculated by adding one to the number of covariates in the model. The Hosmer-Lemeshow statistic is highly dependent on how the observations are grouped. Furthermore, its not as accurate when you have one or two categorical predictor variables and is therefore best used for continuous predictor variables. The statistic is given by:

$$\sum_{l=1}^{g} \sum_{j=1}^{2} \frac{(obs_{ij} - exp_{ij})^2}{exp_{ij}} \tag{2.20}$$

where g= Number of groups and the test used is the chi-square with $g-2$ degrees of freedom.

2.10.7.6 The logistic regression R^2

The R^2 statistic refers to the proportion of the variance in the dependent variable that is explained by the predictor variables, with larger R^2 values indicating that more variation is explained in the model. The coefficient of determination for models with categorical dependent variables is estimated by the Cox and Snell's R^2, the Nagelkerke's R^2 and the McFadden's R^2. These pseudo R^2 values are essential in evaluating multiple models predicting the same outcome on the same data (Veall and Zimmermann, 1996).

2.10.8 Neural network paradigms

In the construction of artificial NN, it is necessary to determine the number of input nodes, output nodes, hidden nodes and the number of hidden layers. The selection of these parameters are problem dependent (Zhang, Patuwo and Hu, 1998). The NN architecture is done on a trial and error basis. The correct number of hidden layers that improves the model fit has to be selected. The problem of over-fitting that can arise from using many hidden layers will be solved through the use of additional techniques such as cross validation. The ideal number of hidden nodes is selected graphically by comparing the use of different number of hidden nodes in the same network. The number of hidden nodes are plotted against the predictive performance of both the training and test set and the hidden nodes which gives the best accuracy results in terms of the root mean square error (RMSE) for both the training and test sets is selected (Costea, 2012).

There have been several suggestions on the formulas to determine the correct number of neurones. Hunter *et al.* (2012) implemented the formula $N_h = N+1$ for a multi layer perceptron (MLP), $N_h = 2N + 1$ for a bridged MLP and $N_h = 2^n - 1$ for a fully connected MLP where N is the input-target relation or parity number (number of input variables), N_h is the number of hidden nodes and n is the total number of neurons in the network. The results obtained in their study had 85% accuracy. Trenn (2008) determined the number of hidden nodes using the formula $N_h = N + N_0 - \frac{1}{2}$, where N is the number of inputs and N_0 is the number of outputs. Xu and Chen (2008) determined that the best number of hidden nodes leads to the minimum RMSE. They implemented the formula $N_h = C_f(\frac{N}{d} logN)^{\frac{1}{2}}$, where N is the number of training pairs, d is the input dimension, and C_f is first absolute moment of Fourier magnitude distribution of target.

Lastly, Masters (1993) suggested that for one hidden layer, the number of neurones is given by $N_h = sqrt(N_{INP}$ x $N_{OUT})$, where N_h is the number of neurones in the hidden layer, N_{INP} is the number of neurones in the input layer and N_{OUT} is the number of neurones in the output layer. The formula for two hidden layers is given by $r = \frac{N_{INP}}{N_{OUT}}^{\frac{1}{3}}$, where $N_h = N_{OUT}$ x r^2 is the number of neurones in the first hidden layer and $N_{h2} = N_{OUT}$ x r is the number of neurones in the second hidden layer. Input selection techniques such as feed forward networks are applied and the backpropagation learning algorithm is used (Cooper, 2000).

2.10.9 Evaluating the model performance

2.10.9.1 Confusion matrix

The confusion matrix (Figure 2.4) is a method that is used to determine if the outcomes being modelled have been classified correctly and incorrectly compared to the true results obtained from the data. The following will be calculated using the confusion matrix:

1. Accuracy: The total proportion of correct predictions;

Confusion Matrix		**Target**			
		Positive	Negative		
Model	Positive	a	b	*Positive Predictive Value*	a/(a+b)
	Negative	c	d	*Negative Predictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	
		a/(a+c)	d/(b+d)		

Figure 2.4: Confusion matrix (Analytics Vidhya, 2015)

2. Positive Predictive Value (PPV): The portion of accurately identified positive outcomes;

3. Negative Predictive Value (NPV): The portion of negative outcomes that were accurately classified;

4. Sensitivity: The portion of actual positive outcomes that were classified accurately;

5. Specificity: The portion of actual negative outcomes that were accurately classified (Akobeng, 2007).

2.10.9.2 Receiver operating characteristic curve

The receiver operating characteristic (ROC) curve is a graphical means of comparing different classification methods. The curve is obtained by plotting the false positive rate which is obtained by subtracting the specificity from one on the X-axis which is the probability of a target outcome being equal to one when its value is zero against the true positive rate also known as the sensitivity on the Y-axis. The sensitivity is the probability of obtaining a target equal to one when its true value is actually equal to one at varying threshold values. The ideal curve converges quickly towards the top left corner, meaning that the model is a perfect classifier and accurately predicted all cases (Fawcett, 2006).

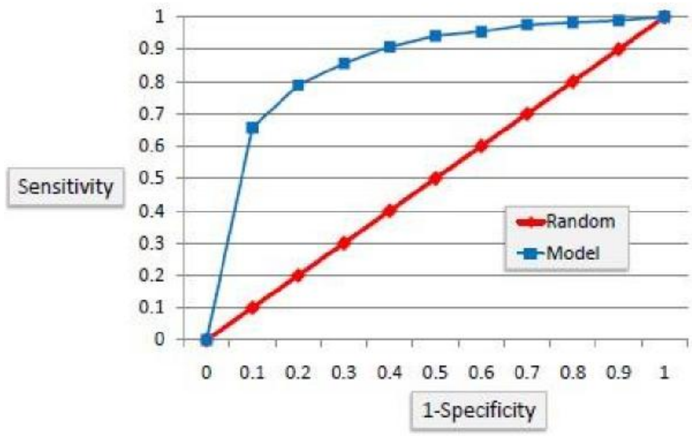


Figure 2.5: Receiver operating characteristic curve (Sayad, 2015)

2.10.9.3 Area under the curve

The area under the curve (AUC) shows the quality and performance of a binary classification method. A random classifier which is not useful will have an area of 0.5 while a perfect test will have an area of 1. Most AUC fall between 0.5 and 1 and this means that they are able to predict the outcomes better than a randomly selected classifier. The classifier can not discriminate between the two groups when the area=0.5 while an area of 1 shows perfect separation between the two groups. Therefore, the preferred AUC should be close to one for the model to be considered as being accurate (Fawcett, 2006).

2.11 Conclusion

1. Most literature shows that research on the application of NN and LR has been mainly in the area of financial and accounting problems such as credit scoring, predicting bankruptcy, detecting fraud and property evaluation (Paliwal and Kumar, 2009). Very little is shown in the application of NN and LR to country risk.

2. The numerous variables, coupled with the wide range of areas they cover (financial, legal, economic and political) make country risk modelling difficult to predict with accuracy. This explains why risk rating agen-

cies with vast resources failed to predict many changes in economies worldwide due to globalisation. Therefore, there is need for constant adaptation and inclusion of better country risk methods as well as new sources of risk that have not been considered in literature (Sviderske, 2014).

3. There are several quantitative methods used to evaluate country risk including; discriminant analysis, principal component analysis, logit models, tobit models, classification and regression trees and artificial NN.

4. LR and artificial NN are said to be the most commonly used classification techniques in biomedicine (Dreiseitl and Ohno-Machado, 2002). According to Gouvea and Gonçalves (2007) the LR and NN models produce better results to the genetic algorithm for credit scoring. Given their application in other areas such as credit risk and in medicine, it is therefore necessary to evaluate these models when applied to country risk data.

Chapter 3

Methodology

3.1 Introduction

The scope of this study seeks to apply the panel LR model and the NN model on political, economic and financial risk factors to aid in investment and loan decisions involving countries. The debt servicing capacity of countries is investigated using the effect of using exports as a percentage of GDP, external debt as a percentage of GDP, inflation, GDP per capita, GDP growth, political stability and no violence, regulatory quality, rule of law, control of corruption, government effectiveness and public debt as a percentage GDP as predictor variables. The data used in the research is obtained from the World Bank.

Data pre-processing techniques including data cleaning and data reduction will be used. Multivariate and univariate EDA will be used to check model assumptions, preliminary selection of variables and in determining the relationships among the explanatory variables. A descriptive data summary will serve as the foundation of these techniques and will help in identifying the general characteristics of patterns within the data as well as detecting the presence of noise and outliers. LR and NN models will be built to compare their performances in identifying countries that will reschedule their debt.

3.2 Data Source

The World Bank is a multilateral financial institution based in Washington. It provides loans and grants to governments and their agencies for use in development projects that are aimed at poverty reduction and in the equitable sharing of prosperity. The Development Data Group is responsible for maintaining macro and financial sector databases as well as the coordination of statistical and data work. The data developed by the World Bank is obtained from the statistical systems of member countries including official sources and bank country management units and the quality of data is dependent on their performance. The World Bank works hand in hand with bodies in countries and invests in statistical activities used in implementing standards and frameworks for collection of data, the methodologies used to analyse it and for disseminating the data. The bank checks for data consistency caused by differences in timing and reporting when combining data from different countries. They make adjustments in the balance of payments to account for fiscal/calender year differences (World Bank, 2015).

The macroeconomic and political data to be used for the analysis were obtained from the World Bank open data which gives free data access on the development of countries around the globe. The macro-economic variables were obtained from the World Development Indicators (WDI) which is a collection of development indicators that cover 214 economies for the period 1960 to 2015. It was compiled from official recognised international sources and includes exports, external debt and public debt as a percentage of GDP, inflation, GDP per capita and GDP growth. It makes available the most accurate and current data including national, regional and global estimates. The political variables were obtained from the Worldwide Governance Indicators (WGI) which has aggregates that pertain to separate governance indicators for 215 countries and territories for the period 1996 to 2014. The WGI dataset are used for research purposes and summarises the views on the quality of governance provided by a vast number of enterprises. The different citizen and expert survey respondents in industrial and developing countries are also summarised. These data are gathered from several institutions that conduct surveys, think tanks which refer to experts that provide advice on varying political and economic problems , non-governmental organisations, international organisations, and private sector firms (World Bank, 2015). The World Bank applies some aggregation rules due to the presence of missing data in which it treats the different aggregates for different grouped economies as estimates of unknown totals or mean values. It applies five methods of aggregation:

1. Missing values are imputed for group or world totals using proxy variables which have complete data for that year. Imputation is not carried out if the variable has more than a third of its values missing for that year;

2. Missing values are not imputed for aggregates concerning sums and these total sums are not calculated if the missing data exceeds a third of the total data in any given year;

3. The ratio aggregates are computed as weighted averages. The value of the denominator or of another indicator can be used as a weight. Variables with missing data are assumed to have the same weights as that from the available data and no ratios can be computed if more than a third of the dataset is missing;

4. The growth rates of aggregates are calculated as weighted averages for growth rates. They are not computed if more than a third of the data is missing. They are computed as annual averages and represented as percentages using either the least squares, exponential or geometric growth rate;

5. Aggregates concerning medians have no values shown if half the observations for countries with a population exceeding one million are missing (World Bank, 2015).

3.3 Description of Variables

The World Bank (2015) gives comprehensive definitions of the variables that were used in the analysis. The data set that was created for use in the research included variables that were commonly used in empirical literature. The ratio of debt service payments to exports, growth rate of GDP, the ratio of exports to GDP and the rate of price increases were used in a study by (Avramovic, 1968). Canuto *et al.* (2012) identified income per capita, public debt to GDP ratio, external debt and inflation as good predictors of debt rescheduling. As such, these variables were analysed in this research. The data was extracted for the years 1996 to 2013 as the political variables considered were available for that particular period. The definitions of the variables are given below:

1. The debt rescheduling variable was obtained from the World Bank and gives the entire quantity of debt rescheduled which includes the total debt value that a country owes, capital sum, interest charges and penalties that are postponed (World Bank, 2015). Countries which have not rescheduled have a value of zero while those that rescheduled have the value of the remaining sum of their debt shown. All countries that rescheduled debt will be given a value of one so that we have a binary variable;

$$\text{Debt rescheduling} = \begin{cases} 1, & \text{if debt is rescheduled.} \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

2. Total debt service as a percentage of exports of goods, services and primary income is the sum of principal repayments and interest actually paid in currency, goods, or services on long-term debt, interest paid on short-term debt, and repayments (repurchases and charges) to the IMF (World Bank, 2015);

3. GDP per capita is obtained by dividing the GDP by midyear population. GDP refers to the sum of gross value added by all resident producers in

the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources (World Bank, 2015);

4. GDP growth is an annual percentage growth rate measure of GDP at market prices and is based on the constant local currency and measures how fast an economy is growing. It is a combination of gross value added by all resident producers in the economy, plus any product taxes and minus any subsidies not included in the value of the products. Deductions arising from depreciation of fabricated assets or from depletion and degradation of natural resources are exempted when making this calculation (World Bank, 2015);

5. Inflation, consumer prices annual percentage as measured by the consumer price index is defined as the annual percentage increase or decrease in the level of prices or cost to the average consumer of acquiring a basket of goods and services that is measured at a fixed time point. Inflation may be fixed or may be tracked at specific time intervals, such as annually (World Bank, 2015);

6. External debt also known as foreign debt is the total amount of debt owed to foreign residents inclusive of governments, banks or financial institutions that may be paid back in the form of money, goods, or services. It is the total sum of all the public, publicly guaranteed, and private non guaranteed long-term debt or IMF credit, and lastly debt accumulated over a short period. The short-term debt refers to all debt having an original maturity of one year or less and interest in arrears on long-term debt (World Bank, 2015);

7. Exports of goods and services as a percentage of GDP is a variable that reflects the value of all goods and market services that are provided to other countries that are produced in the home country. The examples

of exports include the value of merchandise, freight, insurance, transport, travel, royalties and license fees. The other services that can be exported include communication, construction, financial, information, business, personal, and government services. The factor services (now known compensation of employees and investment income) and transfer payments are excluded (World Bank, 2015). The economic and financial variables are monetary quantitative measures while the debt rescheduling variable is a binary variable.

Political variables were considered for the analysis since risk rating agencies that create risk indices combine political, economic and financial information into a single composite rating. Government stability, corruption, law and order were amongst the political variables they listed as causing a country to fail to service its debt (Hoti and McAleer, 2002). These political variables are created from hundreds of variables obtained from 31 different data sources which capture information from survey respondents combining these into political indices. Therefore, the political variables that were used in this study as defined by the World Bank (2015), include:

1. Political stability and absence of violence/terrorism is a variable that reflects perceptions of the likelihood that the government will be destabilized due to its lack of integrity and may be overthrown by unconstitutional or radical means, including politically-motivated violence, coup d'états or revolutions as well as terrorism;

2. Regulatory quality is a variable that captures the perceptions of the ability of the government to formulate, promote and implement sound policies and regulations that permit and promote the development of the private sector;

3. Rule of law refers to the legal principle and of the extent to which individuals , policy makers and private entities trust and abide by the rules of society. In particular, everyone is held accountable by the quality of

contract enforcement, property rights, the police, and the courts, as well as the likelihood and frequency of crime and violence.

3.4 Logistic Regression Methodology

3.4.1 Data extraction

The World Bank data was extracted into an excel worksheet and imported into statistical analysis systems (SAS) 9.4. The debt rescheduling dependent variable was dichotomised so that it could be used in LR model building.

3.4.2 Exploratory data analysis

EDA was conducted to identify the patterns within the data. The procedures that were conducted are described below:

1. The contents procedure was employed in SAS 9.4 to give a description of the data to be used, including the number of observations, the number of variables to be used, the variable types as well as their lengths. The continuous variables wrongly classified by SAS 9.4 as being characters were changed to numeric by multiplying them by one so that SAS 9.4 recognised them as numeric;

2. EDA of the debt rescheduling binary variable was done using frequency tables to determine the frequency of countries that rescheduled debt as opposed to those that paid their debt;

3. Descriptive statistics were used to learn about the data characteristics of the independent variables in terms of central tendency measures such as the mean, maximum and minimum data values. The SAS 9.4 *means* procedure was used to obtain the statistics. The degree to which the numerical variables were dispersed was shown using the variance. These measures were used to understand the distribution of the data;

4. As a general rule, SAS 9.4 procedures deal with incomplete data by excluding the missing values from the analysis, but the way in which missing values are excluded from the dataset varies depending on the procedure being used. Table 3.1 shows how SAS 9.4 deals with missing variables for the different procedures employed in the analysis.

Table 3.1: Handling missing data in SAS 9.4

SAS Procedure	How SAS deals with missing data
proc means	All values that have no missing data are used.
proc freq	Variables with missing values are removed and cumulative frequencies are calculated from the variables that do not have missing data.
proc corr	By default the correlation calculations are based on the number of pairs with complete data. This is also known as pairwise deletion. However, the *nomiss* option can be used for listwise deletion to compute correlations for data that has no missing values.
proc logistic	If any of the variables on the model or var statements are missing they are excluded from the model.

3.4.3 Multivariate exploratory data analysis

Multivariate EDA was conducted to investigate the patterns and relationships between two or more variables in the data. The procedures that were conducted are summarised below:

1. Multicollinearity checks to identify large correlations among a pair of predictor variables were conducted using the *corr* procedure in SAS 9.4. The VIF values were obtained from the *reg* procedure. Variables with VIF values above 10 were identified removed from the model to avoid over-fitting the model (O'Brien, 2007);

2. The PCA variable reduction technique was conducted using the *princomp* procedure to create a smaller number of variables that best described the data (Han and Kamber, 2006);

3. Statistics were calculated for the categorical dependent variable and each of the quantitative independent variables individually. Univariate LR models with one covariate at a time were run and the fits analysed. All the variables whose p-value < 0.25 for the Wald statistic in the univariate analysis were used to develop the final model (Sperandei, 2014). The AIC values for each of the variables were also compared in the univariate analysis;

4. According to (Reitermanova, 2010), the hold out cross validation technique is popular due to its efficiency and simplicity. It separates the data into three mutually disjoint sets and one of its biggest advantages is that the proportions of these three sets are not strictly restricted. The training, validation and test sets used in his research were split into subsets of the proportions 50%, 25% and 25%. Cieslak and Chawla (2007) split the data used by classifiers in their research into 50% training, 20% validation and 30% testing. In their book, Camm *et al.* (2014) encouraged the use of a 50:30:20 split for the training, validation and test sets. Similar proportions 50:30:20 for the training, test and validation sets were applied in this research. Splits of 40:30:30 and 45:35:20 were used for model comparison and these proportions were used due to the flexibility available when splitting data (Reitermanova, 2010). The final models were compared to obtain the most accurate model in predicting debt rescheduling.

3.4.4 Model building

To define the panel multiple LR model, consider a country c observed over t periods of time, where $t = 1, ..., T$ and $c = 1, ..., N$. For this country there exists an unobservable random variable Y^*_{ct}, indicating latent propensity, which is observed indirectly using a binary variable, Y_{ct} such that $Y_{ct} = 1(Y^*_{ct} > 0)$, where Y_{ct} is an indicator variable that takes a value of one if the condition in the brackets is satisfied. The following equation represents the above statements:

$$Y^*_{ct} = \alpha_c + X'_{ct}\beta + u_{ct} \tag{3.2}$$

where Y_{ct} is the total amount of debt rescheduled, X_{ct} represents the predictor variables, u_{ct} is the random error term and Y_{ct} is a dummy variable given by:

$$Y_{ct} = \begin{cases} 1, & \text{if } Y_{ct} > 0. \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

Multiple LR models will be developed and the backward elimination procedure is used. Let $X_{1ct}, ...X_{kct}$ be the predictor variables in the model observed for country c at time t. When the event is equal to one, it denotes countries that have rescheduled their debt and zero denotes countries that have not rescheduled their debt. For the variable β , the likelihood of obtaining the expected outcomes Y_{ct} is dependent on variables X_{ct}. The likelihood that a country c will postpone the payment of its debt at time t is calculated as follows:

$$p(Y_{ct} = 1|X, Y) = \frac{\exp^{\alpha_c + \beta_1 ControlCorruption_{ct} + ... + \beta_k PublicDebtPercGDP_{ct}}}{1 + \exp^{\alpha_c + \beta_1 ControlCorruption_{ct} + ... + \beta_k PublicDebtPercGDP_{ct}}} \tag{3.4}$$

The probability of not rescheduling debt is calculated as follows:

$$p(Event_{ct} = 0|X, Y) = \frac{\exp^{\alpha_c + \beta_1 ControlCorruption_{ct} + ... + \beta_k PublicDebtPercGDP_{ct}}}{1 + \exp^{\alpha_c + \beta_1 ControlCorruption_{ct} + \beta_2 + ... + \beta_k PublicDebtPercGDP_{ct}}} \tag{3.5}$$

where β_k is a $(k \times 1)$ vector of predictor variables that is related to the transposed vector of X_{ct} (Lausev, Stojanovic and Todorovic, 2011).

3.5 Neural Network Methodology

3.5.1 Variable selection

The non-linear nature of the artificial NN hinders the use of significance tests. The most important decision to be made regarding a NN model is its configuration. It is necessary to determine the number of inputs, the number of outputs, the number of hidden layers and the number of neurones in each of the hidden layers.

According to May *et al.* (2011), there is no generally accepted method used in selecting the amount of attributes to apply in building NN. The authors state that selecting the variables to use in a NN model is complex due to some highly correlated variables that cause redundancy. The number of input variables available as well as the low predictive ability of some of these variables makes selection a complex problem. However, even though there is a common notion that NN are capable of removing redundant variables, NN modellers are now aware of the importance of variable selection.

For the purposes of this research, variable selection for the NN will be done using a method proposed by Garson (1991) which selects predictor attributes by comparing their significance for specific dependent variables in a NN by breaking down the model weights. This is known as the relative importance of the variables. It is based on the concept that a specific dependent variable can be determined by identifying all weighted connections between the nodes that are of concern. The scores of the input connections are calculated and scaled relative to all other inputs. One value is obtained for each explanatory variable that describes the relationship with response variable in the model. The function that was developed to calculate the comparative importance of attributes is called *gar.fun* and will be used in the R-statistical package.

3.5.2 Variable normalisation

A NN converges to a solution at a very slow rate and the number of iterations may be many if the attributes in the input layer are not normalised when the backpropagation algorithm is used (Rojas, 1996). Normalisation is a form of data transformation in which the attributes are scaled so that they fall in a specific range. These attributes may fall between -1.0 to 1.0 or 0.0 to 1.0. The min-max transformation performs a linear transformation on the original data. This transformation technique will be used to speed up the training phase of NN modelling. It is summarised below:

Let min_X and max_X be the minimum and maximum variables of an attribute, X. The normalisation procedure maps a value, v, of X to $v\prime$ in the range $[newmin_X - newmax_X]$. The equation is given by:

$$v\prime = \frac{v - min_X}{max_X - min_X}(newmax_X - newmin_X) + newmin_X \tag{3.6}$$

The advantage of using min-max normalisation is that it preserves the relationship amongst the original data values (Han and Kamber, 2006).

3.5.3 Training, testing and validating data

The data used in the research was divided into three subsets. The training set is the largest subset and has the most data points. The training set is used by the NN to learn the patterns in the data. The validation set is used as a final check of the performance of the trained network while the test set is used to evaluate the generalization ability of the network. The NN were trained, tested and validated using similar sets applied in the LR methodology for model comparability using the same training, test and validation test sizes (Jha, 2007).

3.5.4 Determining the neural network architecture

One hidden layer was used and the number of hidden neurones in the hidden layer was determined through plotting the number of hidden nodes and the RMSE. The pyramid rule proposed by Masters (1993) was used to determine the initial number of hidden nodes. The activation function which was implemented was the sigmoid activation function. A weight initialisation function was chosen for the NN model building and the network error was calculated until a specified minimum error was reached.

3.5.5 Neural network model building

The R statistical package, *neuralnet* was used for the network development and made use of the backpropagation algorithm by using the economic, financial and political data to adjust the network's weights and thresholds so as to minimise the error in its predictions in the training set. The pattern connectivity of the network is given by a weight matrix, w, where w_{ij} are the elements of the matrix and denote the weight connection from unit u_i to u_k. The pattern connectivity determines the NN structure. The three step procedure below shows how a hidden component in the output layer determines its activity.

1. The total weighted input X_k is given by

$$X_k = \sum_i D_i w_{ij} \tag{3.7}$$

   where D_i is the activity of the k^{th} unit in the previous layer and w_{ij} is the weight of the connection between the i^{th} and k^{th} unit.

2. The activity D_k is calculated by the unit using an activation function calculated based on input variables that are weighted.

3. The error is then calculated from the following formula:

$$E = \frac{1}{2} \sum_i (D_k - Y_k)^2 \tag{3.8}$$

where D_k is the extent of activity of the k^{th} unit in the first layer and Y_k is the output of the k^{th} unit that is preferred.

According to Jha (2007), the backpropagation algorithm is summarised by the following three steps:

1. A computation is done of the speed of the increase or decrease of the error as the activity of the unit in the output layer changes. Therefore, the derivative of the error is a computation of the difference between the actual and preferred output and is given by:
$$EA_k = \frac{\delta E}{\delta D_k} = D_k - Y_k \tag{3.9}$$

2. A computation is done to calculate how quickly the bias obtained alters as the input transmitted by an output unit is altered. This is obtained by multiplying step one by the frequency at which an output of unit is modified as the total input is altered. It is given by:
$$EI_k = \frac{\delta E}{\delta X_k} = \frac{\delta E}{\delta D_k} \times \frac{dD_k}{dX_k} = EA_k(1 - D_i) \tag{3.10}$$

3. A computation is done of the rate of change of the error as a weight on the network into an output node is altered (EW). EW is obtained by multiplying the activity level of the unit from which the unit emanates by step two.
$$Ew_{ik} = \frac{\delta E}{\delta w_{ik}} = \frac{\delta E}{X_k} \times \frac{\delta X_k}{\delta w_{ik}} = EI_k D_i \tag{3.11}$$
Once each error derivative of s unit is known, step two and three can be used to compute the Ew's of the following connections.

3.5.6 Evaluating model performance

The LR and NN model performances were compared through the use of the ROC curve to determine the discriminative ability of the models (Jaimes, Farbiarz, Alvarez and Martinez, 2005). Calculations of the sensitivity and specificity and model accuracy as well as the AUC are also common measures of discrimination (Dreiseitl and Ohno-Machado, 2002). The discussed methods will be used in evaluating the model to be used in predicting country risk.

3.6 Conclusion

The chapter had a detailed description of the variables selected for use in model development that were identified from previous research. The steps that were taken in building the LR and NN models were summarised including data preparation and descriptive statistics that helped in understanding the patterns within the data.

Chapter 4

Model Development

4.1 Introduction

Research is a process of inquiry that aims at discovery of knowledge, building theories, testing and validating the theories. One of the crucial steps involved in every research is model development. This process involves model selection, model fitting and model validation. These three steps are conducted iteratively until the best model for the data has been developed. This process takes into consideration the assumptions of the different models and the plots of the data, and these are used to help in developing the models. The unknown parameters are estimated and the models developed are assessed to see if they meet the modelling assumptions. This chapter will give a detailed description of the processes involved in the development of the LR and the NN models. This will assist in describing, predicting, testing and understanding complex events. Model development for LR will be conducted using SAS 9.4 and that for the NN will be conducted using the R statistical software package.

4.2 Model Development Steps

The steps that will be conducted in the model building process are shown in Figure 4.1 as well as the location of the SAS and R code.

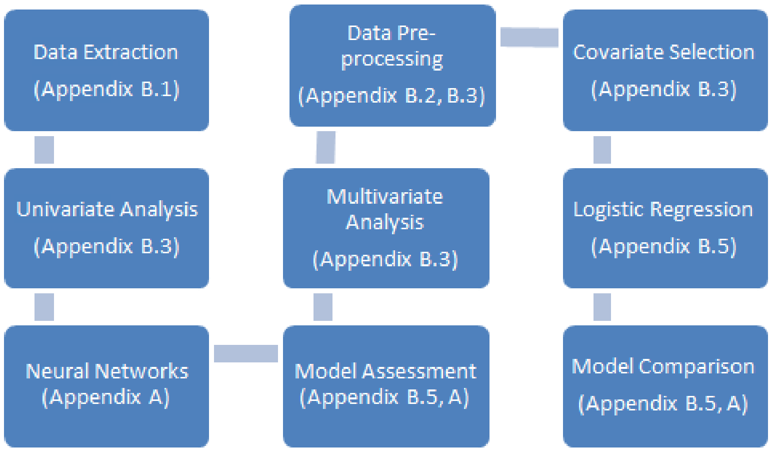


Figure 4.1: Model building process

4.2.1 Data pre-processing

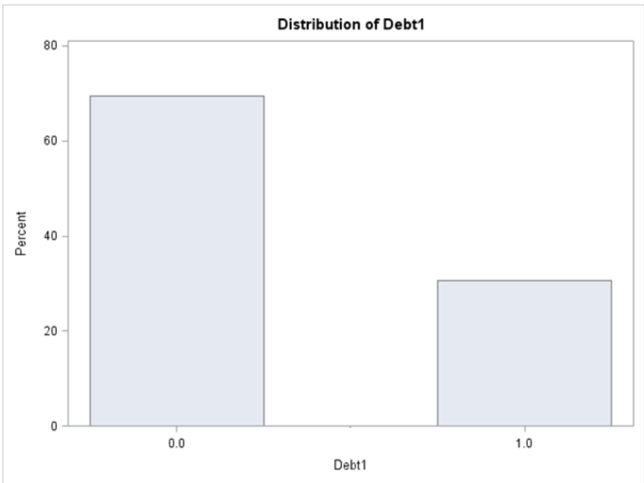


Figure 4.2: Proportion of non debt rescheduling (Debt1=0) to debt rescheduling countries (Debt1=1)

The dependent debt rescheduling variable has a value of zero for countries which have not rescheduled while countries that rescheduled debt have a value of one. The proportion of non debt rescheduling to debt rescheduling countries is shown in Figure 4.2. It indicates that are larger number of countries do not reschedule their debt and a smaller proportion reschedules their debt. The calculation of the proportion does not include the missing data as the missing data was excluded completely from the data set using the SAS software.

Table 4.1: The effect of increasing the value of the independent variables on the debt rescheduling variable

Variable Type	**Variables**	**Positive**	**Negative**
Economic	Exports as a percentage of GDP		X
Economic/Financial	External debt as a percentage of GDP	X	
Financial	Inflation	X	
Economic	GDP per capita		X
Economic	GDP growth		X
Political	Political stability of no violence		X
Political	Regulatory quality		X
Political	Rule of law		X
Political	Control of corruption		X
Political	Government effectiveness		X
Financial/Economic	Public debt as a percentage of GDP	X	

The analysis is started with 11 economic, financial and political variables. The list of each of the variables used and the possible effect that each one is anticipated to have on the debt rescheduling variable is presented in Table 4.1. The X's in Table 4.1 represent the type of impact each variable will have.

1. Positive values in Table 4.1 indicate that an increase in the values of inflation, exports as a percentage of GDP and public debt as a percentage of GDP will result in less stable economic climates that will likely lead to debt rescheduling by the respective countries.

2. An indication of a negative effect in the table shows that the variables are negatively related to the probability of debt rescheduling as these indicate a growing and stable economy.

The descriptive statistics for the countries that have rescheduled their debt and those that have not rescheduled their debt are given in Table 4.2 and Table 4.3.

Table 4.2: Descriptive statistics for countries that have rescheduled debt

Variables	Non Missing	Missing	Mean	Standard Dev
Exports as a % of GDP	727	11	30.9	17.3
External Debt	738	0	280479919572	748159524348
Inflation	690	48	15.2	158.2
GDP Per Capita	736	2	1705.8	2277.5
Public Debt as a % of GDP	432	306	69.0	38.2
GDP Growth	734	4	4.8	4.2

Table 4.3: Descriptive statistics for countries that have not rescheduled debt

Variables	Non Missing	Missing	Mean	Standard Dev
Exports as a % of GDP	1579	90	36.6	20.2
External Debt	1657	12	31774504294	99296468056
Inflation	1508	161	10.9	38.9
GDP Per Capita	1622	47	2750.3	2640.4
Public Debt as a % of GDP	344	1325	51.3	31.5
GDP Growth	1617	52	4.5	5.4

The explanatory variables are compiled from the World Bank's WDI and WGI and they consist of annual data. The descriptive statistics are split according to whether a country rescheduled its debt or not as these countries have different dynamics. Countries that have not rescheduled their debt have a lower mean value for inflation, public debt as a percentage of GDP and external debt as a percentage of GDP which corresponds to the results we expect from Table 4.1.

4.2.2 Covariate selection

According to Bursac, Gauss, Williams and Hosmer (2008), choosing variables from a large set of covariates to include in the model is one of the challenges of model building. This process is necessary so as to reduce the confounding of variables as well as to reduce the dimensionality of the problem. They identified the forward, backward and stepwise methods as the most common variable selection procedures as well as univariate association filtering. Other techniques used for variable selection include the use of the AIC and correlation (Hutmacher and Kowalski, 2015)

Table 4.4 is used to assess if there is any multicollinearity in the data by means of using the VIF. A VIF equal to one indicates that there is no correlation between the n^{th} predictor variable and the remaining variables. Inflation, GDP growth and public debt as a percentage of GDP have VIF values closer to one this indicates that they have low correlation with the other predictor variables. The variables rule of law, control of corruption and government effectiveness require further investigation as they have VIF values above four. All the variables have VIF values below ten and this indicates that multicollinearity is low (O'Brien, 2007).

Table 4.4: Variance inflation factor (VIF) for all the variables

Variance Inflation Factor								
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation	Decision
ExportsPercentageOfGDP	1	0.00381	0.00133	2.88	0.0043	0.43832	2.28146	Included
ExtDebtPercentageofGDP_numeric	1	9.27E-14	3.96E-13	0.23	0.8149	0.56113	1.78212	Included
Inflation_numeric	1	-0.0006963	0.0007671	-0.91	0.3647	0.76263	1.31125	Included
GDPPerCapita	1	-2.462E-05	1.087E-05	-2.26	0.0242	0.36039	2.77479	Included
GDPGrowth	1	0.0136	0.00501	2.71	0.0071	0.88352	1.13183	Included
PoliticalStabilityNoViolence	1	0.03378	0.03494	0.97	0.3344	0.46055	2.17132	Included
RegulatoryQuality	1	0.09749	0.06573	1.48	0.139	0.28447	3.51531	Included
RuleOfLaw	1	-0.07535	0.08296	-0.91	0.3644	0.16901	5.91696	Included
ControlOfCorruption	1	0.01539	0.09021	0.17	0.8646	0.18543	5.39291	Included
GovernmentEffectiveness	1	-0.18122	0.0884	-2.05	0.0412	0.15225	6.56798	Included
PublicDebtPercentageGDP_numeric	1	0.00357	0.000775	4.61	<.0001	0.71106	1.40635	Included

Table 4.5: Correlation matrix with Pearson correlation coefficients

Pearson Correlation Coefficients											
	ExportsPercentageOfGDP	ExtDebtPercentageofGDP_numeric	Inflation_numeric	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP_numeric
ExportsPercentageOfGDP	1	-0.07476	0.01433	0.33489	0.04981	0.3524	0.33853	0.30908	0.30434	0.34144	-0.0384
ExtDebtPercentageofGDP_numeric	-0.0748	1	-0.01566	0.10899	0.04174	-0.0439	0.21811	0.13139	0.09501	0.26801	-0.0626
Inflation_numeric	0.0143	-0.01566	1	-0.0483	0.00886	-0.0888	-0.08105	-0.0805	-0.0668	-0.0622	0.10366
GDPPerCapita	0.3349	0.10899	-0.04825	1	-0.1049	0.48928	0.66971	0.6484	0.71285	0.70999	-0.004
GDPGrowth	0.0498	0.04174	0.00886	-0.1049	1	-0.1186	-0.14887	-0.1605	-0.1555	-0.1426	-0.1847
PoliticalStabilityNoViolence	0.3524	-0.04393	-0.08878	0.48928	-0.1186	1	0.64968	0.79017	0.73617	0.69561	0.03457
RegulatoryQuality	0.3385	0.21811	-0.08105	0.66971	-0.1489	0.64968	1	0.88991	0.86622	0.93162	0.04157
RuleOfLaw	0.3091	0.13139	-0.08051	0.6484	-0.1605	0.79017	0.88991	1	0.93419	0.9301	0.10685
ControlOfCorruption	0.3043	0.09501	-0.06681	0.71285	-0.1555	0.73617	0.86622	0.93419	1	0.93182	0.06007
GovernmentEffectiveness	0.3414	0.26801	-0.06219	0.70999	-0.1426	0.69561	0.93162	0.9301	0.93182	1	0.07752
PublicDebtPercentageGDP_numeric	-0.0384	-0.06262	0.10366	-0.004	-0.1847	0.03457	0.04157	0.10685	0.06007	0.07752	1

The Pearson's correlation coefficient is a measure of strength and direction of linear association between two predictor variables. It follows the following modelling assumptions:

1. Each variable should be continuous;
2. There should be no outliers as these may skew the results;
3. There should be linearity and homoscedasticity;
4. There should be two related pairs ie. each observation should have a pair of values.

It can take a range of values from -1 to +1. A value of zero indicates that there is no linear association between the variables. A value greater than zero indicates positive association while a value less than zero indicates negative association (Norman, 2010).

Table 4.5 shows that there is high positive association between the political variables, confirming the results from the VIF analysis. Government effectiveness has a correlation of approximately 0.93 with rule of law, control of corruption and regulatory quality. All variables with correlation coefficients above 0.60 are highlighted in the diagram. Some of the political variables will thus have to be excluded from the final model as this may lead to over-fitting of the model resulting in unstable model parameters.

4.2.3 Principal component analysis

PCA was used to reduce the initial set of variables by removing redundant variables and retaining variables that explain the most variation in the model. The Kaiser criterion was used to determine the number of components to be extracted. All components with eigenvalues greater than one are extracted according to the rule (Fabrigar *et al.*, 1999). A total of four factors were shown to explain 75% of variation in the model with the factor that has the largest explanatory power explaining 75% of variation in the model.

Table 4.6: Proportion of variance explained by the extracted variables

Eigenvalue	Proportion	Cumulative percentage
4.50089879	0.4092000	0.4092
1.412622408	0.3432402	0.5376
1.27830006	0.1722903	0.6538
1.10707103	0.1006000	0.7544

Table 4.7 below identifies each component while Figure 4.3 is a diagrammatic representation of the variation explained by each component. The table shows that any model which aims at modelling debt rescheduling can be reduced into four factors. PCA analysis does not however, show the effect of each factor on the debt rescheduling variable.

Table 4.7: Identification of extracted components

Factor	Significance
1	Political climate
2	Economic outlook
3	Debt
4	Income

The first component is closely associated with rule of law, control of corruption and government effectiveness. This factor is clearly associated with the political outlook of a country. The second factor is associated with inflation, GDP growth and public debt as a percentage of GDP. This second component is best described as an indication of the economic outlook of a country. The third component consists mainly of the variable external debt and it is summarised mainly as debt. Lastly, the fourth component consists mainly of exports as a percentage of GDP and GDP per capita which represent a country's income.

The scree plot is another method that can be used to determine the components to select in the model. This plot shows each of the eigenvalues of the components. The number of components is therefore selected by identifying the point at which the curve becomes horizontal. According to the scree plot between 5 and 6 components should be retained in the model. The components which are above the selected point are retained (Cattell, 1966). The disadvantage of using the Kaiser criterion as opposed to the scree plot is that it is quite arbitrary, that is if a component has an eigenvalue of 1.01 it is retained in the model while that with an eigenvalue of 0.99 is excluded from the model. This method therefore often leads to selecting either more or less number of factors than is required (Fabrigar *et al.*, 1999). The scree plot, on the other hand has the disadvantage that it is subjective as there is no clear cut rule that is used to govern the number of factors to retain in the model as well as what constitutes a substantial drop on the curve (Kaiser, 1970).

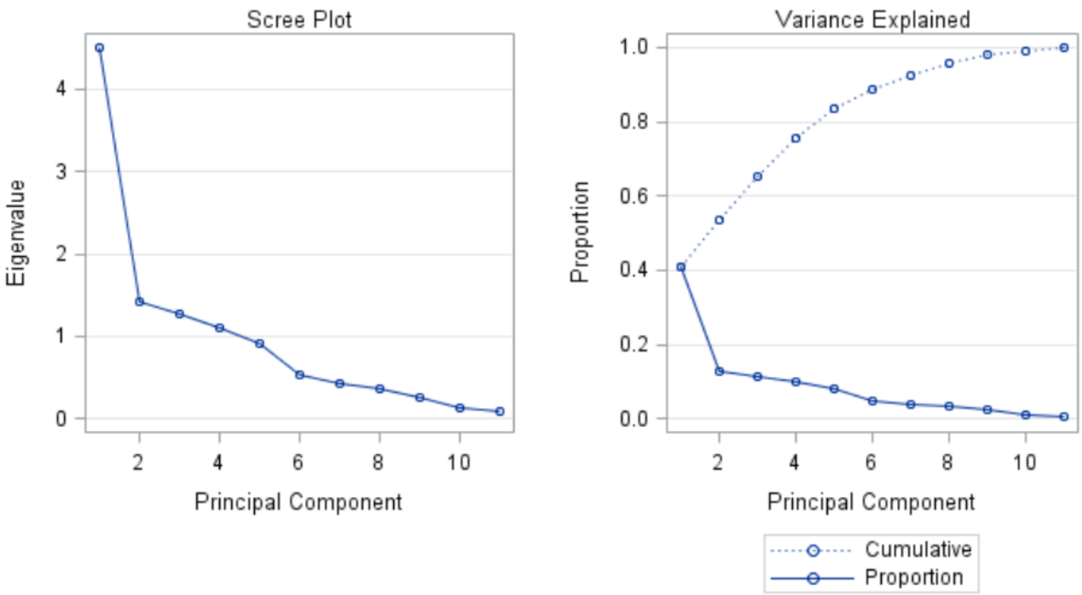


Figure 4.3: Principal component analysis scree plot

4.2.4 Logistic regression model building process

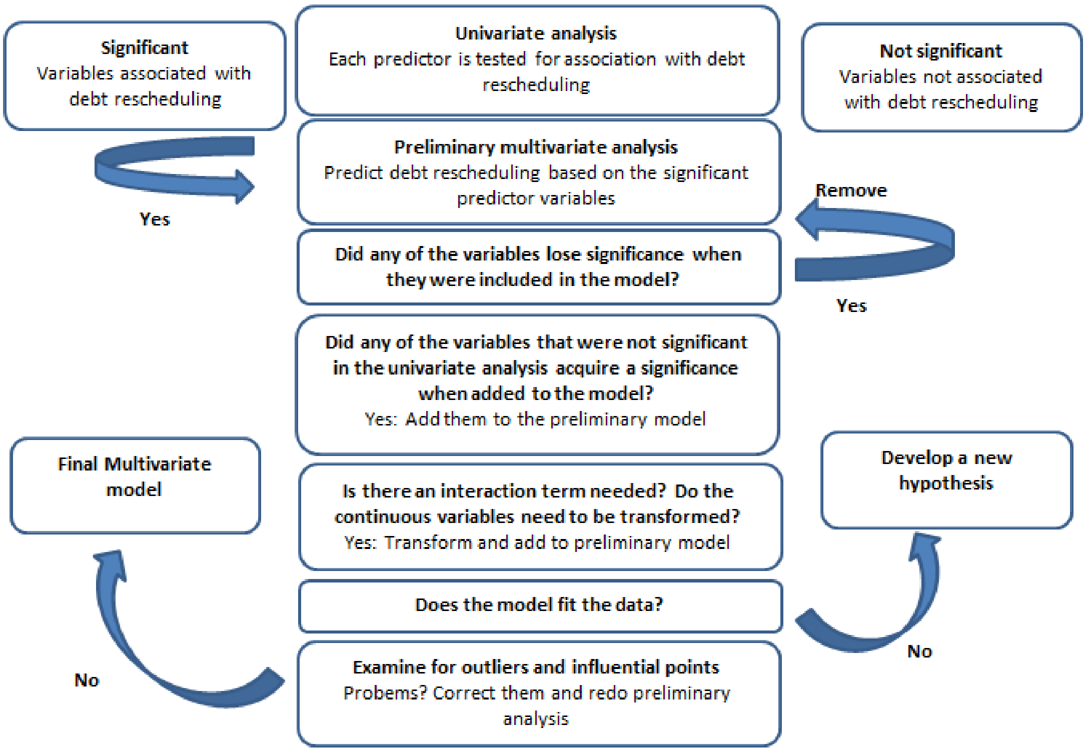


Figure 4.4: Logistic regression model development (Rosell, Olson, Aguirre-Hernandez and Carlquist, 2007)

Figure 4.4 summarises the steps involved in the development of a logistic regression model from the initial stages of univariate analysis until a final model is selected.

4.2.5 Univariate logistic regression analysis

Univariate LR was conducted to investigate the effect of each predictor on the debt rescheduling variable. The Wald test was used to test the significance of explanatory variables in the statistical model. The term ‘significant’ means that a variable adds value to the model. From the univariate LR, inflation was not significant in the model as it had a probability higher than 0.25. This may be due to the presence of an interaction between inflation and the other predictor variables. However, even though it has a probability of 0.382 that is higher than our benchmark it will be included in the multiple regression model to assess how it interacts with other variables in the backward elimination procedure. Table 4.8 illustrates the univariate LR results for each predictor variable, the Wald statistic and its probability.

Table 4.8: Univariate logistic regression results

Parameter	Df	Estimate	Wald χ^2	Pr> χ^2
Exports as a % of GDP	1	-0.0166	41.5213	<0.0001
External debt	1	299E-12	80.2724	<0.0001
Inflation	1	0.000446	0.7643	0.3820
GDP per capita	1	-0.00020	77.4199	<0.001
Political stability	1	-0.1765	8.1414	0.0043
Regulatory quality	1	-0.2474	9.3470	0.0022
Rule of law	1	-0.5076	32.7210	<0.001
Control of corruption	1	-0.4593	20.0211	<0.001
Government effectiveness	1	-0.5628	35.9037	<0.001
Public debt	1	0.0414	17.1010	<0.001

The variables that should be removed from the model taking into account VIF analysis, correlation, PCA and univariate LR are inflation, government effectiveness, control of corruption and rule of law. Inflation has a p-value above 0.25 in the univariate analysis and therefore should be removed from the model. The variables selected for the final model are shown in Chapter 5, Table 5.2 and the results from the variables selected are explained directly after the table.

4.2.6 Multivariate logistic regression

Multivariate LR using the backward elimination procedure was used to select the variables to be included in the final model. The data was divided into training, validation and test sets in the ratio's 50:30:20 (Model 1); 40:30:30 (Model 2) and 45:35:20 (Model 3) to determine the split which would best summarise the data as well as the best model obtained from the results. The split from the first model was proposed by (Cochran *et al.*, 2014). Model 2 and 3 were used for comparison purposes.

Each of the training sets were used to model the predictive relationships between debt rescheduling and the explanatory variables. The validation sets were used for model verification purposes so as to avoid over-fitting. Lastly, the test set was used to assess the models ability to generalise on new data. The AIC's of the three logit models were used as a measure of relative statistical quality and as model selection criteria. The aim of the AIC is to select the model that minimises information loss. Models that have more variables are penalized by this procedure.

Table 4.9: Summary of backward elimination for Model 1

Step	Effect Removed	DF	No. left	Wald χ^2	Pr > ChiSq
1	Exports Percentage of GDP	1	10	0.0001	0.9939
2	Poltical Stability of No Violence	1	9	0.2396	0.6245
3	Control of Corruption	1	8	0.4667	0.4945
4	Government Effectiveness	1	7	0.6070	0.4359
5	GDP Per Capita	1	6	0.6778	0.4104
6	Ext Debt	1	5	1.5733	0.3812
7	GDP Growth	1	4	2.0647	0.2097
8	Rule of Law	1	3	2.0647	0.1507

Table 4.9 shows the backward selection procedure for Model 1. All variables with p-values $> 0.05\%$ were removed in a backward manner starting with the variables with the largest p-values. The results indicate that the highly correlated political variables were part of the first variables to be removed by the backward elimination procedure. This collaborates the results obtained in the study by Erb, Harvey, and Viskanta (1996) in which they used political risk,

economic risk, financial risk, a composite risk, and a country credit rating and discovered that country risk measures were correlated to each other. They further concluded that in terms of country risk related to foreign direct investment, financial risk measures contain the most information and political risk measures contain the least information about future stock returns. The final variables selected in Model 1 are shown in Chapter 5, Table 5.2 and the results that were obtained are discussed in detail.

The respective AIC's for each step in the backward selection procedure are shown in the Figure 4.5. The AIC is used as a means of model selection. Given different models for the data, the preferred model is the one with the minimum AIC value. The AIC works by rewarding the goodness of fit. If the number of parameters in the model increases, a penalty is included to discourage over-fitting. This is because increasing the number of parameters in the model almost always improves the goodness of the fit (Akaike, 1974). The AIC for Model 1 is quite low, with a value of 34.6 as compared to Model 2 and Model 3, which will be shown below.

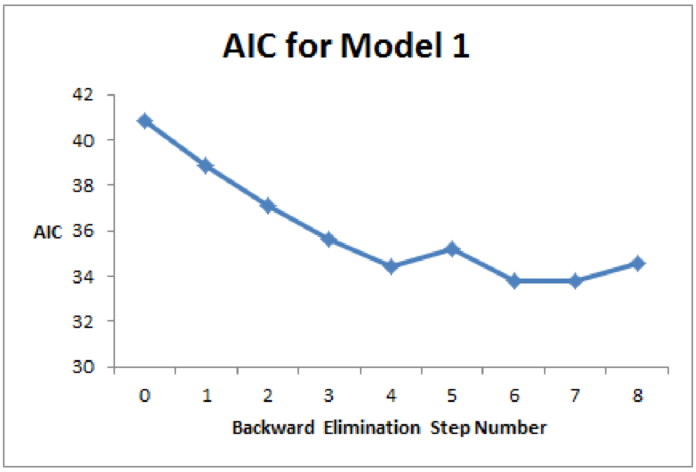


Figure 4.5: AIC for the backward elimination steps for Model 1: 50:30:20 split

Table 4.10: Summary of backward elimination for Model 2

Step	Effect Removed	DF	No. left	Wald χ^2	Pr > ChiSq
1	Poltical Stability of No Violence	1	10	0.0198	0.8881
2	Government Effectiveness	1	9	0.0328	0.8564
3	Control of Corruption	1	8	0.3151	0.5746
4	Rule of Law	1	7	0.0957	0.7570
5	Ext Debt Percentage of GDP	1	6	1.3529	0.2448
6	Regulatory Quality	1	5	1.8889	0.1693
7	Inflation	1	4	1.2988	0.2544
8	Public Debt Perc GDP	1	3	1.2989	0.2544
9	GDP Growth	1	2	0.7536	0.3854

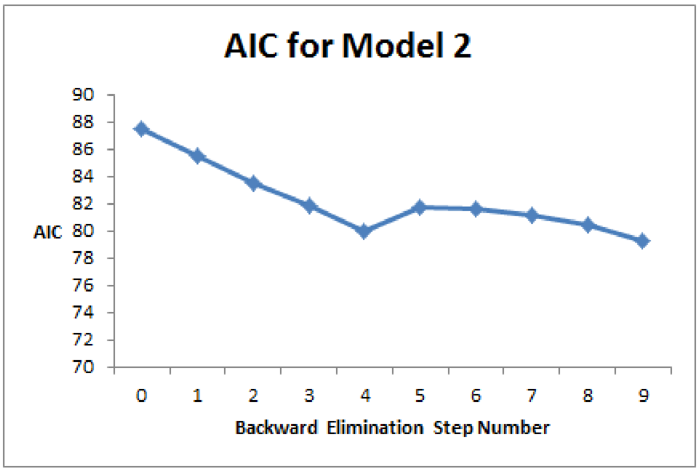


Figure 4.6: AIC for the backward elimination steps for Model 2: 40:30:30 split

A similar pattern was observed from Model 2 represented by Table 4.10, in which the political variables were the first to be removed from the model. However, the AIC obtained from Model 2, shown in Figure 4.6 was much higher than that of Model 1 with AIC values of 79.2 and 34.6 respectively. The results indicate that information loss was not minimised for Model 2 and that Model 1 is the preferred model.

The AIC values of Model 3 were well above the rest of the models but a similar trend was also observed in which the highly correlated political variables were first removed from the model. Model 1 has the best results based on the AIC. Further details on the variables that satisfied the best panel logit model are given in Chapter 5.

Table 4.11: Summary of backward elimination for Model 3

Step	Effect Removed	DF	No. left	Wald χ^2	Pr > ChiSq
1	Control of Corruption	1	10	0.1597	0.6895
2	Poltical Stability of No Violence	1	9	0.2224	0.6372
3	Government Effectiveness	1	8	0.8958	0.3439
4	Exports Percentage of GDP	1	7	1.3495	0.2454
5	Regulatory Quality	1	6	2.4710	0.1160
6	Rule of Law	1	5	1.3151	0.2515
7	GDP Growth	1	4	1.3673	0.2423
8	Ext Debt Percentage of GDP	1	3	2.5017	0.1137

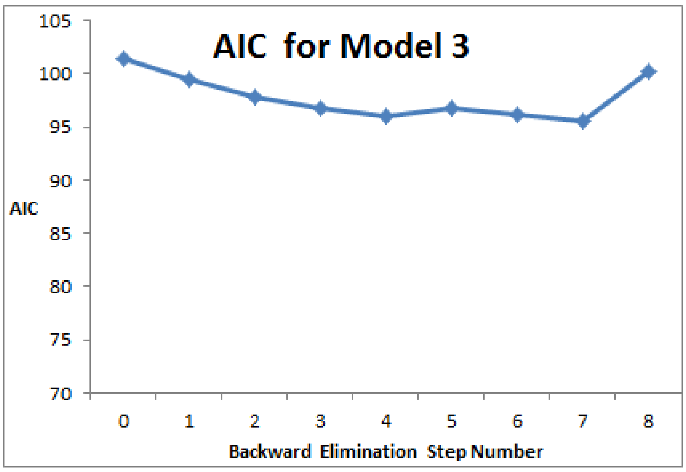


Figure 4.7: AIC for the backward elimination steps for Model 3: 45:35:20 split

4.2.7 Neural network model building process

The NN model building process is summarised by Figure 4.8. The flow diagram shows nine steps that are involved in the development of NN models. The financial, economic and political variables that were used in LR were used in developing the NN. Similar training, test and validation sets sizes were used as those used in LR. The number of hidden nodes used in the hidden layer according to Masters (1993) is:

$$nbrHID = sqr(nbrINP \times nbrOUT) = sqr(11 \times 2) = 4,699 \tag{4.1}$$

Based on the result, NN with four and five hidden nodes were developed.

The RMSE was plotted against the number of hidden nodes to see the nodes that result in the most accurate NN model and to compare if the results obtained are similar to the pyramid rule proposed by (Masters, 1993). Training and validation of the NN was done and tested on new cases. However, if a problem was identified from the validation set, the modelling process was repeated using information from the model validation step to select or fit an improved model (Turban, Sharda, Aronson and King, 2008).

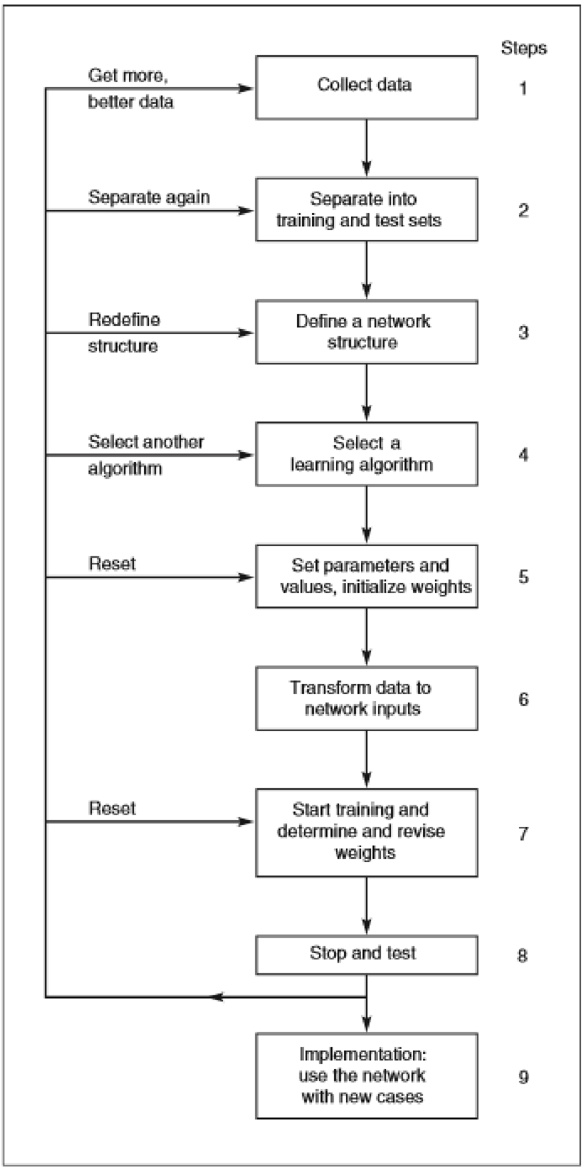
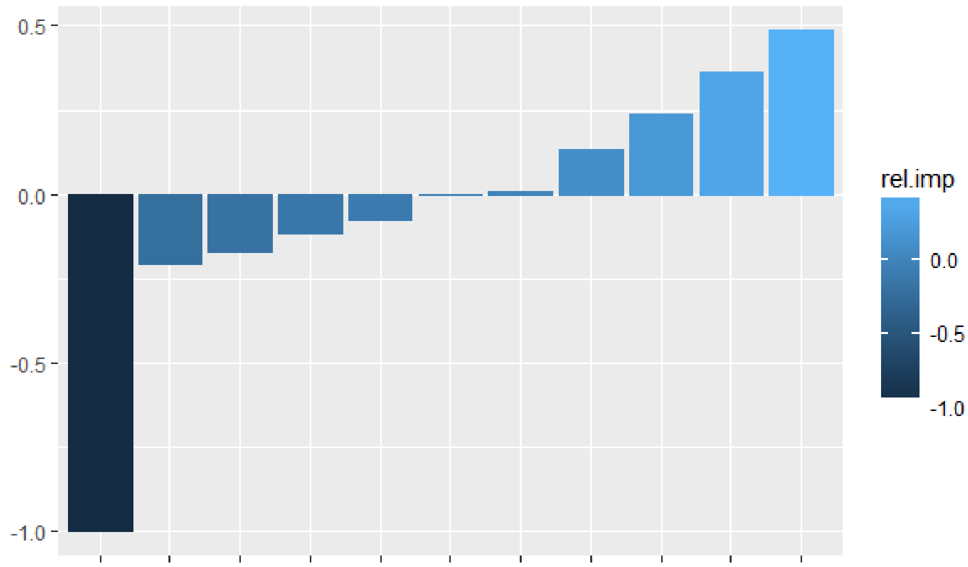


Figure 4.8: Neural network model building process (Turban *et al.*, 2008)

4.2.8 Variable selection

Variable selection is sometimes conducted prior to NN training to save computer resources (May *et al.*, 2011). The variable selection technique proposed by Garson (1991) was used to select fewer variables to use in the neural network model to assess if preselecting the variables has any effect on the neural network model accuracy. The results obtained from the variable selection are shown below:




```
           rel.imp                    x.names
   -1.000000000000                GDPPerCapita
   -0.206180005063 PoliticalStabilityNoviolence
   -0.173385041224        ExtDebtPercentageofGDP
   -0.115076477332       GovernmentEffectiveness
   -0.074179514087                     RuleOfLaw
    0.000000000000            ControlOfCorruption
    0.007767355582                     GDPGrowth
    0.133139474722                     Inflation
    0.241487657974             RegulatoryQuality
    0.363760848796          PublicDebtPercentageGDP
    0.484956404624         ExportsPercentageOfGDP
```

Figure 4.9: Relative importance of the explanatory variables used in neural network model development

GDP per capita, exports as a percentage of GDP, public debt as a percentage of GDP and political stability are considered to have the highest relative importance to debt rescheduling.

4.3 Conclusion

The model development section was used to describe, test, predict and understand the factors that determine country risk, particularly the risk of rescheduling debt. Model development was conducted step by step with model fitting and validation being used to develop models that are highly accurate. The underlying assumptions of the different methods were also taken into consideration during both the univariate and multivariate LR model development process as well as the NN model fitting process. Further detail from the LR and NN results is shown in Chapter 5.

Chapter 5

Analysis and Results

5.1 Introduction

In order to understand the phenomenon of country risk based on the debt rescheduling variable, NN and LR models were developed. The purpose of fitting several panel logit and NN models was to examine the similarities and differences of these models. The final models were assessed for adequacy and accuracy, and the results were presented both graphically and through the use of tables. The final NN and LR models were compared to determine the best modelling technique. Consideration was first given to the overall test of relationship between the dependent and independent variables through an assessment of the model's statistical results. The strength of the LR models was tested and lastly, evaluation of the usefulness of the LR models and NN models and the relationship between the independent and dependent variables was analysed.

5.2 Assessing the Overall Model Fit

5.2.1 Likelihood ratio test

The likelihood ratio test is used to compare the fit of the models. The likelihood ratios of LR model LR1, LR2, and LR3 are summarised in the table below.

The LR models performed to ascertain the effects of the political, economic

Table 5.1: Likelihood ratio test

Model	Test	χ^2	DF	Pr>χ^2	R^2	Max rescaled R^2
1	Likelihood ratio	13.9	3	0.0031	0.29	0.46
2	Likelihood ratio	22.9	2	0.0036	0.17	0.31
3	Likelihood ratio	24.2	3	<0.0001	0.17	0.28

and financial factors on debt rescheduling were statistically significant with χ^2 values of 13.9, 22.9 and 24.2 respectively and each having probabilities < 0.05. The first model explained between 0.29 to 0.46, the second between 0.17 to 0.31 and the last between 0.17 to 0.28 of the variation in the data. These pseudo R^2 measures of the logit models do not give adequate information pertaining to the model accuracy. More useful measures that can be used include the accuracy results obtained from the classification tables. However, the pseudo R^2 gives an indication of the variation explained in the model.

5.2.2 Variables selected in the models

The variables in the model show the predictor variables that were statistically significant in modelling debt rescheduling as well as their contribution to the final model's statistical significance. The variables selected for each of the final models were based on the Wald test and those values with P>0.05 were removed from the model. The variables that were found to be significant for the models developed from different training, validation and test sets are shown in table 5.2.

Model 1 was found to be the most accurate model as it had the highest AUC and the lowest range of AIC values as well as the highest pseudo R^2 value. Inflation, regulatory quality and public debt as a percentage of GDP were the final parameters in the model. The estimates for each of these parameters refer to the value by which the difference in log odds for debt rescheduling=1 changes, given that all other variables in the model remain constant. For our

Table 5.2: Variables in the model

Model	Parameter	DF	Estimate	Odds ratio	Wald χ^2	Pr>χ^2
1	Intercept	1	-5.2921		7.0604	0.0079
	Inflation	1	-0.0334	0.967	4.2899	0.0383
	Regulatory quality	1	-4.8116	0.008	5.7237	0.0167
	Public debt	1	0.0407	1.042	3.9528	0.0468
2	Intercept	1	-1.2674		3.7275	0.00535
	Exports perc of GDP	1	0.00377	1.038	3.9203	0.0477
	GDP per capita	1	-0.00121	0.999	9.4435	0.0021
3	Intercept	1	-2.7460		14.673	0.0001
	Inflation	1	0.0611	1.063	7.1467	0.0075
	GDP per capita	1	-0.00044	1.00	9.0582	0.0026
	Public debt	1	0.0247	1.025	6.8243	0.0090

model we have,

$$\log\frac{\hat{p}}{1-\hat{p}} = -5.2921 - 0.0334\times\text{inflation} - 4.8116\times\text{regulatory quality} + 0.0407\times\text{public debt} \tag{5.1}$$

The intercept is a LR estimate when all the variables in the model are evaluated at zero. If a country's inflation was to increase by one unit, the log-odds for debt rescheduling is anticipated to decrease by 0.0334 given that all other predictors in the model remain constant. We expect the log odds for debt rescheduling to decrease by -4.8116 if a country's regulatory quality were to increase by one unit and lastly, the log odds are expected to increase by 0.0407 if a countries public debt was to increase by one unit, given that the other predictors do not change. The variables with negative estimates have a lower odds ratio and vice-versa. The results obtained are in line with the intuitive results that were given in Table 4.1.

5.2.3 Goodness of fit statistics

The Hosmer-Lemeshow goodness of fit test was used to test whether the observed binary response variable, debt rescheduling, conditional on the explanatory variables, was consistent with the predictions. In other words, the test was used to compare the observed and predicted events where the data was divided into ten equal subgroups. A significant test for the Hosmer-Lemeshow statistic indicates that the model is not a good fit and a non-significant test indicates a good fit. This means that a good model is depicted by a high p-value and low Hosmer-Lemeshow statistic. The results obtained from Model 1 for the Hosmer-Lemeshow test are shown below. The Hosmer-Lemeshow test

Table 5.3: Hosmer-Lemeshow partition for Model 1

		Debt=1		Debt=0	
Group	Total	Observed	Expected	Observed	Expected
1	16	1	0.31	15	15.69
2	16	0	0.78	16	15.22
3	16	0	1.10	16	14.90
4	16	0	1.32	16	14.68
5	16	2	1.50	14	14.50
6	16	2	1.82	14	14.18
7	16	2	2.19	14	13.81
8	16	5	3.27	11	12.73
9	16	6	4.57	10	11.43
10	20	10	11.14	10	8.86

had a χ^2=7.3151, with $Pr > \chi^2 = 0.5031$ and this indicates no lack of fit.

5.2.4 The receiver operating characteristic curve

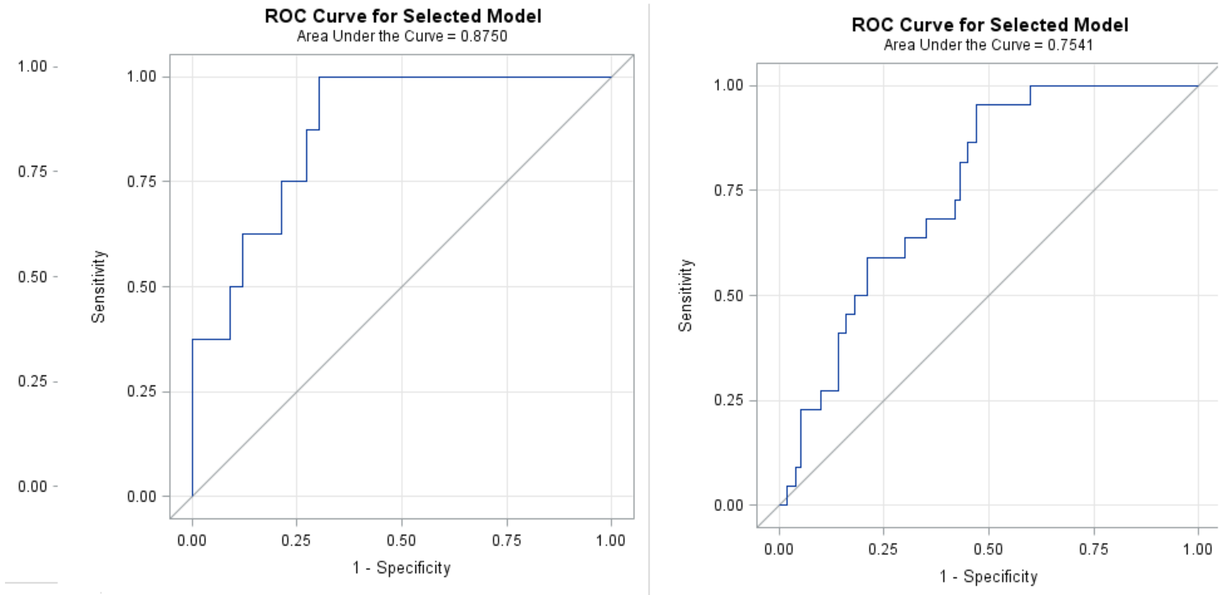


Figure 5.1: Model 1 receiver operating characteristic curve of the validation and test sets

The ROC curve allows the derivation of measures of accuracy, such as the AUC. These measures of accuracy determine the ability of the test to discriminate between countries that will reschedule their debt and those that will settle their debts. The ROC curve is a visual representation of the trade off between the true positive fraction (TPF) and the false positive fraction (FPF). The plot shows TPF(sensitivity) versus FPF (1-specificity) across varying cut-offs. The ROC curves close to the upper left hand conner have higher discriminatory power while that lying on the diagonal line reflects a diagnostic test performance which is no better than chance level.

The AUC obtained for Model 1 on the validation set was 0.875 while that for the test data was 0.75. This shows that Model 1 has a high predictive ability to discriminate between the countries that will reschedule their debt and those that will not reschedule their debt as both the validation and test sets have a high AUC.

Table 5.4: Classification table based on the validation set taking 0.5 as cut-off

Training sample	Expected		
Observed	Debt rescheduling=0	Debt rescheduling=1	Total
Debt rescheduling=0	137	23	160
Debt rescheduling=1	11	6	17

Table 5.5: Classification table based on the test set taking 0.5 as cut-off

Training sample	Expected		
Observed	Debt rescheduling=0	Debt rescheduling=1	Total
Debt rescheduling=0	31	6	37
Debt rescheduling=1	4	8	12

The classification rate is one of the most important metrics that indicates how well a model does on predicting the target variable on out of sample observations. The reliability of the prediction error rate observed in the training set is observed by applying the prediction rule to both the validation and test sets. The prediction error rate from the validation set was 0.19 while that from the test set was 0.20. This shows that there is no significant differences between the errors observed from the validation and the test set and we may conclude that these are reliable indicators of the predictive ability of the logistic regression model that has been developed.

5.2.5 Modelling neural networks using the R neuralnet package

The NN models were developed using R statistical software which is an open source programming language that allows statistical computing and graphics that are supported by its frameworks. The *neuralnet* software package was developed for use in R and was used for the training, validation and testing of the NN models. The Rprop algorithm was used to train the network (Appendix A6 to A15). The function allows flexible settings through custom-choice of error and activation function. Furthermore, the calculation of generalized weights is implemented (Riedmiller and Braun, 1993). A description of some of the arguments used in developing the NN models are given below:

Table 5.6: Arguments used in the *neuralnet* package to develop neural networks

Arguments	Description
hidden	The number of hidden neurones in each hidden layer
threshold	Numeric value which represents the stopping criteria for the partial derivatives of the error
stepmax	The maximum number of steps used for training the NN
rep	The number of training repetitions for the NN
algorithm	The algorithm used to train the network which in our case is Rprop

Table 5.7: Neural network results using Model 1 subset split

No. Hidden Nodes	Error	Threshold	Steps	AIC
1	8.01	0.0098	1574	44.01
2	3.67	0.0096	22625	61.37
3	1.52	0.0094	6236	83.03
4	2.01	0.0094	23271	110.02
5	1.41	0.0099	69196	134.83
6	0.80	0.0094	20236	159.61
7	0.49	0.0096	51260	184.98
8	0.67	0.0095	45177	211.34
9	0.12	0.0096	23813	236.23
10	0.14	0.0092	26750	262.28

The results from NN with a split in the ratio of 50:30:20 are shown in Table 5.7. The number of hidden nodes used to train the NN ranged from one to ten. The results obtained are from networks that produced the best results when tested on new data. Training, testing and validation of the data was done.

The NN model with five hidden nodes was selected as the best. This network had 69196 steps until all absolute partial derivatives of the error function were smaller than 0.01 (the default threshold). If the error function is equal to the negative log-likelihood function, the error refers to the likelihood, as is to calculate the AIC. The AIC's for each of the models was used to determine the model that minimises the information loss.

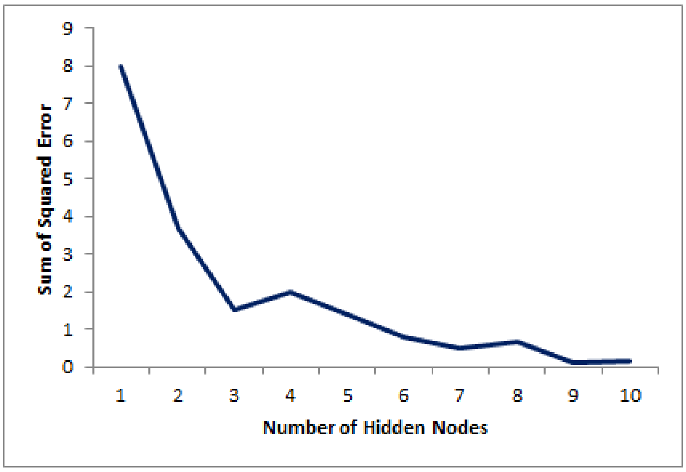


Figure 5.2: Model 1 errors obtained for each number of hidden nodes

The sum of squared error was plotted against the number of hidden nodes to determine the number of hidden nodes that minimises the error. The results from the Figure 5.2 showed better accuracy as the number of hidden nodes was increased.

The calculation from the pyramid rule from Masters (1993) indicated that between four and five hidden nodes should be used. This was confirmed by the ROC results for the NN model with five hidden nodes which had the highest AUC values. The predictive validity of the neural network model with five hidden nodes is shown using the ROC curves obtained from the validation and test sets which are shown below:

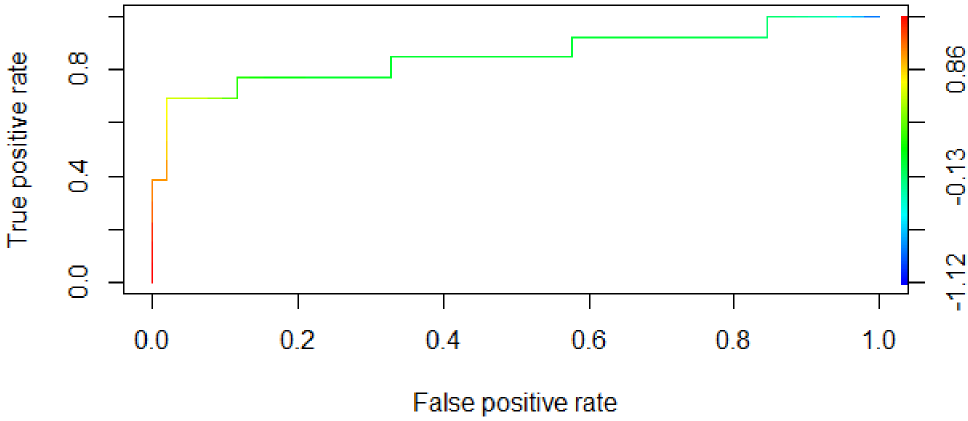


Figure 5.3: Model 1(Validation set) ROC curve with five hidden nodes and an AUC=0.850591716

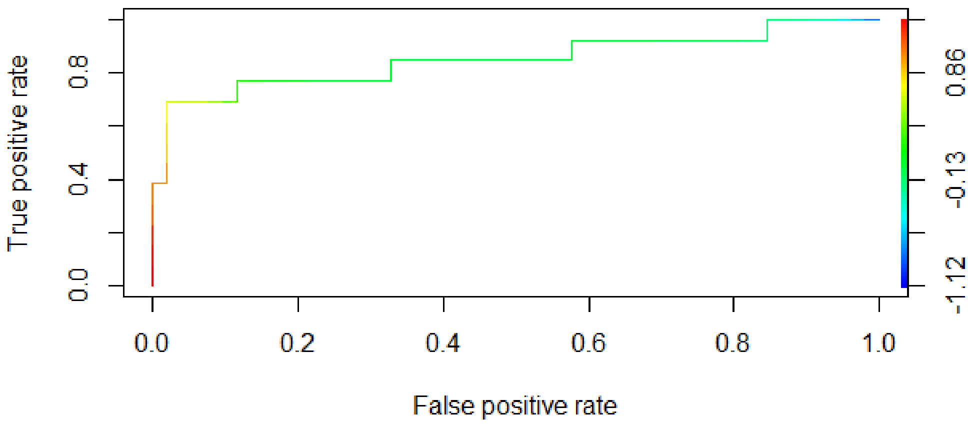


Figure 5.4: Model 1 (Test set) ROC curve with five hidden nodes and an AUC=0.850591716

The NN with five hidden nodes had an AUC of 0.85 for both the validation and the test sets. This shows that the model developed has the ability to generalise its results and as a result there was no over-fitting in the model. The network architecture for the final selected model is summarised below:

Table 5.8: Network architecture for the best neural network model.

Number of hidden layers	1
Number of hidden nodes	5
Number of input variables	11
Number of output variables	1
Activation function	logistic
Algorithm	Rprop
Number of repititions	20
Threshold	0.01

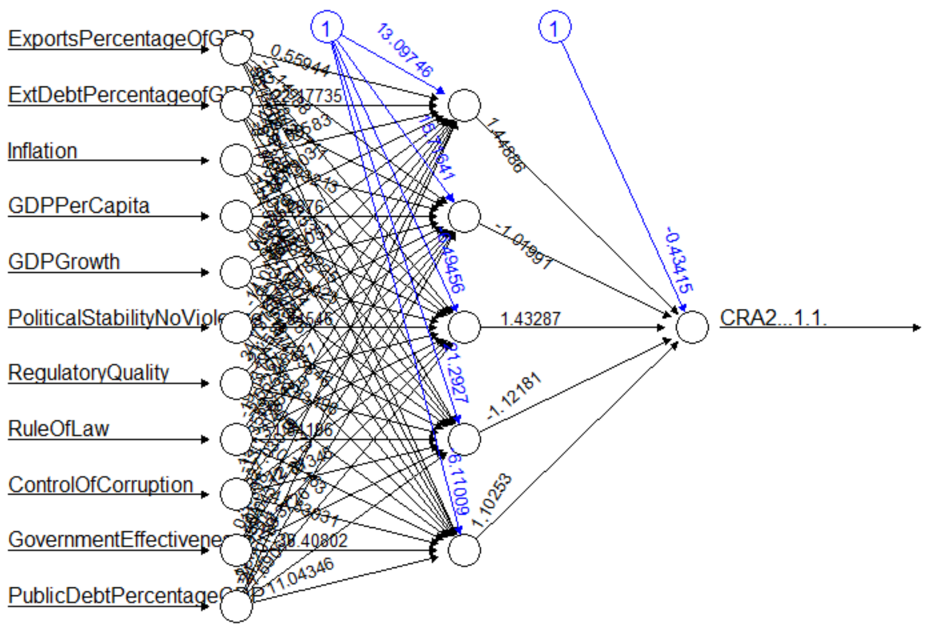


Figure 5.5: Neural network with five hidden nodes

Figure 5.5 is a diagrammatic representation of the network which had the highest predictive ability. It has eleven input nodes, one hidden layer and five hidden nodes in the layer. The output layer also consists of one hidden node. This network was obtained by normalising the variables and applying the Rprop algorithm for the training, test and validation sets. The weights associated with each of the nodes are also shown in Figure 5.5 and the respective R-code that was used is shown in Appendix A1 to A15.

5.2.6 Logistic regression vs. neural network model results

Both the final LR and NN models developed had a training, validation and test set ratio of 50:30:20. The models developed both had high predictive validity, however each of the models had different limitations in the development process.

- The logistic regression model was relatively easier to develop and this may be attributed to the fact that its easy to interpret the model parameters and its easier to use. The neural network predictor variables are difficult to interpret and hence the model is considered a black box;

- The logistic regression model took a relatively short time to develop compared to the neural networks. The neural models had high development time as there is no set method for constructing the network architecture. Furthermore, the networks required very high training times to allow the model to learn the patterns in the data. This resulted in higher computational time required for the neural networks;

- The results obtained from both the logistic regression and neural network models were impressive with AUC values of 0.75 and 0.85 on the test data. Although both AUC values were high, the neural network model outperformed the logistic regression model in terms of predictive ability;

- The results from logistic regression showed that inflation, regulatory quality and public debt were the best predictors of the risk of debt rescheduling. The neural network results can not be broken down however, we used Garson's method separately to establish that GDP per capita, exports as a percentage of GDP and public debt as a percentage of GDP and political stability were important variables to consider in the model;

- The disadvantages of the LR model was that model assumptions had to be met while the neural networks were more flexible as there were no assumptions that needed to be met.

5.3 Conclusion

We measured and compared the discriminative ability of LR and NN in classifying countries that will reschedule their debt and those that will not reschedule their debt with the use of the ROC curves. The AUC for each ROC curve was an indication of how well each discriminates against risky and less risky countries in terms of debt rescheduling. The ROC curves for the best performing LR and NN models were plotted. The results show that the NN model has better discriminatory ability as compared to the LR model as it has a higher AUC value. However, LR allowed us to determine the most predictive explanatory variables associated with debt rescheduling. The model development process showed some of the advantages and disadvantages of the different model building methods.

Chapter 6

Conclusions and Discussions

1. In our study, we reviewed and assessed the application of NN and LR in modelling the risk of debt rescheduling. The study showed how statistical and machine learning models may be used by financial institutions, policy makers and investors to better understand the debt rescheduling factors. The artificial NN model demonstrated a higher discriminatory power as compared to LR, yielding a higher AUC as compared to the latter model.

2. It is difficult to draw general conclusions from this study and prior studies as to which model is superior since the results from each study are based on specific data sets used. The selection of a superior model should therefore be based on the advantages of a particular method as well as its intended purpose of study. NN are prone to over-fitting when it comes to generalizability and their discriminatory ability is dependent on the application data set. Furthermore, it is difficult to interpret the results and identify the important predictors. LR on the other hand requires more statistical knowledge and fails to detect complex relationships between the predictor and explanatory variables unless it is identified by the modeller (Ayer *et al.*, 2010).

3. The variables which were concluded to be the best predictors of debt rescheduling from the LR models were inflation, regulatory quality and public debt which are all indications of a country's financial, political and

economic outlook. The results in this study are in line with the study by Canuto *et al.* (2012) who identified countries that were classified as being less risky in their research as having a high income per capita, low public debt to GDP ratio, low inflation and low external debt. Our work had some limitations in terms of choosing the explanatory variables as most studies use between 10 to 25 explanatory variables, however, we overcame this by selecting the variables that were seen as most significant in the country debt rescheduling literature to select our predictor variables.

4. Despite what the term suggests at first glance, country risk is not only limited to the governmental or sovereign interference and relationship with business operations. Country risk also refers to the effects of an unstable business environment, particularly other sources of risk that may hinder the optimum operation of any foreign organisation abroad. Put in other words, country risk emanates from both the public and private business sectors (Sviderske, 2014).

5. There is not any comprehensive theory of country risk. As such, the list factors that determine country risk is not exhaustive and none of the predictor variables used in the different studies weighs more than the other. This allows flexibility in country risk analysis in the sense that each case can be decided according to its unique traits as opposed to a mechanical application of variables that are predetermined and cast in stone. Furthermore, the lack of a 'closed list' opens up to the possibility of the emergence of new sources of risk and instability and such is crucial considering the rapidly changing environment in which business is being conducted. It also allows for the modelling of different financial, economic and political situations applicable in different countries and helps to prevent debt rescheduling. However, a detailed and exhaustive classification of this concept enables in-depth and extensive discussions and evaluation of the varying and distinct sources of risk, and by so doing developing a better understanding of country risk. Researchers that are

analysing country risk are not adapting the concept to the increasingly phenomenon of globalisation. Globalisation has become a reality and as such, there is need to align the concept of country risk analysis to globalisation (Sviderske, 2014).

6. An analysis of the scientific literature compels one to arrive to the conclusion that country risk appears to be very unsystematic and equally unpredictable in nature taking into consideration the 1960's and 1970's political crises and debt crises as well as the recent Greek debt crisis and Brexit (Bouchet *et al.*, 2003). As such, the outcome of a country risk analysis are highly instructive, but not always conclusive.

6.1 Concluding Remarks

The research used the variables applied in prior studies to assess the important predictor variables of country risk. A NN model and a LR model were built using economic, financial and political risk variables obtained from the World Bank. The results showed that inflation, public debt and regulatory quality are the best predictors of country risk and that a NN model has a higher discriminatory ability when applied on this data sample.

The recommendations for further study are that the countries assessed should be broken down by their different continents to compare the model performances across different economic climates. Furthermore, inclusion of more relevant variables that are more specific for example change of government or binary variables stating whether there was a war or not in the country, as well as indicators to more recent crises may be necessary in a bid to adapt to globalisation.

References

ABC. (2015). *Fact file: The Greek debt crisis explained*. ABC News (Australian Broadcasting Corporation). [online] Available from: http://www.abc.net.au/news/2015-07-20/fact-file-the-greek-debt-crisis/6628928, [11 October 2015].

Agresti, A. (1996). *Categorical data analysis*. New York: John Wiley and Sons, Vol. 990, 721 pages.

Aho, K., Derryberry, D. and Peterson, T. (2014). Model selection for ecologists: The worldviews of AIC and BIC. *Ecology*, Vol. 95, pp. 631-636.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Vol. 19, No. 6, pp. 716–723.

Akobeng, A.K. (2007). Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta paediatrica*, Vol. 96, No. 3, pp. 338-341.

Analytics Vidhya (2015). *7 important Model Evaluation Metrics Everyone should know.* [online] Available from: https://www.analyticsvidhiya.com/wp-content/uploads/2015/01/Confusion_matrix.png, [16 January 2017].

Angelini, E., di Tollo, G. and Roli, A. (2008). A neural network approach for credit risk evaluation. *The quarterly review of economics and finance*, Vol. 48, No. 4, pp. 733-755.

Aguiar, S., Aguiar-Conraria, L. and Gulamhussen, M.A. (2006). Foreign Direct Investment in Brazil and Home Country Risk. NIPE WP 7/2006, pp. 1-27.

Asiri, B.K. and Hubail, R.A. (2014). An Empirical Analysis of Country Risk Ratings. *Journal of Business Studies Quarterly*, Vol. 5, No. 4, p. 52.

Avramovic, D. (1968). *Economic Growth and External Debt.* John Hopkins Press: Baltimore, MD, 207 pages.

Ayer, T., Chhatwal, J., Alagoz, O., Kahn Jr, C.E., Woods, R.W. and Burnside, E.S., (2010). Comparison of logistic regression and artificial neural network models in breast cancer estimation 1. *Radiographics*, Vol. 30, No. 1, pp. 13-22.

Balkan, E.M. (1992). Political instability, country risk and probability of default. *Applied Economics*, Vol. 24, No. 9, pp. 999-1008.

Baser, O., Crown, W.H. and Pollicino, C. (2006). Guidelines for selecting among different types of bootstraps. *Current Medical Research and OpinionR*, Vol. 22, No. 4, pp. 799-808.

Basu, S., Deepthi, D. and Reddy, J. (2011). Country Risk Analysis in Emerging Markets: The Indian Example. *IIM Bangalore Research Paper*, 326 pages.

Becerra-Fernandez, I., Zanakis, S.H. and Walczak, S. (2002). Knowledge discovery techniques for predicting country investment risk. *Computers and Industrial Engineering*, Vol. 43, No. 4, pp. 787-800.

Behrens, J.T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods*, Vol. 2, No. 2, p. 131.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley Series in Probability and Statistics. New York: Wiley Interscience, 267 pages.

Bennell, J.A., Crabbe, D., Thomas, S. and Ap Gwilym, O. (2006). Modelling sovereign credit ratings: Neural networks versus ordered probit. *Expert Systems with Applications*, Vol. 30, No. 3, pp. 415-425.

Bouchet, M.H., Clark, E. and Groslambert, B. (2003). *Country risk assessment: A guide to global investment strategy.* John Wiley and Sons, 265 pages.

Bulow, J. and Rogoff, K. (1989). Sovereign Debt: Is to forgive to forget? *The American Economic Review*, Vol. 79, No. 1, pp. 43-50.

Bursac, Z., Gauss, C.H., Williams, D.K. and Hosmer D.W. (2008). Purposeful selection of variables in logistic regression. *Source code for biology and medicine*, Vol. 3, No. 1, p. 17.

Calverley, J. (1985). *Country risk Analysis.* London, Butterworths, 181 pages.

Camm, J., Cochran, J., Fry, M., Ohlmann, J. and Anderson, D. (2014). *Essentials of Business Analytics.* Nelson Education, 865 pages.

Cantor, R. and Packer, F. (1996). Determinants and impact of sovereign credit ratings. *Economic policy review*, Vol. 2, No. 2, pp. 37-53.

Canuto, O., Dos Santos, P.F.P. and de Sa Porto, P.C. (2012). Macroeconomics and sovereign risk ratings. *Journal of International Commerce*, Economics and Policy, Vol. 3, No. 2, pp. 1-25.

Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioural Research*, Vol. 12, No. 3, pp. 245-276.

Christl, J. and Spanel, T. (2001). 7 Country risk analysis in the light of emerging market crises. *Adapting to Financial Globalisation*, pp. 14-83.

Citron, J.T. and Nickelsburg, G. (1987). Country risk and political instability. *Journal of Development Economics*, Vol. 25, No. 2, pp. 385-392.

Cieslak, D.A. and Chawla, N.V. (2007). Detecting fractures in classifier performance. *In Data Mining , 2007. ICDM 2007. Seventh IEEE International Conference on*, pp. 123-132. IEEE.

Cooper, J.C.B. (2000). Artificial neural networks versus multivariate statistics: An application from economics. *Journal of Applied Statistics*, Vol. 26, No. 8, pp. 1-15.

Cosset, J.C., Siskos, Y. and Zopounidis, C. (1992). Evaluating country risk: A decision support approach. *Global Finance Journal*, Vol. 3, No. 1, pp. 79-95.

Costea, A. (2012). Applying Fuzzy C-means and Artificial Neural Networks for Analysing the Non-Banking Financial Institutions' Sector in Romania. *Journal of Applied Quantitative Methods*, Vol. 7, No. 3, pp. 26-32.

DiCiccio T.J. and Efron, B. (1996). Bootstrap confidence intervals (with Discussion). *Statistical Science*, No. 11, pp. 189-228.

Dreiseitl, S. and Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodological review. *Journal of Biomedical Informatics*, Vol. 35, pp. 352-359.

Eaton, J. and Gersovitz, M. (1980). LDC Participation in International Markets: Debt and Reserves. *Journal of Development Economics*, Vol. 7, pp. 3-21.

Edwards, D. (2012). *Introduction to Graphic Modelling.* Springer Science and Business Media, 335 pages.

Erb, C.B., Harvey, C.R. and Viskanta, T.E.(1996). Political Risk, Economic Risk, and Financial Risk. *Financial Analysts Journal*, pp. 29-46.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. and Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, Vol. 4, No. 3, pp. 272-299.

Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognition letters*, Vol. 27, No. 8, pp. 861-871.

Feder, G. and Just, R. E. (1977). A study of debt servicing capacity applying logit analysis. *Journal of Development Economics*, Vol. 4, No.1, pp. 25-38.

Fensterstock, F. (2005). Credit Scoring the Next Step. *Business Credit*, Vol. 107, No. 3, pp. 46-49. New York: National Association of Credit Management.

Garson, G.D. (1991). Interpreting neural network connection weights. *Artificial Intelligence Expert*, Vol.6, No.4, pp.46–51.

Ghose, T.K. (1988). How to analyse country risk. *Asian Finance*, October, pp. 61-63.

Gouvea, M. and Goncalves, E.B. (2007). *Credit risk analysis applying logistic regression, neural networks and genetic algorithms models.* Paper presented at the Production and Operations Management Society (POMS), Dallas, Texas, U.S.A.

Greene, W.H. (1993): *Econometric Analysis.* Englewood Cliffs, NJ: Prentice Hall, 801 pages.

Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques (2nd ed.).* San Franscisco: Elsevier, 770 pages.

Heinrichs, M. and Stanoeva, I. (2012). Country risk and sovereign risk building clearer borders. S and P Capital IQ, pp. 1-10.

Henisz, W.J. and Zelner, B.A. (2010). The Hidden Risks in Emerging markets. *Harvard Business Review*, Vol. 88, No. 4, pp. 2-8.

Hosmer Jr, D.W., Lemeshow, S. and Sturdivant, R. X. (2000). Model-building strategies and methods for logistic regression. *Applied Logistic Regression*, Third Edition, pp. 89-151.

Hoti, S. and McAleer, M. (2002). Country risk ratings: an international comparison. *Semminars of Department of Economics of University of Western Australia.*

Hunter, D., Yu, H., Pukish, M. S., Kolbusz, J. and Wilamowski B. M. (2012). Selection of proper neural network sizes and architectures: a comparative study. *IEEE Transactions on Industrial Informatics*, Vol. 8, No. 2, pp. 228-240.

Hutmacher, M.M and Kowalski, K.G. (2015). Covariate selection in pharmacometric analyses: a review of methods. *British journal of clinical pharmacology*, Vol.79, No. 1, pp. 132-147.

Jaimes, F., Farbiarz, J., Alvarez. and Martinez, D. (2005). Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. *Critical care*, Vol. 9, No. 2, pp. 150-156.

Jones, E. (2004). An Introduction to Neural networks. Ph.D, Visual Numerics, Inc, 39 pages.

Jha, G.K. (2007). Artificial neural networks and its application. *IARI, New Dehli, girish_iasri@rediffmail.com*, pp. 1-10.

Kaastra, I. and Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, Vol. 10, No. 3, pp. 215-236.

Kaiser, H.F. (1970). A second generation little jiffy. *Psychometrika*, Vol. 35, pp. 401-415.

Kosmidou, K., Doumpos, M. and Zopounidis, C. (Eds.). (2008). *Country Risk Evaluation*. Springer, 127 pages.

Lausev, J., Stojanovic, A. and Todorovic, N. (2011). Determinants of debt rescheduling in Eastern European countries. *Economic Annals*, Vol. 56, No. 188, pp. 7-31.

Leon-Soriano, R. and Mun˜oz-Torres, M. J. (2012). Using neural networks to model sovereign credit ratings: Application to the European Union. *In Modeling and Simulation in Engineering, Economics and Management*, pp. 13-23. Springer Berlin Heidelberg.

Liao, D. and Valliant, R. (2012). Variance inflation factors in the analysis of complex survey data. *Survey Methodology*, Vol. 38, No. 1, pp. 53-62.

Masters, T. (1993). *Practical neural network recipes in C++*. Academic Press, New York, 491 pages.

May, R., Dandy, G. and Maier, H. (2011). Review of Input Variable Selection Methods for Artificial Neural Networks. *Artificial Neural Networks- Methodological Advances and Biomedical Applications*, Prof. Kenji Suzuki (Ed.), InTech, DOI: 10.5772/16004. [online] Available from: `http://www.intechopen.com/books/artificial-neural-networks-methodological-advances-and-biomedical/-applications-review-of-input-variable-selection-methods/-for-artificial-neural-networks`, [4 February 2017].

McAleer, M., da Veiga, B. and Hoti, S. (2011). Value-at-Risk for country risk ratings. *Mathematics and Computers in Simulation*, Vol. 81, No. 7, pp. 1454-1463.

Meldrum, D. (2000). Country risk and foreign direct investment. *Business Economics*, Vol. 35, No. 1, pp. 33-40.

Mellios, C. and Paget-Blanc, E. (2006). Which factors determine sovereign credit ratings? *The European Journal of Finance*, Vol. 12, No. 4, pp. 361-377.

Nagy, P. J. (1988). *Country risk: How to Assess, Quantify, and Monitor it.* Euromoney Publications, London.

Nath, H.K. (2009). Country risk analysis: A survey of the quantitative methods. *Economia Internazionale/International Economics*, Vol. 62, No. 1, pp. 69-94.

Norman, G. (2010). Likert scales, levels of measurement and the laws of statistics. *Advances in health sciences education*, Vol. 15, No. 5, pp. 625-632.

O'Brien, R. (2007). A caution regarding the rules of thumb for variance inflation factors. *Quality and Quantity*, Vol. 41, No 5, pp. 673-690.

Paliwal, M. and Kumar, U.A. (2009). Neural networks and statistical techniques: A review of applications. *Expert systems with applications*, Vol. 36, No. 1, pp. 2-17.

Park, H. (2013). An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*, Vol. 43, No. 2, pp. 154-164.

Penn State Eberly College of Science. (2016). *STAT 504, Lesson 12: Advanced Topics 1 Generalized Estimating Equations (GEE)*. [online] Available from: `https://onlinecourses.science.psu.edu/stat504/node/180`, [11 November 2016].

Reitermanova, Z. (2010) Data Splitting. *InWDS*, Vol. 10, pp. 31-36.

Riedmiller M. and Braun H. (1993). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks (ICNN)*, pp. 586-591. San Francisco.

Robock, S. H. (1971). Political Risk: Identification and Assessment. *Columbia Journal of World Business*, Vol. 6, No. 4, pp. 6-20.

Rojas, R. (1996). *Neural Networks: A Systematic Introduction.* Springer-Verlag Berlin, 509 pages.

Rosell, J.A., Olson, M.E., Aguirre-Hernandez, R. and Carlquist, S. (2007). Logistic regression in comparative wood anatomy: Tracheid types, wood anatomical terminology and new inferences from Carlquist and Hoekman southern Californian data set. *Botanical Journal of the Linnean Society*, Vol. 154, No. 3, pp. 331-351.

Saini, K.G. and Bates, P.S. (1984). A survey of the quantitative approaches to country risk analysis. *Journal of Banking and Finance*, Vol. 8, No. 2, pp. 341-356.

Sayad, S. (2015). *Model Evaluation- Classification*. [online] Available from: http://www.saedsayad.com/images/Chart_ROC.png, [11 February 2017].

Simon, J.D. (1992). Political-risk analysis for international banks and multinational enterprises. *Country risk analysis: a handbook*, pp. 118-133.

Seltman, H.J. (2015). *Experimental Design and Analysis*. Carnegie Mellon University, Pittsburgh, PA , 407 pages.

Somerville, R.A. and Taffler, R.J. (1995). Banker judgement versus formal forecasting models: The case of country risk assessment. *Journal of Banking and Finance*, Vol. 19, No. 2, pp. 281-297.

Song, J.H., Venkatesh, S.S., Conant, E.A., Arger, P.H. and Sehgal, C.M. (2005). Comparative analysis of logistic regression and artificial neural network for computer aided diagnosis of breast masses. *Academic radiology*, Vol. 12, No. 4, pp. 487-495.

Sperandei, S. (2014). Understanding Logistic regression analysis. *Biochema medica*, Vol. 24, No 1, pp. 12-18.

Stoltzfus, J.C. (2011). Logistic regression: a brief primer. *Academic Emergency Medicine*, Vol. 18, No. 10, pp. 1099-1104.

Sviderske, T. (2014). *Country Risk Assessment in Economic Security and Sustainability Context*. Doctoral dissertation, VGTU leidykla Technika.

The New York Times. (2016). *Brexit, Explained: 7 Questions About What it Means and Why It Matters*. [online] Available from: http://mobile.nytimes.com/2016/06/21/world/europe/brexit-britain-eu-explained.html, [16 November 2016].

The Telegraph. (2016). *Dow Jones ends up 1.4pc up after markets endure rocky day of trading following Trump triumph*. [online] Available from: https://www.google.co.za/amp/www.telegraph.co.uk/business/2016/11/09/peso-dollar-and-stocks-sinking-while-gold-rises-as-markets-watch/amp/,[16 November 2016].

Topping, B.H.V., Sziveri, J., Bahreinejad, A., Leite, J.P.B. and Cheng, B. (1998). Parallel processing, neural networks and genetic algorithms. *Advances in Engineering Software*, Vol. 29, No. 10, pp. 763–786.

Trenn, S. (2008). Multilayer perceptrons: approximation order and necessary number of hidden units. *IEEE Transactions on Neural Networks*, Vol. 19, No. 5, pp. 836-844.

West, D. (2000). Neural network credit scoring models. *Computers and Operations Research*, Vol. 27, pp. 1131-1152.

Tu, J.V. (1996). Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes. *J Clin Epidemid*, Vol. 49, No. 11, pp. 1225-1231.

Tucker, J. (1996). Neural networks versus logistic regression in financial modelling: A methodological comparison. *in Proceedings of the 1996 World First Online Workshop on Soft Computing (WSC1).*

Turban, E., Sharda, R., Aronson, J.E. and King, D. (2008). *Business Intelligence: A Managerial Approach.* Pearson Prentice Hall, 233 pages.

Veall, M.R. and Zimmermann, K.F. (1996). Pseudo-R2 Measures For Some Common Limited Dependent Variable Models. *Journal of Economic surveys*, Vol. 10, No. 3, pp. 241-259.

Verisk Maplecroft. (2015). Country risk map. [online] Available from: `https://maplecroft.com/media/v_development/updatable/news/pvi-2015.jpg`, [15 January 2017].

Weesie, J. (2001). Testing for omitted variables. *In North Aemrican Stata Users' Group Meetings 2001*, No. 2.2, pp. 1-8. Stata Users Group.

World Bank. (2015). World Bank Indicators. [online] Available from `http://data.worldbank.org/indicator`, [15 December 2016].

Worrell, C., Brady, S.M. and Bala, J.W. (2012). Comparison of data classification methods for predictive ranking of banks exposed to risk of failure. *In Computational Intelligence for Financial Engineering and Economics (CIFEr)*, 2012 IEEE Conference on (pp. 1-6). IEEE.

Xu, S. and Chen, L. (2008). A novel approach for determining the optimal number of hidden layer neurones in multivalued multi-threshold neural networks. *in Proceedings of the 2nd International Symposium on Intelligent Information Technology Application*, pp. 103-107.

Yim, J. and Mitchell, H. (2005). Comparison of country risk models: hybrid neural networks, logit models, discriminant analysis and cluster techniques. *Expert Systems with Applications*, Vol. 28, No. 1, pp. 137-148.

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998). Forecasting with artificial neural networks: The state of the art. *International journal of forecasting*, Vol. 14, No. 1, pp. 35-62.

Zhang. G. (2000). Neural networks for classification: A Survey. *IEEE transactions on systems, man and cybernetics- part c: Applications and reviews*, Vol. 30, No. 4, pp. 451-462.

APPENDIX A

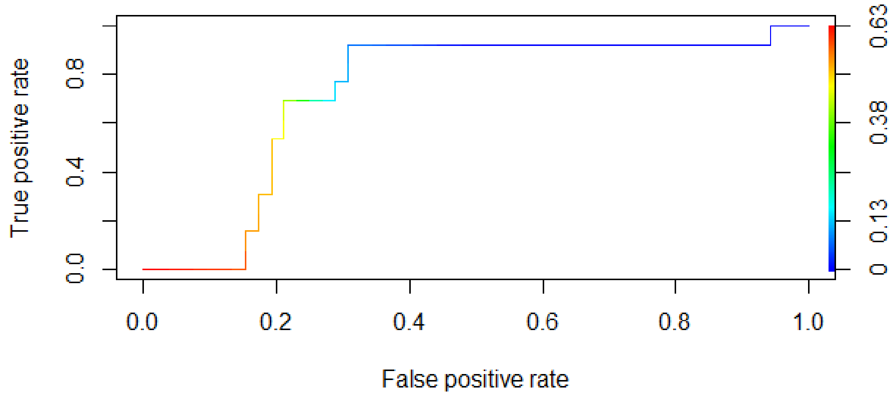


Figure 6.1: Model 1 ROC curve with one hidden node and an AUC= 0.7307692308

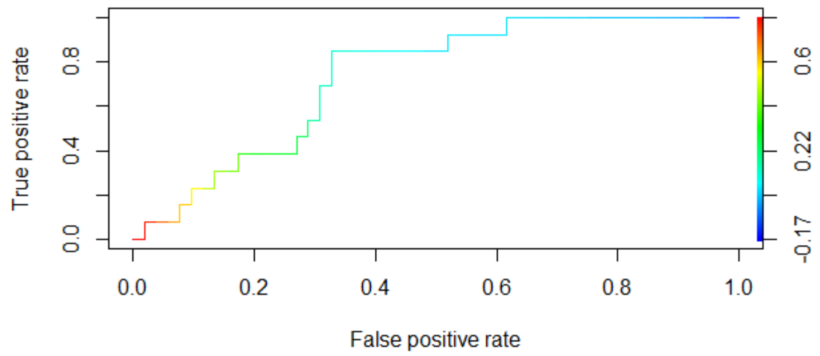


Figure 6.2: Model 1 ROC curve with two hidden nodes and an AUC=0.7337278107

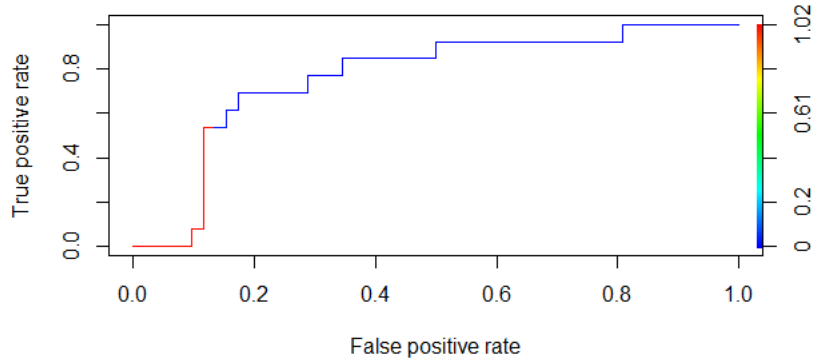


Figure 6.3: Model 1 ROC curve with three hidden nodes and an AUC=0.7647928994

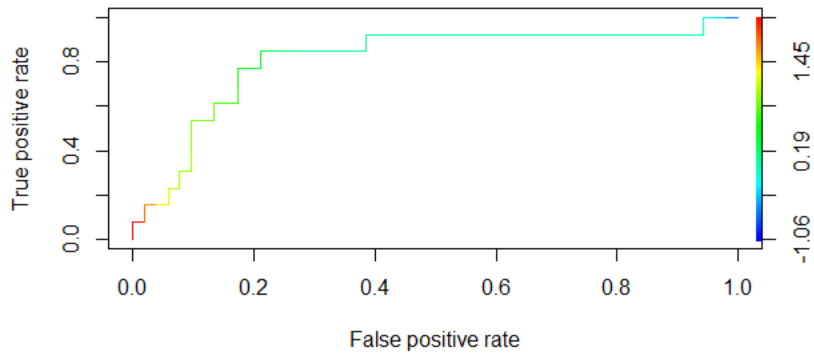


Figure 6.4: Model 1 ROC curve with four hidden nodes and an AUC=0.8106508876

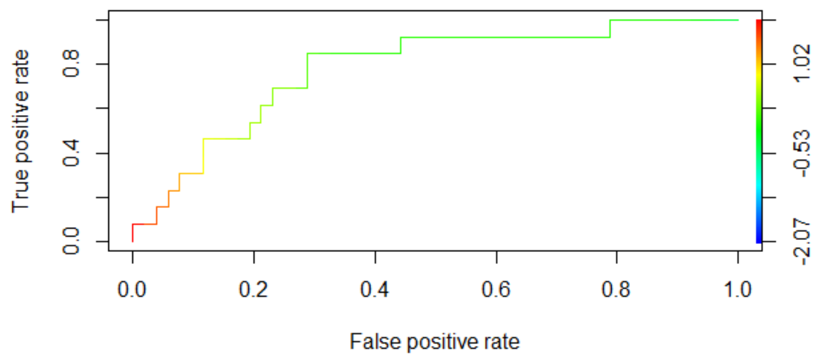


Figure 6.5: Model 1 ROC curve with six hidden nodes and an AUC=0.7810650888

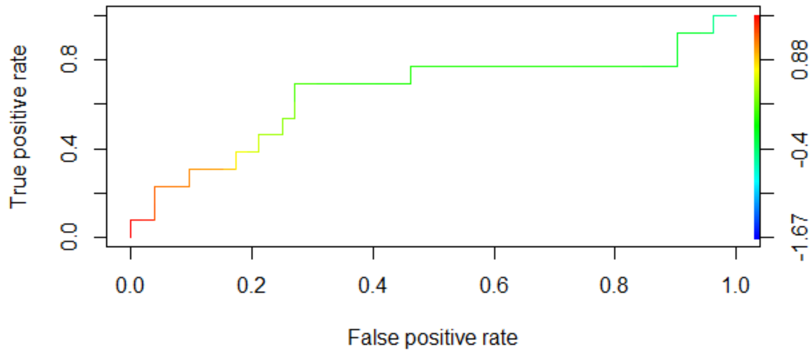


Figure 6.6: Model 1 ROC curve with seven hidden nodes and an AUC=0.6479289941

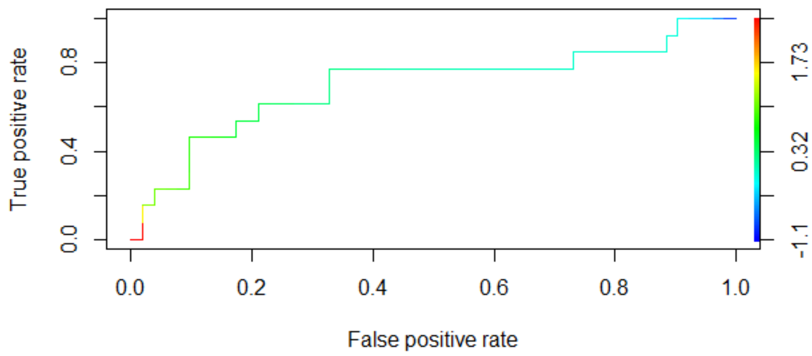


Figure 6.7: Model 1 ROC curve with eight hidden nodes and an AUC=0.6982248521

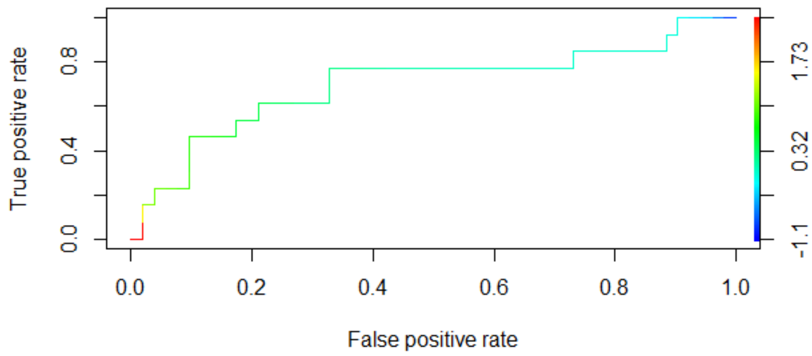


Figure 6.8: Model 1 ROC curve with nine hidden nodes and an AUC=0.6982249

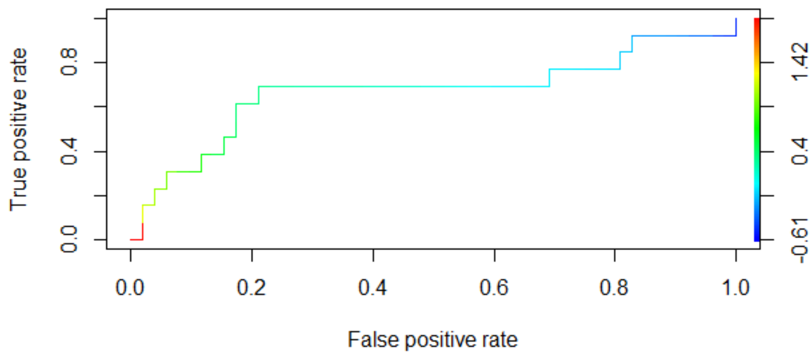


Figure 6.9: Model 1 ROC curve with ten hidden nodes and an AUC=0.6701183432

APPENDIX A1

R Code for Neural Networks

A.1 Importing Data

```
library("neuralnet", lib.loc="C:/Program Files/R/R-3.3.2/library")

library(readxl)

countryrisk <- read_excel("C:/Users/C Ncube/Desktop/Analysis and Results/countryrisk.xlsx", sheet = "Sheet1", na = "empty")

View(countryrisk)
```

A.2 Data Preparation

```
set.seed(240)

CR<-countryrisk

View(CR)

CRA <- data.frame(CR[,-1])

View(CRA)

CRA1<-data.frame(CRA[,-1])

View(CRA1)

CRA2<-na.omit(CRA1)

View(CRA2)

CRA3<-as.data.frame(CRA2[,1:1])

CRA3$debt1<-ifelse(CRA3>0,1,0)

View(CRA3)

CRA4<-data.frame(CRA3[,-1])
```

A.3 Normalising data using min max normalisation

```
maxs<-apply(CRA2[,2:12],2,max)

mins<-apply(CRA2[,2:12],2,min)
```

```
C<-as.data.frame(scale(CRA2 [,2:12],center=mins, scale=maxs-mins))

A<-cbind.data.frame(CRA4,C)

View(A)
```

A.4 Training, validation and test sets (50:30:20)

```
trainsample<-floor(0.50*nrow(A))

set.seed(123)

View(trainsample)

remsample<-sample(seq_len(nrow(A)), size=trainsample)

View(remsample)

trainsample <- A[remsample,]

View(trainsample)

valid<-A[-remsample,]

View(train)

View(valid)

Validsample<-floor(0.60*nrow(valid))

validsample<-sample(seq_len(nrow(valid)), size=Validsample)

View(validsample)

validset<-valid[validsample,]

testset<-valid[-validsample,]

variable.names(trainsample)

View(testset)
```

A.5 Variable selection using Garson’s method

```
cols<-colorRampPalette(c('lightgreen','lightblue'))

G<-gar.fun("CRA2...1.1.",CR.nn1,bar.plot=T,struct=NULL,x.lab=NULL, y.lab=NULL,
wts.only = F)

xlab(G)
```

A.6 One hidden node

```
CR.nn1 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 1, threshold = 0.01,

                stepmax = 1e+05, rep = 20, startweights = NULL,

                 learningrate.limit = NULL,

                 learningrate.factor = list(minus = 0.5, plus = 1.2),

                 learningrate=NULL, lifesign = "none",

                 lifesign.step = 2000, algorithm = "rprop+",

                 err.fct = "sse", act.fct = "logistic",

                 linear.output = FALSE, exclude = NULL,

                 constant.weights = NULL, likelihood = TRUE)

T1 <- subset(testset, select = c("ExportsPercentageOfGDP",
"ExtDebtPercentageofGDP","Inflation","GDPPerCapita","GDPGrowth","PoliticalStabilityN
oViolence","RegulatoryQuality" , "RuleOfLaw", "ControlOfCorruption"
,"GovernmentEffectiveness","PublicDebtPercentageGDP"))

a1 <- compute(CR.nn1, T1, rep=1)

results<-data.frame(actual = testset$CRA2...1.1., prediction = a1$net.result)

results$prediction <- round(results$prediction)

n1=table(results)

accuracy1=sum(n1[1,1])/sum(n1)

plot(CR.nn1, rep="best")

pred.nn1 <- prediction(results$prediction, testset$CRA2...1.1.)

perf.nn1 <- performance(pred.nn1, 'tpr', 'fpr')

plot(perf.nn1,colorize=TRUE)

auc.tmp1<- performance(pred.nn1,"auc")

auc1<-as.numeric(auc.tmp1@y.values)
```

A.7 Two hidden nodes

```
CR.nn2 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption +GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 2, threshold = 0.01,

        stepmax = 1e+05, rep = 20, startweights = NULL,

        learningrate.limit = NULL,

        learningrate.factor = list(minus = 0.5, plus = 1.2),

        learningrate=NULL, lifesign = "none",

        lifesign.step = 2000, algorithm = "rprop+",

        err.fct = "sse", act.fct = "logistic",

        linear.output = TRUE, exclude = NULL,

        constant.weights = NULL, likelihood = TRUE)

CR.nn2

a2 <- compute(CR.nn2, T1)

View(a1)

results2<-data.frame(actual = testset$CRA2...1.1., prediction = a2$net.result)

results2$prediction <- round(results2$prediction)

n2=table(results)

accuracy2=sum(n2[1,1])/sum(n2)

View(accuracy2)

pred.nn2 <- prediction(results2$prediction, testset$CRA2...1.1.)

perf.nn2 <- performance(pred.nn2, 'tpr', 'fpr')

plot(perf.nn2,colorize=TRUE)

auc.tmp2<- performance(pred.nn2,"auc")

auc2<-as.numeric(auc.tmp2@y.values)

auc2
```

A8 Three hidden nodes

```
CR.nn3 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 3, threshold = 0.01,

          stepmax = 1e+05, rep = 20, startweights = NULL,

          learningrate.limit = NULL,

          learningrate.factor = list(minus = 0.5, plus = 1.2),

          learningrate=NULL, lifesign = "none",

          lifesign.step = 2000, algorithm = "rprop+",

          err.fct = "sse", act.fct = "logistic",

          linear.output = FALSE

          , exclude = NULL,

          constant.weights = NULL, likelihood = TRUE)

CR.nn3

a3 <- compute(CR.nn3, T1)

results3<-data.frame(actual = testset$CRA2...1.1., prediction = a3$net.result)

View(results3)

results3$prediction <- round(results3$prediction)

n3=table(results)

accuracy3=sum(n3[1,1])/sum(n3)

pred.nn3 <- prediction(results3$prediction, testset$CRA2...1.1.)

perf.nn3 <- performance(pred.nn3, 'tpr', 'fpr')

plot(perf.nn3,colorize=TRUE)

auc.tmp3<- performance(pred.nn3,"auc")

auc3<-as.numeric(auc.tmp3@y.values)

auc3
```

A.9 Four hidden nodes

```
CR.nn4 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption +GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 4, threshold = 0.01,

        stepmax = 1e+05, rep = 20, startweights = NULL,

        learningrate.limit = NULL,

        learningrate.factor = list(minus = 0.5, plus = 1.2),

        learningrate=NULL, lifesign = "none",

        lifesign.step = 2000, algorithm = "rprop+",

        err.fct = "sse", act.fct = "logistic",

        linear.output = TRUE, exclude = NULL,

        constant.weights = NULL, likelihood = TRUE)

CR.nn4

a4 <- compute(CR.nn4, T1)

results4<-data.frame(actual = testset$CRA2...1.1., prediction = a4$net.result)

View(results4)

results4$prediction <- round(results4$prediction)

n4=table(results)

accuracy4=sum(n4[1,1])/sum(n4)

View(accuracy4)

pred.nn4 <- prediction(results4$prediction, testset$CRA2...1.1.)

perf.nn4 <- performance(pred.nn4, 'tpr', 'fpr')

plot(perf.nn4,colorize=TRUE)

auc.tmp4<- performance(pred.nn4,"auc")

auc4<-as.numeric(auc.tmp4@y.values)

auc4
```

A.10 Five hidden nodes

```
CR.nn5 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 5, threshold = 0.01,
        stepmax = 1e+05, rep = 20, startweights = NULL,
        learningrate.limit = NULL,
        learningrate.factor = list(minus = 0.5, plus = 1.2),
        learningrate=NULL, lifesign = "none",
        lifesign.step = 2000, algorithm = "rprop+",
        err.fct = "sse", act.fct = "logistic",
        linear.output = TRUE, exclude = NULL,
        constant.weights = NULL, likelihood = TRUE)
CR.nn5
a5 <- compute(CR.nn5, T1)
results5<-data.frame(actual = testset$CRA2...1.1., prediction = a5$net.result)
results5$prediction <- round(results5$prediction)
n5=table(results)
accuracy5=sum(n5[1,1])/sum(n5)
plot(CR.nn5, rep="best")
pred.nn5 <- prediction(results5$prediction, testset$CRA2...1.1.)
perf.nn5 <- performance(pred.nn5, 'tpr', 'fpr')
plot(perf.nn5,colorize=TRUE)
auc.tmp5<- performance(pred.nn5,"auc")
auc5<-as.numeric(auc.tmp5@y.values)
auc5
```

A.11 Six hidden nodes

```
CR.nn6 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 6, threshold = 0.01,

          stepmax = 1e+05, rep = 20, startweights = NULL,

          learningrate.limit = NULL,

          learningrate.factor = list(minus = 0.5, plus = 1.2),

          learningrate=NULL, lifesign = "none",

          lifesign.step = 2000, algorithm = "rprop+",

          err.fct = "sse", act.fct = "logistic",

          linear.output = TRUE, exclude = NULL,

          constant.weights = NULL, likelihood = TRUE)

CR.nn6

a6 <- compute(CR.nn6, T1)

results6<-data.frame(actual = testset$CRA2...1.1., prediction = a6$net.result)

View(results6)

results6$prediction <- round(results6$prediction)

n6=table(results)

accuracy6=sum(n6[1,1])/sum(n6)

View(accuracy6)

pred.nn6 <- prediction(results6$prediction, testset$CRA2...1.1.)

perf.nn6 <- performance(pred.nn6, 'tpr', 'fpr')

plot(perf.nn6,colorize=TRUE)

auc.tmp6<- performance(pred.nn6,"auc")

auc6<-as.numeric(auc.tmp6@y.values)

auc6
```

A.12 Seven hidden nodes

```
CR.nn7 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption +GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 7, threshold = 0.01,

          stepmax = 1e+05, rep = 20, startweights = NULL,

          learningrate.limit = NULL,

          learningrate.factor = list(minus = 0.5, plus = 1.2),

          learningrate=NULL, lifesign = "none",

          lifesign.step = 2000, algorithm = "rprop+",

          err.fct = "sse", act.fct = "logistic",

          linear.output = TRUE, exclude = NULL,

          constant.weights = NULL, likelihood = TRUE)

a7 <- compute(CR.nn7, T1)

results7<-data.frame(actual = testset$CRA2...1.1., prediction = a7$net.result)

pred.nn7 <- prediction(results7$prediction, testset$CRA2...1.1.)

perf.nn7 <- performance(pred.nn7, 'tpr', 'fpr')

plot(perf.nn7,colorize=TRUE)

auc.tmp7<- performance(pred.nn7,"auc")

auc7<-as.numeric(auc.tmp7@y.values)

auc7

CR.nn7
```

A.13 Eight hidden nodes

```
CR.nn8 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 8, threshold = 0.01,

          stepmax = 1e+05, rep = 20, startweights = NULL,

          learningrate.limit = NULL,

          learningrate.factor = list(minus = 0.5, plus = 1.2),

          learningrate=NULL, lifesign = "none",

          lifesign.step = 2000, algorithm = "rprop+",

          err.fct = "sse", act.fct = "logistic",

          linear.output = TRUE, exclude = NULL,

          constant.weights = NULL, likelihood = TRUE)

a8 <- compute(CR.nn8, T1)

results8<-data.frame(actual = testset$CRA2...1.1., prediction = a8$net.result)

pred.nn8 <- prediction(results8$prediction, testset$CRA2...1.1.)

perf.nn8 <- performance(pred.nn8, 'tpr', 'fpr')

plot(perf.nn8,colorize=TRUE)

auc.tmp8<- performance(pred.nn8,"auc")

auc8<-as.numeric(auc.tmp8@y.values)

auc8

CR.nn8
```

A.14 Nine hidden nodes

```
CR.nn9 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 9, threshold = 0.01,

        stepmax = 1e+05, rep = 20, startweights = NULL,

        learningrate.limit = NULL,

        learningrate.factor = list(minus = 0.5, plus = 1.2),

        learningrate=NULL, lifesign = "none",

        lifesign.step = 2000, algorithm = "rprop+",

        err.fct = "sse", act.fct = "logistic",

        linear.output = TRUE, exclude = NULL,

        constant.weights = NULL, likelihood = TRUE)

a9 <- compute(CR.nn9, T1)

results9<-data.frame(actual = testset$CRA2...1.1., prediction = a8$net.result)

pred.nn9 <- prediction(results9$prediction, testset$CRA2...1.1.)

perf.nn9 <- performance(pred.nn9, 'tpr', 'fpr')

plot(perf.nn9,colorize=TRUE)

auc.tmp9<- performance(pred.nn9,"auc")

auc9<-as.numeric(auc.tmp9@y.values)

auc9
```

A.15 Ten hidden nodes

```
CR.nn10 <- neuralnet(CRA2...1.1.~
ExportsPercentageOfGDP+ExtDebtPercentageofGDP+Inflation
+GDPPerCapita+GDPGrowth+PoliticalStabilityNoViolence+RegulatoryQuality+RuleOfLaw
+ControlOfCorruption+GovernmentEffectiveness+PublicDebtPercentageGDP,trainsample,
hidden = 10, threshold = 0.01,

            stepmax = 1e+05, rep = 20, startweights = NULL,

            learningrate.limit = NULL,

            learningrate.factor = list(minus = 0.5, plus = 1.2),

            learningrate=NULL, lifesign = "none",

            lifesign.step = 2000, algorithm = "rprop+",

            err.fct = "sse", act.fct = "logistic",

            linear.output = TRUE, exclude = NULL,

            constant.weights = NULL, likelihood = TRUE)

a10 <- compute(CR.nn10, T1)

results10<-data.frame(actual = testset$CRA2...1.1., prediction = a10$net.result)

pred.nn10 <- prediction(results10$prediction, testset$CRA2...1.1.)

perf.nn10 <- performance(pred.nn10, 'tpr', 'fpr')

plot(perf.nn10,colorize=TRUE)

auc.tmp10<- performance(pred.nn10,"auc")

auc10<-as.numeric(auc.tmp10@y.values)

auc10
```

APPENDIX B

SAS Code

B.1 Importing Data

```
/*Data Preparation: Importing Data into SAS*/

PROC IMPORT OUT= WORK.Countryrisk
        DATAFILE= "E:\countryrisk.xlsx"
        DBMS=EXCEL REPLACE;
   RANGE="Sheet1$";
   GETNAMES=YES;
   MIXED=NO;
   SCANTEXT=YES;
   USEDATE=YES;
   SCANTIME=YES;
RUN;
```

B.2 Data Preparation

```
/*Data Preparation: Making the dependent variable dichotomous*/

DATA Debtrescheduling;
  SET Countryrisk;
  Debt1=.;
IF (.<Debt<=0)THEN Debt1=0;
IF (Debt>0) THEN Debt1=1;
drop Debt;
```

```
/* Assessing the variable attributes*/

PROC CONTENTS VARNUM DATA=Debtrescheduling;
RUN;

/* Converting variables that were coded in SAS as characters to numeric*/

data Debtrescheduling1; set Debtrescheduling;
    Inflation_numeric = input(Inflation,4.);
    ExtDebtPercentageofGDP_ac = ExtDebtPercentageofGDP_numeric*1;
    ExtDebtPercentageofGDP_numeric = input(ExtDebtPercentageofGDP,$13.);
    PublicDebtPercentageGDP_ac = PublicDebtPercentageGDP_numeric*1;
    PublicDebtPercentageGDP_numeric = input(PublicDebtPercentageGDP,$6.);
    drop ExtDebtPercentageofGDP;
    drop PublicDebtPercentageGDP;
    drop Inflation;
 run;
```

B.3 Descriptive statistics

```
/*Descriptive statistics for countries that rescheduled their debt*/

proc contents data=Debtrescheduling1;
run;
proc means data=Debtrescheduling1 (where=(Debt1=1)) n nmiss min max mean std;
 var
 ExportsPercentageOfGDP
 ExtDebtPercentageofGDP_numeric
 Inflation_numeric
 GDPPerCapita
```

```
GDPGrowth
 PoliticalStabilityNoViolence
 RegulatoryQuality
 RuleOfLaw
 ControlOfCorruption
 GovernmentEffectiveness
 PublicDebtPercentageGDP_numeric;
 run;

/*Descriptive statistics for countries that did not reschedule their debt*/

proc means data=Debtrescheduling1 (where=(Debt1=0)) n nmiss min max mean std;
 var
 ExportsPercentageOfGDP
 ExtDebtPercentageofGDP_numeric
 Inflation_numeric
 GDPPerCapita
 GDPGrowth
 PoliticalStabilityNoViolence
 RegulatoryQuality
 RuleOfLaw
 ControlOfCorruption
 GovernmentEffectiveness
 PublicDebtPercentageGDP_numeric;
 run;
proc freq Data=Debtrescheduling1;
 tables Debt1;
run;
```

```
DATA Debt1;
 SET Debtrescheduling1;

 /*Histogram showing the percentage of debt rescheduling countries to non-debt
rescheduling countries*/

proc univariate data=Debtrescheduling1 noprint;
  histogram Debt1 / midpoints=0.0 0.5 hoffset=10;
title 'Frequency of debt rescheduling vs. non-debt rescheduling countries';
run;

/*Testing for Multicollinearity*/

proc corr data = Debtrescheduling1;
var
 ExportsPercentageOfGDP
 ExtDebtPercentageofGDP_numeric
 Inflation_numeric
 GDPPerCapita
 GDPGrowth
 PoliticalStabilityNoViolence
 RegulatoryQuality
 RuleOfLaw
 ControlOfCorruption
 GovernmentEffectiveness
 PublicDebtPercentageGDP_numeric;
run;

/* Calculating the variance inflation factor*/

proc reg data=Debtrescheduling1;
```

```
model Debt1=
  ExportsPercentageOfGDP
  ExtDebtPercentageofGDP_numeric
  Inflation_numeric
  GDPPerCapita
  GDPGrowth
  PoliticalStabilityNoViolence
  RegulatoryQuality
  RuleOfLaw
  ControlOfCorruption
  GovernmentEffectiveness
  PublicDebtPercentageGDP_numeric
  /tol vif;
  run;

/*Principal Component Analysis*/

proc princomp data =Debtrescheduling1;
var ExportsPercentageOfGDP
  ExtDebtPercentageofGDP_numeric
  Inflation_numeric
  GDPPerCapita
  GDPGrowth
  PoliticalStabilityNoViolence
  RegulatoryQuality
  RuleOfLaw
  ControlOfCorruption
  GovernmentEffectiveness
  PublicDebtPercentageGDP_numeric ;
run;
```

B.4 Univariate Analysis

```
/*Univariate Analysis*/

proc logist data=Debtrescheduling1 descending;
title 'Exports as a Percentage of GDP as a predictor';
model Debt1 = ExportsPercentageOfGDP/ rsquare;
run;

proc logist data=Debtrescheduling1 descending;
title 'External debt as a percentage of GDP as a predictor';
model Debt1 = ExtDebtPercentageofGDP_numeric/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'Inflation as a predictor';
model Debt1 = Inflation_numeric/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'GDP Per Capita as a predictor';
model  Debt1=  GDPPerCapita/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'GDP growth as a predictor'
model Debt1 =  GDPGrowth / rsquare;
run;
```

```
proc logist data=Debtrescheduling1  descending;
title 'Political stability no violence as a predictor';
model Debt1 = PoliticalStabilityNoViolence/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'Regulatory quality as a predictor';
model Debt1 =  RegulatoryQuality/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'Rule of Law as a predictor';
model Debt1 = RuleOfLaw/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'Control Of Corruption as a predictor';
model Debt1 = ControlOfCorruption/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'Government effectiveness as a predictor';
model Debt1 = GovernmentEffectiveness/ rsquare;
run;

proc logist data=Debtrescheduling1  descending;
title 'Public Debt as a Percentage GDP as a predictor';
model Debt1 = PublicDebtPercentageGDP_numeric/ rsquare;
run;
quit;
```

B.5 Multiple logistic regression

```
/* Splitting the data into Training (50%), Validation (30%) and Test (20%) sets*/

data split1;
retain seed 384747;
set Debtrescheduling1;
if ranuni(seed) < .5 then selected = 1;
else if 0.5<ranuni(seed)<0.8 then selected = 2;
else selected=3;
run;
 ods graphics on;
DATA Training;
Set split1(where=(selected=1));
proc logistic data=Training descending plots=all;
 model Debt1(event="1")=
  ExportsPercentageOfGDP
  ExtDebtPercentageofGDP_numeric
  Inflation_numeric
  GDPPerCapita
  GDPGrowth
  PoliticalStabilityNoViolence
  RegulatoryQuality
  RuleOfLaw
  ControlOfCorruption
  GovernmentEffectiveness
  PublicDebtPercentageGDP_numeric
  /selection=backward
```

```
lackfit
  rsquare
  ctable
  link=logit
  scale=deviance
  outroc=roc1;
  run;
  ods graphics off;
  ods graphics on;
 DATA Validation;
 Set Split1(where=(selected=2));
proc logistic data=Validation descending plots=all;
 model Debt1 (event="1")=
  ExportsPercentageOfGDP
  ExtDebtPercentageofGDP_numeric
  Inflation_numeric
  GDPPerCapita
  GDPGrowth
  PoliticalStabilityNoViolence
  RegulatoryQuality
  RuleOfLaw
  ControlOfCorruption
  GovernmentEffectiveness
  PublicDebtPercentageGDP_numeric
  /selection=backward
  lackfit
  rsquare
  ctable
  link=logit
  scale=deviance
  outroc=roc2;
```

```
run;
  ods graphics off;
 ods graphics on;
DATA Test00;
 Set Split1(where=(selected=3));
proc logistic data=Test00 descending plots=all;
 model Debt1 (event="1")=
  ExportsPercentageOfGDP
  GDPPerCapita
  RuleOfLaw
  PublicDebtPercentageGDP_numeric
  /selection=backward
   lackfit
   rsquare
   ctable
   link=logit
   scale=deviance
   outroc=roc3;
   run;
   ods graphics off;
ods graphics on;
DATA Test01;
 Set Split1(where=(selected=3));
proc logistic data=Test01 descending plots=all;
 model Debt1 (event="1")=
  Inflation_numeric
  RegulatoryQuality
  PublicDebtPercentageGDP_numeric
  /selection=backward
   lackfit
   rsquare
```

```
  ctable
  link=logit
  scale=deviance
  outroc=roc4;
  run;
  ods graphics off;

/* Splitting the data into Training (40%), Validation (30%) and Test (20%) sets*/

data split2;
retain seed 384747;
set Debtrescheduling1;
if ranuni(seed) < .4 then selected1 = 1;
else if 0.4<ranuni(seed)<0.7 then selected1 = 2;
else selected1=3;
run;
 ods graphics on;
DATA Training1;
Set split2(where=(selected1=1));
proc logistic data=Training1 descending plots=all;
 model Debt1(event="1")=
  ExportsPercentageOfGDP
  ExtDebtPercentageofGDP_numeric
  Inflation_numeric
  GDPPerCapita
  GDPGrowth
  PoliticalStabilityNoViolence
  RegulatoryQuality
  RuleOfLaw
  ControlOfCorruption
  GovernmentEffectiveness
```

```
scale=deviance
  outroc=roc6;
  run;
  ods graphics off;
 ods graphics on;

DATA Test10;
 Set Split2(where=(selected1=3));
proc logistic data=Test10 descending plots=all;
 model Debt1 (event="1")=
  ExportsPercentageOfGDP
  PoliticalStabilityNoViolence
  PublicDebtPercentageGDP_numeric
  /selection=backward
   lackfit
   rsquare
   ctable
   link=logit
   scale=deviance
   outroc=roc7;
   run;
   ods graphics off;
ods graphics on;

DATA Test11;
 Set Split2(where=(selected1=3));
proc logistic data=Test11 descending plots=all;
 model Debt1 (event="1")=
  ExportsPercentageOfGDP
  GDPPerCapita
  /selection=backward
```

```
lackfit
  rsquare
  ctable
  link=logit
  scale=deviance
  outroc=roc7;
  run;
  ods graphics off;
 ods graphics on;

/* Splitting the data into Training (45%), Validation (35%) and Test (20%) sets*/

data split3;
retain seed 384747;
set Debtrescheduling1;
if ranuni(seed) < .45 then selected2 = 1;
else if 0.45<ranuni(seed)<0.80 then selected2 = 2;
else selected2=3;
run;
 ods graphics on;
DATA Training2;
Set split3(where=(selected2=1));
proc logistic data=Training2 descending plots=all;
 model Debt1(event="1")=
  ExportsPercentageOfGDP
  ExtDebtPercentageofGDP_numeric
  Inflation_numeric
  GDPPerCapita
  GDPGrowth
  PoliticalStabilityNoViolence
  RegulatoryQuality
```

```
 RuleOfLaw
 ControlOfCorruption
 GovernmentEffectiveness
 PublicDebtPercentageGDP_numeric
 /selection=backward
  lackfit
  rsquare
  ctable
  link=logit
  scale=deviance
  outroc=roc8;
  run;
  ods graphics off;
  ods graphics on;
DATA Validation2;
Set Split3(where=(selected2=2));
proc logistic data=Validation2 descending plots=all;
model Debt1 (event="1")=
 ExportsPercentageOfGDP
 ExtDebtPercentageofGDP_numeric
 Inflation_numeric
 GDPPerCapita
 GDPGrowth
 PoliticalStabilityNoViolence
 RegulatoryQuality
 RuleOfLaw
 ControlOfCorruption
 GovernmentEffectiveness
 PublicDebtPercentageGDP_numeric
 /selection=backward
  lackfit
```

```
rsquare
  ctable
  link=logit
  scale=deviance
  outroc=roc9;
  run;
  ods graphics off;
 ods graphics on;
 DATA Test20;
 Set Split3(where=(selected2=2));
proc logistic data=Test20 descending plots=all;
 model Debt1 (event="1")=
  ExportsPercentageOfGDP
  GovernmentEffectiveness
  PublicDebtPercentageGDP_numeric
  /selection=backward
  lackfit
  rsquare
  ctable
  link=logit
  scale=deviance
  outroc=roc10;
  run;
  ods graphics off;
 ods graphics on;
 DATA Test21;
 Set Split3(where=(selected2=2));
proc logistic data=Test21 descending plots=all;
 model Debt1 (event="1")=
  Inflation_numeric
  GDPPerCapita
```

```
PublicDebtPercentageGDP_numeric
/selection=backward
 lackfit
 rsquare
 ctable
 link=logit
 scale=deviance
 outroc=roc11;
 run;
 ods graphics off;
```

Appendix C

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	12.31652	4.92E+08	12.72548	951.1321	9.1	-0.43156	-0.4192	-0.93323	-1.09098	-0.80171	39.70914
4542000	10.7614	6.24E+08	20.64286	871.8937	12.7	-0.66037	-0.18658	-1.20273	-1.00859	-0.69108	51.69209
0	65.77614	4.47E+09	20.79158	5574.604	10.77242	-0.33087	-0.34869	-0.7609	-1.04131	-0.76347	4.576538
0	51.63575	4.75E+09	1.401056	4950.295	9.410654	-0.29473	-0.2988	-0.82823	-1.11004	-0.62995	6.295012
0	54.30472	7.03E+09	5.667923	5842.806	4.854339	-0.25328	-0.36835	-0.85301	-1.17902	-0.79395	6.385576
0	12.40997	1.66E+10	3.332565	399.4275	3.833124	-1.07861	-1.01193	-0.90038	-1.17711	-0.69569	31.86399
0	11.43115	1.84E+10	5.668708	432.2197	4.739567	-1.14156	-0.92398	-1.02458	-1.33367	-0.70552	31.20445
0	46.34996	1.95E+09	52.71208	1452.447	2.8	0.004019	-1.09181	-0.72875	-0.93073	-0.40837	11.52012
0	59.05114	2.37E+09	72.86972	1511.77	8.4	0.021018	-1.85146	-0.8017	-0.63306	-0.68202	20.18054
50376000	69.21082	2.6E+09	168.6202	1273.049	5.8	0.058345	-1.7527	-1.01941	-0.50318	-0.64778	15.01445
2573000	63.62752	3.38E+09	42.53755	1479.465	5.045267	0.225317	-1.57958	-1.27187	-0.77919	-0.96043	12.33669
0	59.79781	5.28E+09	10.33888	3126.368	9.4	0.3475	-1.47526	-1.19446	-0.87231	-1.10024	6.570221
0	60.06131	6.54E+09	7.033029	3848.586	10	0.13161	-1.63851	-1.2891	-0.62791	-1.17045	6.644402
0	60.94305	1.25E+10	8.4215	4735.957	8.6	0.30698	-1.43219	-1.14304	-0.68161	-1.12612	8.909261
0	60.93776	1.51E+10	14.83788	6376.173	10.2	0.502491	-1.2522	-1.00552	-0.63776	-1.12191	10.65663
0	50.52975	2.21E+10	12.94566	5176.045	0.2	0.468292	-1.14992	-0.99775	-0.64095	-1.14927	19.18518
0	53.24372	2.84E+10	7.735748	5818.855	7.740781	-0.13286	-1.15691	-1.04114	-0.7295	-1.13527	19.59372
0	81.12894	3.39E+10	53.2287	6305.774	5.543711	-0.1211	-1.20667	-1.08315	-0.72488	-1.09521	40.75485
0	81.34082	3.38E+10	59.21974	6721.835	1.731391	0.021188	-1.09617	-0.92145	-0.52316	-0.93961	25.22436
0	50.01219	2.82E+08	6.397682	3019.524	1.428723	0.460658	0.164441	0.007658	-0.01462	0.382717	38.8391
0	51.62401	1.27E+09	-1.08002	4441.482	0.713251	0.027834	-0.4698	-0.37534	-0.04063	-0.47266	82.00438
0	58.21243	1.29E+09	5.579629	4527.337	3.323871	0.063642	-0.44597	-0.3589	-0.08175	-0.43968	80.1679
0	60.09441	1.26E+09	-3.65348	4701.548	2.103951	0.146383	-0.53591	-0.49336	-0.26026	-0.36109	76.87422
0	62.45286	1.2E+09	1.30645	4856.716	3.824749	0.180854	-0.48373	-0.4425	0.00605	-0.17679	74.47115
0	35.48813	1.13E+08	8.789659	617.2933	5.565173	0.615882	-0.44161	0.049381	0.443431	0.646414	36.29985
0	33.10001	1.71E+08	10.57702	706.2562	5.914031	0.443429	-0.4267	0.201989	0.720498	0.559306	29.96927
0	28.9844	2.12E+08	4.010994	778.1664	6.933024	0.41291	-0.3741	0.178184	0.384495	0.777606	39.12591
0	24.92009	3.86E+08	2.48343	897.3943	10.72784	0.597655	-0.46536	0.091756	0.58038	0.733384	56.36897
0	26.19336	4.94E+08	1.566152	1009.155	7.664334	0.907077	-0.0098	0.234251	0.75395	0.38277	68.78929
0	31.29266	6.01E+08	-18.1086	1108.518	5.896408	1.174361	-0.80632	0.356159	0.618482	-0.13942	75.23595
0	38.24984	6.57E+08	5.311513	1258.992	7.12256	1.304768	-0.42999	0.363544	0.747862	0.280391	79.53392
0	54.41945	7.21E+08	5.000454	1348.82	6.849366	1.30769	-0.59925	0.255707	0.658192	0.175545	74.88324
0	54.97042	8.01E+08	5.156111	1760.603	17.92582	0.62162	-0.72686	0.3715	0.740649	0.190863	62.79525
0	46.55944	6.94E+08	8.32716	1817.982	4.768354	0.751302	-0.82748	0.372823	0.765826	0.218838	60.41026
0	44.70198	7.23E+08	4.361122	1795.149	6.657224	0.815643	-1.09638	0.17795	0.810364	0.480322	56.77631
0	53.82844	6.27E+08	10.08286	2990.617	5.8298	0.923111	0.757556	0.502153	0.585983	0.465285	11.15676
0	38.52648	1.8E+09	6.948877	6492.868	8.563257	0.960708	0.456155	0.666208	1.003335	0.463565	21.63043
0	44.40475	2.4E+09	8.458166	7893.957	6.048549	1.050536	0.500329	0.665027	0.988278	0.461956	19.85778
0	42.44091	2.49E+09	7.540284	7381.764	4.831387	1.08013	0.693962	0.661444	0.930109	0.43293	18.71509

Table 1: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	43.78611	5.54E+10	2.753172	6738.1	-5.01194	0.31842	0.662095	-0.0739	-0.24712	0.161331	13.38255
0	55.14014	5.03E+10	2.438991	6580.814	0.655325	0.327432	0.641008	-0.10395	-0.20698	0.110422	14.64958
0	63.63874	4.72E+10	4.219903	7588.809	1.982003	0.279012	0.537228	-0.13671	-0.22504	0.108624	14.83066
0	64.60203	5.07E+10	2.954568	7198.518	0.49219	0.350884	0.544536	-0.12377	-0.23703	0.136272	17.52433
0	5.820321	1.13E+09	26.43678	138.0628	-8	-2.24305	-1.67297	-1.7185	-1.38879	-1.72686	144.0977
0	7.99975	1.12E+09	12.50041	138.6191	4.75	-2.42884	-1.59482	-1.48169	-1.15436	-1.65561	151.2772
4.94E+08	21.43038	1.13E+10	3.170752	637.3913	4.89536	-0.82482	-0.63709	-1.11523	-1.04016	-0.73754	97.49886
0	16.5753	3.62E+10	7.131186	2261.284	3.918272	-2.39011	-0.07836	-0.76008	-0.17041	-0.13695	52.53542
0	17.78948	4.62E+10	6.996991	5403.456	3.546805	-1.84202	0.261596	-0.43576	-0.2222	-0.0313	64.76494
0	16.03081	5.3E+10	4.202933	5104.991	1.651549	-1.83112	0.151948	-0.42732	-0.30545	-0.22894	68.57884
0	15.93598	6.38E+10	2.27822	6179.77	3.971801	-1.53231	0.258239	-0.34544	-0.40863	-0.04285	72.33779
0	18.73611	7.69E+10	3.411616	7124.549	6.589512	-1.26548	0.367974	-0.28601	-0.30012	0.063943	62.77181
0	18.25876	7.99E+10	3.176934	7748.96	4.043944	-1.39749	0.386963	-0.39212	-0.43424	0.010877	65.45031
0	30.00276	1.28E+10	492.4419	133.8514	-1.02317	-2.83376	-1.8343	-1.93181	-2.05746	-1.68694	243.5955
0	80.81237	5.55E+09	-0.63184	1039.38	0.813264	-1.19618	-1.10129	-1.18608	-0.93613	-1.31389	0.213918
790000	30.24973	3.42E+10	9.318969	1757.776	7.087827	-0.59258	-0.2789	-0.18468	-0.67432	-0.37517	85.78853
0	26.36667	5.77E+09	1.865525	2381.153	2.340823	0.222369	0.018091	-0.5607	-0.75836	-0.50881	55.2776
0	27.06773	8.1E+09	2.120391	2495.569	2.300038	-0.19768	-0.11546	-0.48626	-0.37257	-0.33617	50.97264
0	26.9586	8.57E+09	4.451944	2611.161	1.850536	-0.06868	0.099613	-0.37677	-0.39547	-0.2679	49.08503
0	25.63971	9.32E+09	4.690949	2814.935	3.562815	-0.02946	-0.04692	-0.46938	-0.41567	-0.32166	48.33682
0	25.68205	1.02E+10	4.037124	3042.748	3.911949	-0.14878	0.068174	-0.64152	-0.18615	-0.19106	43.56655
0	25.88275	9.85E+09	4.578086	3283.531	3.839765	0.00441	0.196811	-0.69019	-0.2868	-0.19924	41.02333
0	26.88162	1.07E+10	6.707923	3483.709	1.274227	0.052919	0.215693	-0.71244	-0.2979	-0.16345	40.64439
0	23.19684	1.04E+10	1.05595	3341.32	-3.13305	-0.02132	0.353225	-0.8004	-0.19571	-0.01895	49.53681
0	25.92456	1.11E+10	0.907807	3444.456	1.364784	0.055753	0.375746	-0.86744	-0.23387	0.002475	49.52074
0	27.98003	1.2E+10	5.128924	3698.546	2.216836	0.114121	0.487637	-0.75734	-0.21159	-0.1074	47.82986
0	9.176547	1.01E+10	-8.48425	145.3358	12.42617	-1.0974	-1.34148	-0.90987	-1.15256	-1.28224	72.15001
34734000	13.26575	1.04E+10	0.894802	125.3846	-3.45814	-0.72327	-1.17575	-0.78776	-0.691	-0.93778	77.96168
0	20.74702	6.67E+08	4.838622	434.4954	-0.94168	0.249529	-0.51626	-0.28051	-0.70791	-0.6676	115.3361
0	14.73657	7.18E+08	2.056503	441.9197	1.1241	-0.0155	-0.38324	-0.30326	-0.74082	-0.70607	118.4142
0	14.73448	7E+08	5.369135	522.3399	3.631026	0.064142	-0.35988	-0.24025	-0.7715	-0.59163	23.86084
7360000	16.17402	3.73E+08	4.443655	612.0286	5.734642	0.080116	-0.38467	-0.35622	-0.75296	-0.72038	69.44844
23976000	22.90781	5.06E+08	4.561582	553.1065	6.449696	0.143736	-0.32211	-0.44333	-0.5638	-0.6277	65.0552
0	16.45873	1.65E+09	3.566808	805.2727	3.104903	-1.76541	-0.47903	-1.45592	-0.78348	-0.63921	56.90445
0	22.99463	1.83E+09	4.063962	691.9977	1.838342	-0.98919	-0.39476	-1.1193	-0.8776	-0.72571	69.93811
48038000	29.22917	2.03E+09	5.587837	779.3846	5.473838	-1.36108	-0.81443	-1.16702	-1.13873	-0.87782	64.9582
0	31.83762	2.15E+09	4.785253	922.0297	11.0581	-1.35019	-0.66331	-1.00089	-0.64721	-0.49982	53.80605
89761000	31.55575	2.32E+09	5.656354	1186.873	5.85733	-0.8646	-0.45817	-0.66544	-0.6083	-0.4859	43.83622
79789000	33.74824	2.15E+09	8.247064	1469.927	9.599641	-0.6824	-0.50918	-0.72648	-0.35926	-0.42436	40.9235
67129000	32.86555	2.57E+09	9.160958	1761.12	9.383276	-0.93658	-0.1248	-0.46592	-0.04376	-0.1812	27.9581

Table 2: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	31.20572	2.9E+09	9.244919	2318.127	12.344	-0.62388	0.28353	-0.33584	-0.24913	0.108092	22.68713
0	28.61997	7.63E+09	9.9995	2918.711	2.314047	-0.90984	0.484929	-0.25951	-0.2236	0.295664	27.01777
0	29.73941	8.56E+09	1.726046	2440.961	-3.77577	-0.9429	0.51722	-0.21042	-0.22422	0.277763	34.61136
0	34.95113	9.52E+09	7.119216	2613.757	6.253029	-0.71669	0.585655	-0.2126	-0.11907	0.292477	36.79972
0	36.24242	1.14E+10	8.53495	3219.606	7.199999	-0.65516	0.651458	-0.12751	-0.02456	0.552616	32.45976
0	38.15052	1.31E+10	-0.93463	3528.732	6.182117	-0.67421	0.676574	-0.02881	0.253194	0.565481	32.5307
0	17.81012	3.39E+09	11.05692	1544.999	2.95778	-1.07674	-0.20633	-1.18811	-0.82384	-0.49982	11.24183
0	18.17161	3.69E+09	6.613461	1814.003	4.993528	-0.77185	0.094027	-1.13282	-0.84581	-0.41903	11.28017
0	20.19503	3.95E+09	5.977577	1721.729	3.608869	-0.82618	-0.11357	-0.9188	-0.67828	-0.52682	16.98382
0	26.29896	4.53E+09	8.132631	1765.851	3.866638	-0.89625	-0.16104	-1.0081	-0.49896	-0.49323	18.3936
0	25.7737	5.16E+09	5.603477	1816.95	2.530838	-0.81812	-0.34075	-1.10456	-0.66528	-0.4475	20.71104
0	26.97552	8.09E+09	7.578622	1937.716	3.152031	-0.83715	-0.18568	-1.0137	-0.53754	-0.64077	21.78594
0	25.05476	9.51E+09	9.10865	2146.179	3.260163	-0.85641	-0.39353	-1.09521	-0.63211	-0.69747	21.26247
0	24.93169	1.09E+10	6.560853	2326.308	5.379614	-0.74004	-0.20125	-1.11004	-0.74005	-0.5924	21.87669
0	25.5654	1.27E+10	6.821618	2561.441	6.30425	-0.76587	-0.19829	-1.14476	-0.68894	-0.54835	21.63558
0	24.71812	1.38E+10	11.35576	2867.499	3.281094	-0.72792	-0.15631	-1.15426	-0.61525	-0.56348	20.05391
0	23.9753	1.48E+10	1.859103	2697.383	0.526193	-0.94431	-0.11823	-1.07299	-0.47813	-0.69382	23.01721
0	25.80571	1.5E+10	3.859509	2882.386	2.869372	-0.87314	-0.13028	-0.99882	-0.4813	-0.69875	24.41982
0	26.62528	1.63E+10	6.215342	3240.376	4.161946	-0.76554	-0.112	-1.05808	-0.47008	-0.69945	24.16256
0	24.86869	1.41E+10	3.7825	3340.783	2.969849	-0.65577	-0.1767	-1.09722	-0.61526	-0.76328	24.31127
0	41.93154	2.72E+10	23.42814	4504.676	0.033921	0.913693	0.876682	0.829139	0.578486	0.83692	76.49041
0	53.00151	2.84E+10	14.17507	4728.792	4.208478	1.124261	1.014151	0.789046	0.654127	0.941032	65.45136
0	66.89463	3.08E+10	9.780586	4613.706	4.240404	0.81988	1.066013	0.852988	0.687797	0.960691	62.00981
0	58.22528	4.19E+10	5.261541	6631.448	4.491147	1.180384	1.310162	0.929167	0.524351	1.020233	59.88229
0	56.5973	5.75E+10	4.648458	8365.465	3.778347	1.105742	1.118526	0.890325	0.603236	0.961748	62.15006
0	60.02429	8.31E+10	6.77996	10206.33	4.789353	0.810015	1.176149	0.891287	0.650468	0.898782	65.57796
0	63.15296	8.62E+10	3.550808	11092.43	4.259658	0.982656	1.112534	0.825876	0.615889	0.79512	67.87026
0	74.62753	1.33E+11	3.878312	11342.89	3.962944	0.95821	1.208113	0.96134	0.608527	0.875363	69.40759
0	78.6111	1.75E+11	7.935009	13781.14	0.511279	0.723845	1.18874	0.919356	0.558783	0.719478	69.5759
0	79.95559	2.23E+11	6.066157	15598.32	0.878582	0.718135	1.191159	0.892088	0.384221	0.713956	72.83976
0	75.06223	2.40E+11	4.20919	12906.75	-6.55103	0.517169	1.08032	0.758685	0.34122	0.676862	81.09508
0	82.62343	2.16E+11	4.881345	12958.27	0.789117	0.670444	1.017244	0.747584	0.252409	0.665541	81.65163
0	87.53358	2.11E+11	3.957267	13983.5	1.806605	0.73476	1.029384	0.744916	0.319202	0.675183	81.22527
0	87.41585	2.01E+11	5.705829	12784.3	-1.47794	0.665254	0.969599	0.595087	0.279199	0.618975	84.65846
0	10.20618	9.49E+10	8.977149	410.8184	7.549522	-0.91241	-0.44069	0.259185	-0.40251	-0.08173	44.93528
0	10.82857	9.88E+10	13.23084	425.4453	6.184416	-1.14237	-0.40985	0.289573	-0.28305	-0.09057	49.45198
0	12.77323	1.01E+11	4.009434	457.2835	3.840991	-0.99123	-0.15845	0.279676	-0.36516	-0.13923	54.05627
0	14.01854	1.06E+11	4.3922	486.6405	3.803975	-1.24437	-0.37928	-0.03574	-0.49742	-0.13084	61.47469
0	14.69029	1.19E+11	3.805866	565.3355	7.860381	-1.52688	-0.36484	0.098476	-0.43556	-0.0708	61.11823
0	17.55133	1.24E+11	3.767238	649.7106	7.922937	-1.21908	-0.39915	0.038159	-0.41397	-0.10433	61.51421

Table 3: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	19.28015	1.21E+11	4.246353	740.1143	9.284832	-0.99096	-0.24292	0.159548	-0.39707	-0.08404	61.19467
0	21.06948	1.60E+11	6.145522	830.1632	9.263965	-1.05837	-0.23357	0.185133	-0.29668	-0.04564	59.10994
0	20.43089	2.04E+11	6.369997	1068.678	9.80136	-1.14687	-0.26645	0.105312	-0.42499	0.111004	56.47807
0	23.60125	2.27E+11	8.351816	1042.084	3.890957	-1.10107	-0.35728	0.086872	-0.36129	-0.02984	56.11267
0	20.04962	2.56E+11	10.87739	1147.238	8.479787	-1.32833	-0.30323	0.02349	-0.47737	-0.00521	54.30595
0	21.9703	2.92E+11	11.9923	1417.074	10.25996	-1.23315	-0.36918	-0.0414	-0.51284	0.016961	50.60013
0	24.27455	3.37E+11	8.857845	1503.341	6.638348	-1.29616	-0.33498	-0.1119	-0.57281	-0.00503	44.59655
0	24.42506	3.95E+11	9.312446	1481.201	5.081418	-1.252	-0.47244	-0.10471	-0.56553	-0.18126	50.31137
0	25.82455	1.29E+11	7.96848	1153.588	7.642786	-1.17699	0.194289	-0.366	-0.56062	-0.41701	23.90756
3.89E+09	52.96814	1.51E+11	58.38709	470.1961	-13.1267	-1.76889	-0.26419	-0.71933	-1.08211	-0.59606	53.79358
5.07E+09	32.68762	1.28E+11	11.87876	909.8873	4.499475	-1.62067	-0.64309	-0.96559	-1.13394	-0.42742	32.22413
3.09E+09	30.47766	1.34E+11	6.585719	1076.219	4.780369	-2.11807	-0.7811	-0.89168	-0.95995	-0.45118	29.7174
0	32.21669	1.37E+11	6.243521	1160.615	5.030874	-1.86905	-0.66707	-0.76884	-0.89463	-0.37552	56.60271
2.61E+09	34.06727	1.42E+11	10.45196	1273.465	5.692571	-1.48486	-0.54405	-0.82031	-0.85506	-0.44258	47.33803
0	31.03472	1.36E+11	13.10942	1601.031	5.500952	-1.40214	-0.33734	-0.72871	-0.81047	-0.33871	38.99597
0	29.43572	1.48E+11	6.407448	1871.288	6.345022	-1.20161	-0.32374	-0.67624	-0.58287	-0.28279	35.16698
0	29.80828	1.58E+11	9.776585	2178.198	6.013704	-1.08532	-0.31825	-0.65957	-0.56275	-0.23873	33.07422
0	24.15912	1.79E+11	4.813524	2272.042	4.628871	-0.75852	-0.33309	-0.59514	-0.81583	-0.27657	28.37321
0	39.74123	1.04E+10	9.293255	4805.879	1.4	-0.20259	0.314088	-0.45525	-0.48557	0.297552	113.0256
0	41.94871	1.03E+10	22.02093	5130.227	-0.71407	-0.26461	0.334917	-0.40398	-0.48628	0.293842	120.4325
3.87E+08	52.87237	7.39E+09	6.501218	1601.933	2.085493	-0.15249	0.026383	0.276333	-0.12444	0.12165	106.2211
2.11E+08	44.85409	7.56E+09	3.091667	1733.265	2.996114	-0.05112	0.402503	0.357825	-0.00994	0.079839	101.1937
1.94E+08	41.88793	1.11E+10	0.666881	1763.169	4.239817	-0.09063	0.248454	0.3708	0.040744	-0.00831	93.74273
2.13E+08	47.48418	1.3E+10	1.832994	1901.58	5.792386	-0.55758	0.029744	0.134742	-0.10398	0.123936	85.09265
5.63E+08	47.43563	1.27E+10	1.63	1973.862	4.163752	-0.11698	0.213587	0.375004	0.314324	0.236361	88.93871
2.57E+08	52.26628	1.3E+10	3.361868	2156.44	8.562737	-0.2324	0.317657	0.370598	0.339598	0.124381	89.05454
2.2E+08	52.70598	1.29E+10	3.493685	2326.495	8.16378	-0.13203	0.169825	0.408534	0.328823	0.074761	79.95192
2.08E+08	53.87449	1.4E+10	6.251725	2719.822	8.093331	-0.76934	0.343752	0.379852	0.297587	0.184097	72.62513
1.96E+08	54.23437	1.5E+10	5.386824	3022.543	8.175719	-0.31229	0.315041	0.450412	0.308095	0.221619	69.95882
0	56.50592	1.4E+10	14.92782	3797.593	7.232408	-0.3646	0.328387	0.458087	0.405914	0.219199	56.78488
0	45.87899	1.44E+10	-0.67818	4026.766	5.476581	-0.35597	0.306866	0.280175	0.220112	0.281031	61.01055
0	48.22828	1.72E+10	5.013942	4370.721	2.33683	-0.3101	0.250487	0.202606	0.064747	0.131783	61.37346
0	47.65491	1.78E+10	4.409979	4665.954	2.560809	-0.51653	0.300633	0.259721	0.095688	0.095251	61.85455
0	46.24668	1.88E+10	4.76795	4896.688	2.651203	-0.52085	0.177755	0.37285	0.069341	-0.03649	66.82353
0	30.3442	6.07E+09	7.146327	1468.702	-1.9	0.054811	-0.36846	-1.07109	-0.9403	-0.82093	16.82658
0	56.60243	1.29E+10	13.18089	1229.001	9.8	0.011201	-0.59443	-1.07787	-1.06216	-0.71345	21.61816
0	46.99064	1.84E+10	5.836925	1658.031	9.8	0.291217	-0.72668	-1.12108	-1.0607	-0.9229	15.35629
0	48.41702	2.32E+10	6.438218	2068.124	9.3	0.324801	-0.47608	-1.05339	-0.98378	-0.65469	13.21759
0	20.16926	6.82E+09	6.722437	474.5095	3.290214	-0.96329	-0.34502	-1.12448	-1.02177	-0.49175	52.80179
1.95E+08	30.7381	1.14E+09	31.94734	394.8601	7.084502	-0.32204	-0.27232	-0.76258	-0.47268	-0.40837	71.43419

Table 4: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
41003000	36.48329	1.51E+09	10.45738	345.1381	2.121835	-0.29756	-0.07041	-0.74084	-0.48243	-0.12619	111.2169
32225000	41.84772	1.94E+09	18.70073	279.6203	5.426674	-0.33955	-0.09504	-0.87648	-0.72645	-0.54284	114.5344
0	29.89189	7.01E+08	9.330315	458.604	5.213676	0.116533	-0.37183	-0.00171	-0.47268	-0.14468	68.67731
5625000	27.49957	4.48E+08	8.4866	376.9283	9.131659	-0.42384	-1.09477	-0.95996	-0.61573	-1.25195	0.466855
29021000	32.36432	4.87E+08	6.834768	413.7584	10.23864	-0.47547	-1.05284	-0.91528	-0.58154	-1.18152	0.444567
0	30.67932	4.72E+09	11.85968	246.2826	4.760065	0.043881	-0.44987	-0.25204	-0.13393	-0.63712	111.8795
0	108.3053	4.83E+10	1.807872	4130.678	5.390988	0.458356	0.527368	0.496208	0.211443	0.992509	43.04734
0	106.9434	5.08E+10	0.992816	4427.46	5.788499	0.455911	0.60385	0.570833	0.391868	1.173955	45.07664
0	115.3731	6.05E+10	1.518542	4918.167	6.783438	0.311378	0.494295	0.586404	0.425679	1.128076	45.69654
0	112.899	6.49E+10	2.960865	5553.993	5.332139	0.551897	0.611302	0.574833	0.268954	1.134693	42.06756
0	112.1859	7.01E+10	3.609236	6179.63	5.584847	0.257698	0.551699	0.52939	0.254842	1.196392	40.58853
0	106.1687	8.42E+10	2.027353	7218.265	6.298786	0.17347	0.535039	0.501285	0.280442	1.24741	40.08819
0	99.49932	1.07E+11	5.440782	8453.932	4.83177	0.078235	0.358578	0.396455	0.016727	1.155139	39.7996
0	91.41679	1.20E+11	0.583308	7277.981	-1.51369	-0.07022	0.307567	0.487884	-0.03116	0.995223	50.83575
0	93.31604	1.36E+11	1.710037	8754.243	7.42597	0.122589	0.594664	0.526331	0.133005	1.128926	51.05825
0	91.51535	1.47E+11	3.2	10060.39	5.187251	0.080936	0.58751	0.519978	0.051385	1.028537	51.52009
0	85.25315	1.95E+11	1.655362	10429.46	5.644607	-0.00714	0.549164	0.50613	0.299762	1.009908	53.25304
0	52.69868	5.56E+08	3.46216	4869.803	19.91912	0.797889	0.290966	0.116961	-0.53108	0.039627	31.55413
0	103.3162	8.32E+08	7.372987	5664.432	10.64095	0.092966	-0.02499	-0.02129	-0.83344	0.004673	31.22993
0	93.01278	8.99E+08	12.25763	6747.876	10.93389	-0.14339	-0.39074	-0.15887	-0.86937	-0.27489	32.54782
0	79.04151	9.65E+08	3.982837	6776.982	-6.04526	-0.2176	-0.41326	-0.16846	-0.6814	-0.45342	50.00969
0	86.06341	9.94E+08	6.609167	7161.527	7.196445	-0.1295	-0.39577	-0.32898	-0.52622	-0.2125	60.90254
0	98.35694	9.74E+08	12.82983	7412.549	10.83084	-0.20615	-0.4023	-0.57263	-0.52483	-0.3075	64.60157
0	48.96088	2.27E+09	2.549811	7082.296	3.048982	0.661059	0.869707	0.952071	0.629706	0.762369	37.77375
0	52.49056	2.73E+09	2.892969	7772.1	4.100202	0.582814	0.897997	0.862825	0.652759	0.848314	37.86787
0	53.42357	9.57E+09	6.531354	8984.63	3.886294	0.936932	0.849142	0.896201	0.59429	0.843472	36.36012
0	54.58665	1.02E+10	3.85216	9110.805	3.2	0.96183	0.983676	0.934985	0.333444	0.934179	37.23187
0	26.72859	1.56E+11	34.37766	4088.455	5.874767	-0.96659	0.390961	-0.75947	-0.44694	0.072263	26.02544
212000	25.74893	1.59E+11	15.928	4986.253	4.701842	-0.48622	0.284818	-0.58225	-0.38114	0.34839	23.32884
0	26.2767	1.52E+11	9.495019	6581.537	5.296474	-0.23096	0.288571	-0.44748	-0.23751	0.231028	19.74615
1.19E+08	55.26713	8.41E+08	23.5135	462.172	-5.2	-0.12653	0.017541	-0.18699	-0.20455	-0.37301	40.12299
0	48.52829	1.06E+09	7.786486	448.8414	-6.5	0.213316	-0.1969	-0.12304	-0.28694	-0.45342	85.82479
2.46E+08	49.77801	1.84E+09	31.14565	354.0013	2.1	-0.56328	-0.27633	-0.56059	-0.53624	-0.5829	73.01939
67838000	52.74043	1.96E+09	5.253304	458.6778	7.8	-0.26268	-0.40708	-0.64321	-0.94551	-0.60919	59.55925
3339000	53.4812	2.11E+09	11.62356	548.2897	6.6	-0.18183	-0.4724	-0.5919	-0.82478	-0.7105	52.53392
17733000	50.7068	2.11E+09	12.47907	720.9409	7.410062	-0.2681	-0.4385	-0.36836	-0.99003	-0.88644	51.95466
0	51.1443	2.22E+09	11.76593	831.2036	7.487165	-0.44137	-0.45988	-0.40093	-0.63885	-0.72651	32.42522
87514000	45.25597	2.62E+09	12.87406	950.6482	4.798522	-0.39872	-0.34623	-0.53784	-0.57594	-0.78719	29.17905
31836000	47.45246	3.37E+09	12.1355	1230.435	3.067944	-0.05075	-0.2779	-0.52547	-0.60159	-0.80748	23.23975
31488000	40.81904	3.66E+09	12.89708	1695.973	7.764846	-0.27043	-0.17391	-0.43243	-0.5677	-0.76126	18.44265

Table 5: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	36.87212	3.75E+09	-0.05869	1525.526	-5.98958	-0.5923	-0.1285	-0.47263	-0.6646	-0.57191	27.63223
0	39.22522	4.84E+09	7.352251	1631.536	7.094072	-0.38592	-0.10048	-0.39485	-0.68601	-0.64185	26.32298
0	44.97096	5.32E+09	7.610607	1970.571	6.414351	-0.07426	-0.08171	-0.37167	-0.62602	-0.60223	23.70906
0	43.48292	5.88E+09	4.639519	2046.537	-0.7	0.020272	-0.11363	-0.35924	-0.59781	-0.55232	24.32591
0	35.52433	5.34E+08	46.88995	580.9033	2.235094	0.6635	-0.18729	-0.03788	-0.1245	-0.39347	49.47963
0	47.84351	7.27E+08	9.357042	477.3441	3.339937	0.267135	-0.106	-0.0093	-0.23334	-0.29656	71.57533
0	53.99675	9.6E+08	11.59554	474.206	1.146062	0.754955	-0.09625	-0.12096	-0.40494	-0.30729	78.41833
0	52.35205	1.54E+09	5.128995	646.2369	7.004635	0.960644	-0.45745	0.13628	-0.16772	-0.28976	95.72553
2872000	59.44327	1.49E+09	5.095167	1333.878	8.556235	0.680718	-0.29961	-0.33867	-0.57931	-0.40839	50.49448
0	59.61445	1.74E+09	9.045246	1631.942	10.24802	0.684551	-0.26132	-0.38513	-0.65614	-0.51598	43.51401
0	50.278	2.99E+09	6.279233	1715.37	-1.2686	0.601512	-0.27676	-0.2761	-0.75531	-0.6583	59.50712
0	30.15159	1.81E+10	2.79562	1362.532	3.316036	-0.34848	-0.15844	-0.01063	-0.17606	-0.13932	56.46437
0	34.20127	1.79E+10	3.284762	2128.171	7.759852	-0.47488	-0.17652	-0.25337	-0.40364	-0.14268	57.31107
0	35.74857	2.07E+10	2.042085	2416.175	2.705774	-0.51117	-0.19717	-0.26181	-0.32318	-0.16347	53.52063
0	37.47806	2.1E+10	3.707317	2827.178	5.587056	-0.60018	-0.18266	-0.28804	-0.38	-0.17393	47.29757
0	28.70384	2.47E+10	0.994826	2860.992	4.758347	-0.41044	-0.04838	-0.19243	-0.30911	-0.13146	48.54353
0	33.237	2.72E+10	0.987355	2822.719	3.642975	-0.38302	-0.06833	-0.1568	-0.1753	-0.09059	52.3361
0	35.57532	2.99E+10	0.92236	3044.097	4.985647	-0.39455	-0.1058	-0.21514	-0.39747	-0.12718	56.83854
0	35.91191	3.38E+10	1.278741	2899.981	2.669166	-0.4623	-0.09218	-0.20571	-0.43675	-0.04544	59.65907
0	22.81758	2.4E+09	9.220467	214.1379	5.328284	-0.1512	-0.54606	-0.18896	-0.01462	-0.40837	65.19788
0	22.82205	2.67E+09	11.24447	218.9926	3.016389	-0.80587	-0.43434	-0.16801	-0.51479	-0.71839	66.35144
0	23.284	2.88E+09	2.47882	236.9828	6.2	-1.23486	-0.56162	-0.32785	-0.53622	-0.49631	64.57332
0	17.73713	3E+09	3.029399	251.0439	0.120267	-1.79486	-0.57629	-0.53223	-0.31914	-0.49068	63.93546
0	15.69986	3.22E+09	5.707009	258.1179	3.945038	-1.94807	-0.42366	-0.62379	-0.33956	-0.57007	62.5882
0	16.6827	3.37E+09	2.841811	291.869	4.682603	-2.12129	-0.52736	-0.76095	-0.80815	-0.75364	59.41567
0	14.58369	3.19E+09	6.836333	321.455	3.479181	-2.11257	-0.49766	-0.83631	-0.63117	-0.84019	51.76462
0	13.44661	3.4E+09	6.920336	352.801	3.364615	-1.92247	-0.50084	-0.62906	-0.62714	-0.78439	49.68697
0	12.85571	3.62E+09	5.74591	397.904	3.41156	-1.9154	-0.55421	-0.64124	-0.71451	-0.6548	42.97866
0	12.77582	3.7E+09	9.878396	477.9322	6.104639	-1.84045	-0.61827	-0.72207	-0.72257	-0.7589	43.79999
0	9.582536	3.79E+09	9.324001	596.0905	4.816415	-1.59619	-0.74329	-1.0081	-0.64831	-0.86062	33.86252
0	39.7879	3.41E+10	14.03178	510.4169	10.35418	-1.65142	-1.24022	-1.52285	-1.32013	-0.96479	66.43297
1.09E+08	30.16075	3.67E+10	14.99803	645.9257	33.73578	-1.72088	-1.32287	-1.43203	-1.3045	-0.91259	53.6304
0	31.65697	2.05E+10	17.86349	804.1524	3.444667	-1.64829	-0.76769	-1.3613	-1.15879	-0.88329	28.64545
1144000	43.11133	3.96E+09	8.239527	1014.757	8.210965	-2.03632	-0.88703	-1.08112	-1.07408	-0.96124	11.78421
0	33.72852	3.75E+09	5.382224	1130.88	6.828398	-2.01304	-0.86435	-1.06521	-0.9835	-1.04088	12.4049
0	39.88313	4.04E+09	11.57798	1376.016	6.270264	-1.86165	-0.78015	-1.05953	-0.81105	-0.96723	11.52877
0	30.76862	6.77E+09	11.53767	1090.747	6.934416	-1.95247	-0.72846	-1.16418	-0.97647	-1.20063	15.13073
0	25.26412	7.21E+09	13.7202	2310.861	7.839739	-2.19361	-0.71375	-1.17335	-0.99733	-1.15112	9.449543
0	31.32981	8.96E+09	10.84079	2507.683	4.887387	-1.9469	-0.6694	-1.21265	-1.13051	-1.0684	10.2327
0	31.43875	1.01E+10	12.21701	2730.223	4.279277	-2.05749	-0.72248	-1.17788	-1.14037	-0.992	10.41934

Table 6: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	16.48479	3.22E+10	6.228004	453.4948	2.550234	-1.17621	-0.49336	-0.76509	-0.96052	-0.44895	79.08469
0	59.40451	2.51E+09	11.62418	1064.959	7.733696	-0.81291	-0.50683	-0.7049	-0.3507	-0.28099	55.64337
0	54.04149	2.71E+09	13.57419	742.3678	-3.76911	-0.23576	-0.56815	-0.68842	-0.90141	-0.53077	66.35015
0	66.17836	2.3E+09	15.59589	654.6199	-2.49484	-0.3263	-0.48833	-0.89497	-0.81651	-0.42603	57.74153
0	60.91581	1.87E+09	11.79945	529.9295	-0.1589	-0.62842	-0.60846	-1.04255	-0.90961	-0.46525	69.73446
0	29.97211	2.93E+10	2.002258	3134.527	7.528817	-0.85038	0.15896	-0.73553	-0.21683	-0.58014	32.82874
0	30.49846	3.22E+10	1.779986	3606.707	8.518442	-0.76005	0.275959	-0.76491	-0.26052	-0.5098	28.52882
0	28.39296	3.47E+10	5.785876	4246.948	9.143148	-0.89731	0.349068	-0.74538	-0.2042	-0.35724	25.85717
0	25.18603	3.74E+10	2.936232	4188.529	1.049232	-1.18094	0.401356	-0.66275	-0.34216	-0.42241	26.32041
0	26.55976	4.22E+10	1.528321	5075.477	8.450747	-0.97689	0.460709	-0.60022	-0.25043	-0.20314	23.51761
0	29.65525	4.5E+10	3.370685	5759.409	6.452216	-0.73778	0.476881	-0.60998	-0.24931	-0.14568	20.86108
0	27.13283	5.41E+10	3.653732	6425.182	5.950346	-0.8719	0.489412	-0.61237	-0.39492	-0.15763	19.21191
0	40.50666	4.4E+10	7.476104	1163.848	5.845873	-0.47865	0.258618	-0.00538	-0.1764	-0.176	53.18962
0	44.76423	5.36E+10	9.234934	970.6139	-0.57672	-0.29563	0.343132	0.006995	-0.14522	0.034731	60.12527
0	51.36929	5.85E+10	3.977125	1043.456	4.411213	-1.41428	0.157829	-0.43861	-0.45494	-0.14404	49.58106
0	46.74701	6.01E+10	2.722772	1004.991	3.645898	-0.90837	-0.0934	-0.42673	-0.45415	-0.08814	62.76997
0	47.15706	6.28E+10	2.289157	1015.78	4.970364	-1.5832	-0.03197	-0.52647	-0.5332	-0.03925	67.74914
0	48.57265	6.11E+10	4.829211	1084.765	6.697636	-1.67803	-0.26461	-0.57445	-0.60339	-0.20291	73.94497
0	30.78059	6.14E+10	3.172086	2587.617	6.801331	-1.16435	-0.05798	-0.54554	-0.58429	0.077567	51.45198
3504000	101.3908	1.5E+09	36.96476	11123	-2.14052	0.773309	-0.71218	0.225711	0.245717	0.093551	176.79
75640000	107.9906	1.71E+09	31.75444	9707.265	-1.10718	0.617699	-0.62431	0.061972	0.313525	0.101325	100.3392
2.43E+08	93.79552	1.49E+09	-2.40464	10805.1	5.945431	0.884741	-0.56993	0.023537	0.292838	0.181529	75.53937
35517000	95.76979	1.78E+09	2.559268	12188.98	7.894706	0.96473	-0.43543	-0.02557	0.254846	0.255884	78.84189
8617000	92.7854	2.02E+09	7.110371	12844.83	6.039788	0.756869	-0.30984	-0.0373	0.333867	0.375671	72.99266
0	27.15889	2.54E+10	5.338953	3099.132	4.2	-0.25486	0.402365	0.101098	0.612061	0.691411	44.11324
0	34.97334	8.3E+09	15.93583	757.9482	3.8	-1.80418	0.079618	0.172926	-0.09766	-0.25518	92.34398
0	36.24382	9.03E+09	9.364243	840.8738	4.698397	-1.51895	0.264753	0.083203	-0.19009	-0.4466	89.19582
0	39.01571	9.17E+09	6.176276	854.9267	6	-1.93095	0.24772	0.171398	-0.26153	-0.35368	96.90406
0	34.91333	9.78E+09	9.551032	903.8964	3.964656	-0.84681	0.182453	0.324318	-0.24004	-0.06122	102.0322
0	34.6535	1.05E+10	6.314638	984.8102	5.940269	-0.87738	0.102125	0.243077	-0.25909	-0.20106	102.2707
0	35.3309	1.11E+10	7.575926	1063.161	5.445061	-1.05855	-0.04425	0.203612	-0.13885	-0.39765	102.3285
2.65E+08	32.33687	1.15E+10	11.63969	1242.428	6.241748	-1.19272	-0.35152	0.146073	-0.36986	-0.28768	90.60491
0	30.12853	1.2E+10	10.02018	1424.101	7.668292	-1.42647	-0.24905	0.186362	-0.17649	-0.17982	88.69948
0	29.115	1.45E+10	15.84211	1614.364	6.796826	-1.737	-0.26369	0.144896	-0.09768	-0.08964	84.99442
0	22.37683	1.99E+10	6.217649	2399.92	8.015959	-0.92311	-0.20471	-0.07555	-0.39923	-0.18335	81.92648
0	23.05506	2.21E+10	6.716768	2835.813	8.24591	-0.69914	-0.10634	-0.07131	-0.37384	-0.10161	78.45208
0	22.83374	2.35E+10	7.542914	2921.66	6.341362	-0.7033	-0.11831	-0.10789	-0.23981	-0.23652	79.17225
0	49.52459	1.36E+08	2.142406	2945.172	5.175199	1.026516	0.28005	0.392322	0.1257	-0.16784	47.06363
0	6.702094	1.68E+10	17.10506	345.6092	4.308504	-2.29818	-1.35283	-1.62613	-1.01586	-1.21629	7.869492
0	65.68064	5.85E+10	1.80435	2211.874	7.139975	-0.15415	0.370559	0.147993	-0.1536	0.376201	28.87396

Table 7: Country risk data

Debt	ExportsPercentageOfGDP	ExtDebtPercentageofGDP	Inflation	GDPPerCapita	GDPGrowth	PoliticalStabilityNoViolence	RegulatoryQuality	RuleOfLaw	ControlOfCorruption	GovernmentEffectiveness	PublicDebtPercentageGDP
0	70.69705	5.84E+10	2.759149	2478.818	6.344073	-0.69127	0.251331	0.113795	-0.15463	0.362202	26.14378
0	73.56771	5.86E+10	4.540369	2689.951	4.604699	-0.8519	0.455116	0.087956	-0.10134	0.429161	27.32998
0	73.6469	6.25E+10	4.637474	3143.237	5.092899	-1.14219	0.253418	-0.02684	-0.34489	0.404782	26.10769
0	73.42447	6.27E+10	2.241541	3737.719	5.044316	-1.1526	0.146416	-0.0835	-0.36254	0.372619	24.47986
0	76.4446	6.66E+10	5.468489	4118.401	2.4843	-1.28478	0.236602	-0.13192	-0.41993	0.25338	23.99906
0	68.35168	8.08E+10	-0.84572	3978.908	-2.32985	-1.41542	0.240433	-0.22393	-0.28047	0.280718	28.60696
0	42.10938	1.14E+10	3.725145	2154.987	7.146081	0.156455	0.012376	-0.20212	-0.22141	0.409988	55.29652
0	38.51979	1.08E+10	3.125366	2336.084	4.783763	0.200897	-0.08034	-0.14501	-0.03078	0.552958	54.28174
0	39.5467	1.14E+10	2.962308	2247.913	4.70986	0.276054	-0.06224	-0.15064	-0.02804	0.523814	56.72147
0	40.78459	1.55E+10	2.721033	2373.781	1.717043	0.074198	-0.15476	-0.0392	0.550672	0.625516	55.93409
0	39.48345	1.84E+10	2.712592	2790.004	5.379599	0.305225	0.002591	-0.0862	0.322893	0.554804	55.02116
0	42.22114	1.96E+10	3.63228	3139.537	5.615627	0.136262	-0.07513	0.145145	0.239532	0.476308	53.83742
0	44.93229	1.79E+10	2.017786	3217.886	3.819175	0.04597	-0.10959	0.097936	-0.08597	0.421916	52.42053
0	46.02701	1.87E+10	4.490514	3394.365	5.65215	0.239225	0.119924	0.201595	-0.07464	0.577048	48.56232
0	51.08611	2.05E+10	3.416547	3805.277	6.22676	0.18994	0.056642	0.172598	-0.10592	0.466197	45.78954
0	56.1743	2.14E+10	4.920696	4342.676	4.736813	0.121079	0.06825	0.139754	-0.18315	0.309018	43.29193
0	45.83365	2.27E+10	3.524814	4162.587	3.608145	0.058036	0.002306	0.199438	-0.11001	0.403822	42.92979
0	50.05189	2.25E+10	4.416269	4212.155	3.249399	-0.03957	-0.02044	0.121903	-0.14903	0.238128	40.31929
0	49.18812	2.26E+10	3.544029	4305.064	-0.5083	-0.36881	-0.1923	-0.12741	-0.17766	0.058434	44.4886
0	49.18368	2.53E+10	5.138117	4197.528	4.661335	-0.74212	-0.20776	-0.14671	-0.15069	-0.01599	44.46773
0	23.90792	2.89E+11	10.44413	10379.49	0.658838	-0.84521	0.266911	0.080984	0.08151	0.262214	43.95434
0	23.31643	2.77E+11	6.250977	8626.398	-4.82588	-1.0319	0.298808	0.101826	0.07408	0.289299	54.01662
0	21.20914	2.99E+11	8.566444	10135.75	9.156953	-0.92075	0.308698	0.117665	0.029938	0.308	51.33859
0	23.97664	3.05E+11	6.47188	10604.55	8.772748	-0.95561	0.381906	0.079893	0.055267	0.362708	46.93753
0	26.29611	3.37E+11	8.89157	10660.73	2.127461	-1.19409	0.416153	0.035491	0.168109	0.40377	45.14453
1.49E+08	9.639048	3.94E+09	0.068804	289.0556	4.905265	-1.2184	0.249619	-0.63805	-0.91764	-0.38781	57.04699
34869000	10.65141	3.54E+09	3.392022	255.1219	3.141907	-1.34297	0.078954	-0.79174	-0.85238	-0.3827	42.92659
3476000	12.69688	4.77E+09	3.721287	285.9642	6.807233	-1.296	0.004128	-0.63475	-0.74706	-0.34451	69.34236
1.54E+08	14.1797	4.44E+09	8.448726	313.799	6.332565	-1.43329	-0.1789	-0.56264	-0.8463	-0.53846	64.94707
10131000	15.27541	1.28E+09	7.310676	334.6392	10.78474	-1.1594	-0.20421	-0.33599	-0.75394	-0.4811	57.92775
0	16.72507	1.63E+09	6.138511	400.0428	8.412426	-0.9619	-0.1958	-0.37946	-0.79768	-0.40908	26.4924
0	24.28014	2.27E+09	12.05086	448.0669	8.708752	-0.90864	-0.2156	-0.37791	-0.83054	-0.57443	33.1932
0	19.8144	2.74E+09	13.01726	517.0613	7.251045	-0.98957	-0.14938	-0.41831	-0.88713	-0.61643	28.9164
24500000	17.45698	2.97E+09	3.976553	553.2626	5.170343	-1.00575	-0.15477	-0.39135	-0.90439	-0.52032	27.54934
0	18.43656	3.26E+09	18.6929	530.934	9.673225	-0.98697	-0.14052	-0.34618	-0.89532	-0.50403	35.43992
0	19.90915	3.78E+09	14.01606	652.7496	4.41126	-0.8824	-0.24379	-0.35694	-0.98421	-0.56257	33.2118
2.18E+09	62.44488	1.39E+10	28.2031	635.709	5.9	-0.50408	-0.52272	-1.14185	-1.07161	-0.74846	45.28747
0	46.92328	9.83E+10	25.23191	3891.038	2.3	0.032038	-0.51925	-0.69342	-0.78759	-0.70791	13.8267
0	46.37503	1.04E+11	15.89457	2545.48	-14.8	-0.31069	-0.57082	-0.77416	-1.00693	-0.79552	24.87512
0	50.74642	1.25E+11	9.378589	2973.996	4.2	-0.02388	-0.51552	-0.81299	-0.97517	-0.74673	29.88033
0	49.81634	1.36E+11	7.960095	3569.757	5.2	-0.08006	-0.6063	-0.82814	-0.99623	-0.80661	27.48276
0	47.72078	1.36E+11	0.555556	3855.421	0.2	-0.10219	-0.61147	-0.79457	-1.02875	-0.58328	33.70309
0	42.28952	64166000	3.28216	1472.99	4.301466	1.026516	-0.46053	-0.48889	-0.22805	-0.44098	28.4023
2.03E+08	28.47891	7.06E+09	43.0731	396.4618	6.218547	-0.226	-0.42184	-0.64543	-1.02881	-1.06042	175.9345
77529000	24.61776	6.87E+09	24.45846	369.3811	-0.38575	0.092711	-0.11721	-0.52713	-0.88067	-0.85642	161.2322

Table 8: Country risk data