

**EFFECT OF COST OF DOING BUSINESS ON THE FORMALIZATION OF
BUSINESSES IN KENYA**


GATHUMBI LILIAN WANGARI

**A RESEARCH PROJECT SUBMITTED TO THE DEPARTMENT OF
APPLIED ECONOMICS IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE AWARD OF THE MASTER OF ECONOMICS
(FINANCE) DEGREE OF KENYATTA UNIVERSITY**

MAY, 2023

DECLARATION

This research project is my original work and has not been presented for a degree in any other university or any other award.

Signature 
Gathumbi Lilian Wangari

Date 17/05/2023

Masters of Economics (Finance), Kenyatta University
K102/CTY/PT/27751/2014

I confirm that the work reported in this project has been carried out by the candidate under my supervision

Signature 

Date 17/05/2023

Dr. James Maingi

Department of Applied Economics

School of Economics

Kenyatta University

Nairobi, Kenya

DEDICATION

This research project is dedicated to my loving parents who have been very supportive in this journey.

ACKNOWLEDGEMENT

I wish to thank God and Kenyatta University, School of Economics for giving me the opportunity to do my Master's degree. I also acknowledge the efforts of my supervisor Dr. James Maingi, for his guidance during the writing of this project. I thank my parents for the support they have given me up to this level. May God bless them. I also thank my classmates and workmates for the support and encouragement they gave me during this journey.

TABLE OF CONTENTS

DECLARATION.....	ii
DEDICATION.....	iii
ACKNOWLEDGEMENT.....	iv
LIST OF FIGURES	viii
LIST OF TABLES	ix
ABBREVIATIONS AND ACRONYMS.....	x
OPERATIONAL DEFINITION OF TERMS.....	xi
ABSTRACT	xii
CHAPTER ONE	1
INTRODUCTION.....	1
1.1 Background	1
1.1.1 Overview of the Informal Sector in the World	4
1.1.2 The informal economy in Africa.....	6
1.1.3 Informal sector overview in Kenya.....	8
1.1.4 Measures undertaken to encourage formalization in Kenya.....	11
1.1.5 Overview of Cost of doing business in Kenya.....	12
1.2 Statement of the Problem	14
1.3 Research Questions	16
1.4 Objectives of the Study	16
1.5 Significance of the Study	16
1.6 Scope of the Study	17
1.7 Organization of the Study	17

CHAPTER TWO	18
LITERATURE REVIEW	18
2.1 Introduction.....	18
2.2 Theoretical Literature Review	18
2.2.1 Modernization Theory	18
2.2.2 Dependency Theory	19
2.2.3 Structuralism Theory	20
2.2.4 Institutional theory	21
2.2.5 Neo-liberalism Theory	22
2.2.6 Discrete Choice Random Utility Model	23
2.3 Empirical Literature	24
2.3.1 Global Empirical Literature	24
2.3.2 Empirical literature from Kenya	31
2.4 Overview of the Literature	32
CHAPTER THREE	34
METHODOLOGY	34
3.1 Introduction	34
3.2 Research Design.....	34
3.3 Theoretical Framework	34
3.4 Specification of the model	38
3.5 Definition and Measurement of variables.....	39
3.6 Data Type and Sources	40
3.7 Data Analysis	40
3.8 Diagnostics Tests	41

CHAPTER FOUR.....	42
EMPIRICAL FINDINGS.....	42
4.1 Introduction.....	42
4.2 Descriptive Statistics.....	42
4.3 Diagnostic Tests.....	46
4.3.1 Multi collinearity test.....	47
4.3.2 Classification report.....	47
4.3.3 Hosmer Lemeshow Test	48
4.3.4 Specification Test.....	49
4.4 Empirical Findings.....	49
4.4.1 Sectoral distribution of formal and informal enterprises in Kenya.....	49
CHAPTER FIVE: SUMMARY, CONCLUSION AND POLICY IMPLICATION	57
5.1 Introduction.....	57
5.2 Summary	57
5.3 Conclusion	59
5.4 Policy Implications	59
5.5 Scope for Further Research.....	60
REFERENCES.....	62
APPENDICES	71

LIST OF FIGURES

Figure 1.1: The scope of the informal economy for different country income groups..	4
Figure 1.2: The scope of the informal economy for different World regions.....	5
Figure 1.3: The magnitude of the informal economy for different regions in African..	6
Figure 1.4: Trends in the number of jobs in both formal and informal economy in Kenya.	8
Figure 1.5: Trends in proportion of both formal and informal sector in the total employment in Kenya.	9
Figure 1.6: Trends in the number of new jobs in both the formal and informal economy in Kenya.	10
Figure 1.7: Trends in employment growth rate in formal and informal economy in Kenya.	10
Figure 4.1: Mean monthly cost by formality status	46
Figure 4.2: AUC curve.....	48
Figure 4.3: Distribution of top 10 sectors by formality status	51

LIST OF TABLES

Table 3.1: Definition and measurement of variables	39
Table 4.1: Descriptive statistics of the categorical variables	43
Table 4.2: Descriptive statistics of the categorical variables	44
Table 4.3: Descriptive statistics of the continuous variables	45
Table 4.4: Multi collinearity test results	47
Table 4.5: Sectoral distribution of formal and informal enterprises in Kenya	50
Table 4.6: Logistic regression results	51
Table 4.7: Marginal Effects results	53
Table 4.8: Sales revenues by formality	54

ABBREVIATIONS AND ACRONYMS

CFED:	Centre for Free Enterprise and Democracy
COFEMER:	Federal Commission for Regulatory Improvement
DTF:	Distance to Frontier
GDP:	Gross Domestic Product
IFS:	Informal Enterprise Surveys
ILD:	Institute for Liberty and Democracy
ILO:	International Labour Organization
INEGI:	Institute of Statistics, Geography and Informatics
IMSS:	Mexican Social Security Institute
LIC:	Low Income Countries
MIMIC:	Multiple Indicator-Multiple Cause
MSEs:	Micro and Small Enterprises
OECD:	Organization for Economic Co-operation Development
SFOF:	System of Fast Opening of Firms
USAID:	United States Agency for International Development

OPERATIONAL DEFINITION OF TERMS

Cost of doing business	This is all the expenses a business incurs while producing and selling goods and services. Some of these costs include start-up costs which include; registration and licensing cost, raw materials cost, renting cost, tax compliance cost, employees' salary cost, marketing costs, office utilities & supplies, insurance costs, inventory costs among others.
Informal sector	It involves all the economic activities that are income generating and are operating outside the official legal framework.
Formal Sector	It involves all the economic activities that are income generating and are operating within the official legal framework.
Jua Kali	It refers to the informal sector of business owners and traders running shops in open market spaces and streets.

ABSTRACT

The informal sector has been growing at a higher pace than the formal one in the last three decades. A large informal sector is undesirable because it does not pay tax, it faces a lower cost of business which leads to unfair competition and is not protected by the legal framework provided by the state. Over the years, the government of Kenya and policymakers have come up with measures to encourage the growth of the formal sector such as making business registration cheaper, easier, and more attractive; rationalizing business licenses, and creating a good business environment. These measures have been intended to increase the growth of the formal sector, which is more productive and provides more decent jobs. However, despite these efforts, the growth of the formal sector has been slow compared to the informal. According to various Kenya Economic survey reports, the sector has been increasing over the years with over 80 percent of all the new businesses formed being in the informal sector. The goal of this study was therefore to find out the cost of doing business on the growth of the formal and informal sector in Kenya. The objectives were to: find out the sectoral distribution of the formal and informal enterprises in Kenya, to investigate the effect of the high cost of electricity, licenses, and taxes on the cost of doing business on the growth of the formal and informal sector in Kenya. The study utilized the secondary cross-sectional household survey data gathered by the Kenya National Bureau of Statistics in 2016. Descriptive analysis was used to find out the sectoral and sub-sectoral distributions of firms while the second objective was answered using a logit model. The sectoral and sub sectoral distribution found that majority of the businesses were in the wholesale, retail and motor vehicle repair. Furthermore, it was found that the cost of credit and formality were negatively related, but positively related with the costs of electricity, rent, taxes and licenses. The conclusion was that high cost of business impacts the formalization of businesses in Kenya. The research recommends reduction of cost and improvement of business environment constraints that hinder business formalization. Since low education and business size contributed largely to businesses becoming informal, the government should enhance higher education opportunities for its citizens and provide training opportunities to small businesses to educate them on how to improve on their competitiveness.

CHAPTER ONE

INTRODUCTION

1.1 Background

An individual can conduct activities in the economy in two ways. These are formal and informal. Different authors and researchers have defined the informal sector in a context framework of their specialization. Due to its heterogeneous nature, it has been difficult for the researchers to state a single definition of the informal sector (Farazi, 2014). Therefore, there has been no definition of the informal sector that holds for all stakeholders including economists, operators in the informal sector, governments and policymakers, and donors (Dada, 1995)

Schneider, Montenegro, and Buehn, (2010) describes the informal sector as the production of legal market-based goods and services but the production operates outside the legal frameworks to avoid paying tax, to escape social security contribution, avoid meeting legal requirements and market standards. Smith (1994) describes the informal sector as the production of legal or illegal market-based goods and services while avoiding detection by the official GDP estimates. Del'Anno (2003) and Del'Anno and Schneider (2004) illustrates the informal sector as the income generating economic activities that are concealed from government regulations. De Soto (1989) describes informal economy as the collection of businesses, labours, and activities that function outside the legal framework. Loayza, Luis, and Naotaka (2009) adopts a similar description of the informal sector to De Soto (1989). The United Nations (1993) describes informal sector as the components for producing goods and services for individual consumption. These components have low organizational levels, limited labour and capital, no division of labour and capital, and operate at a small-scale level.

The informal sector is not governed by the formal establishments of society, despite the fact that there are different definitions of it. Low resources, family ownership, ease of entry, labour-intensive, illegitimate systems for skill acquisition, and adapted technology are other characteristics of the informal economy (Dada, 1995; Schneider, Buehn & Montenegro, 2010). Informal activities are found in but are not limited to, unregistered enterprises, self-employed manufactures, street vendors, and casual retailers. Also, if a registered business engages in an undocumented transaction or uses a regulatory framework for an alternative purpose, these activities are termed informal. These activities are found all over the world. They are found in the rural and city areas, in the developed nations, in the developing nations, in people earning below minimum wage, in people earning above minimum wage. Although informal activities bring about benefits in the economy, some activities in the sector can be criminal (Dada, 1995). However, this study does not deal with the criminal activities of the informal sector.

Most of the literature done focuses on explaining the causes of informality. One strand of literature views informality as a result of the cost of doing business. The cost of doing business is defined as all the expenses a business incurs in producing and selling goods and services. Some of these costs include start-up cost, registration and licensing cost, raw materials cost, renting cost, tax compliance cost, employees' salary cost, among others (Johnson, Kaufmann, and Shleifer, 1997). The individuals who operate in the informal sector try to avoid these burdens. These business owners believe that the advantages of working in the unregulated market outweigh the disadvantages (Johnson, Simon & Kaufman, 2000). An important and growing component of the economies of developing nations is made up of the informal sector. In underdeveloped nations, the majority of enterprises are informal. These are the

small, medium-sized, and microbusinesses (La Porta & Shleifer, 2008). Additionally, the informal sector accounts for 40–50% of the GDP of emerging nations (Schneider, Buehn & Montenegro, 2010). A large portion of the population relies on the informal sector for their livelihood, which is estimated to employ around 60% of the non-agricultural workforce in developing nations (ILO, 2012). The informal sector is multidimensional, involving several markets that carry out a variety of operations, much like the formal sector. These include the informal financial sector, the informal corporate or commercial sector, and the informal labor market, among others (Garcia-Bolivar, 2006). Therefore, this depicts informal economy as a vital component of economic growth in the developing countries.

Policymakers in developing countries have been apprehensive about the portion of their country's informal economy for several reasons. To start with, the informal sector avoids/evades tax thus limiting the government in its mandate to provide goods and services (Levy, 2008). Second, the co-occurrence of both the formal and informal businesses in the same industry could lead to competition, and firms competing formally could face increased marginal production costs. This could result in inefficient allocation or misallocation of resources (Hsieh & Klenow, 2009). Third, informal firms have a cost advantage and this could lead to inequitable competition with the formal firms hence restricting economic growth (Farrell, 2004). Fourth, informal businesses are not protected by the legal framework provided by the state. Finally, Informal enterprises are unable to utilize the services that are only available to the businesses that have fully complied with the rules and regulations. Such services are credit facilities by the banks, training of managers and employees, trade fairs, infrastructure services, export products, among others. The lack of these services limits the informal enterprises to grow.

The informal sector is not a new phenomenon in Kenya. The sector also known as *Jua Kali* in the country spreads across various industries. Some of the informal activities in Kenya are street vending and non-confined trading activities, carpentry, welding, handicrafts, food processing, metal works which constitutes small scale manufacturing, maintenance, and repair activities, construction and transport, restaurant & catering, hairdressing, and beauty, childcare, and guarding, among others (Hope, 2014).

1.1.1 Overview of the Informal Sector in the World

The informal economy exists everywhere. However, developing nations continue to experience the informal economy (La Porta & Shleifer, 2014). For instance, it has been asserted that the informal sector predominates in the largest and fastest-growing industry in Africa, such as construction. Figure 1.1 demonstrates the portion of the informal economy for different country income groups.

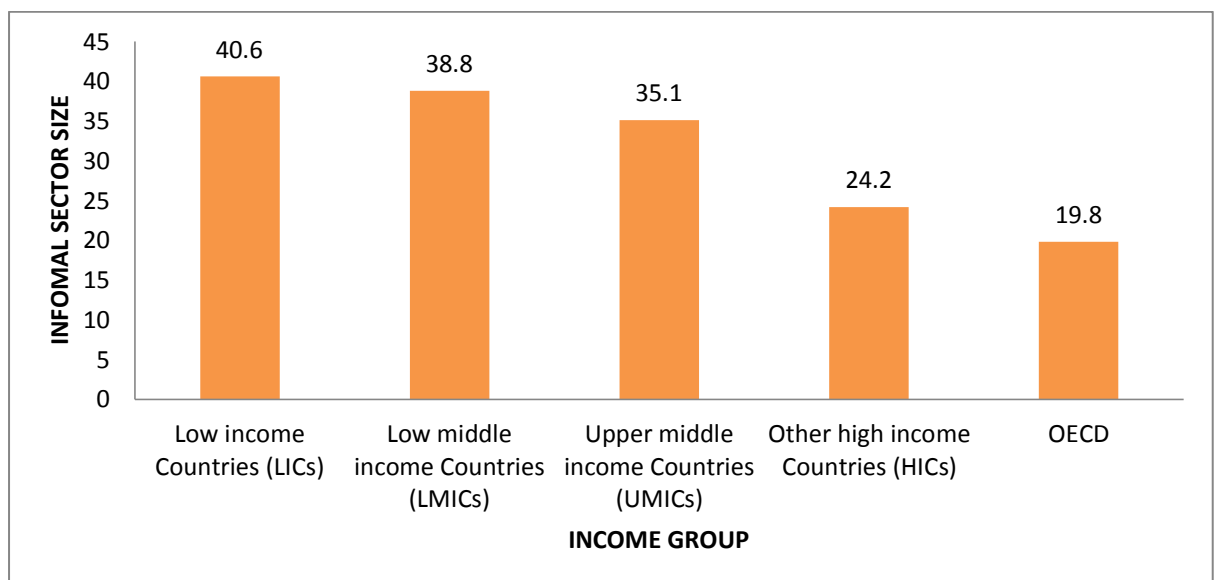


Figure 1.1: The scope of the informal economy for different country income groups

Data Source: Schneider & Buehn (2018)

According to Figure 1.1, when income levels rise, the size of the informal economy shrinks. The informal economy is largest in low-income (LIC) nations. When compared to Organization for Economic Co-operation and Development (OECD) countries, the size of the informal sector in low-income countries is twice as large. According to Medina's (2017) assessment, governments in high-income nations have a greater capacity and motive to ensure formalization.

When informality is compared across the different regions in the World, Africa and Latin America have the highest levels of informality at 39.9 percent and 37.8 percent, respectively. Figure 1.2 displays the dispersal of the informal sector for the different regions in the World.

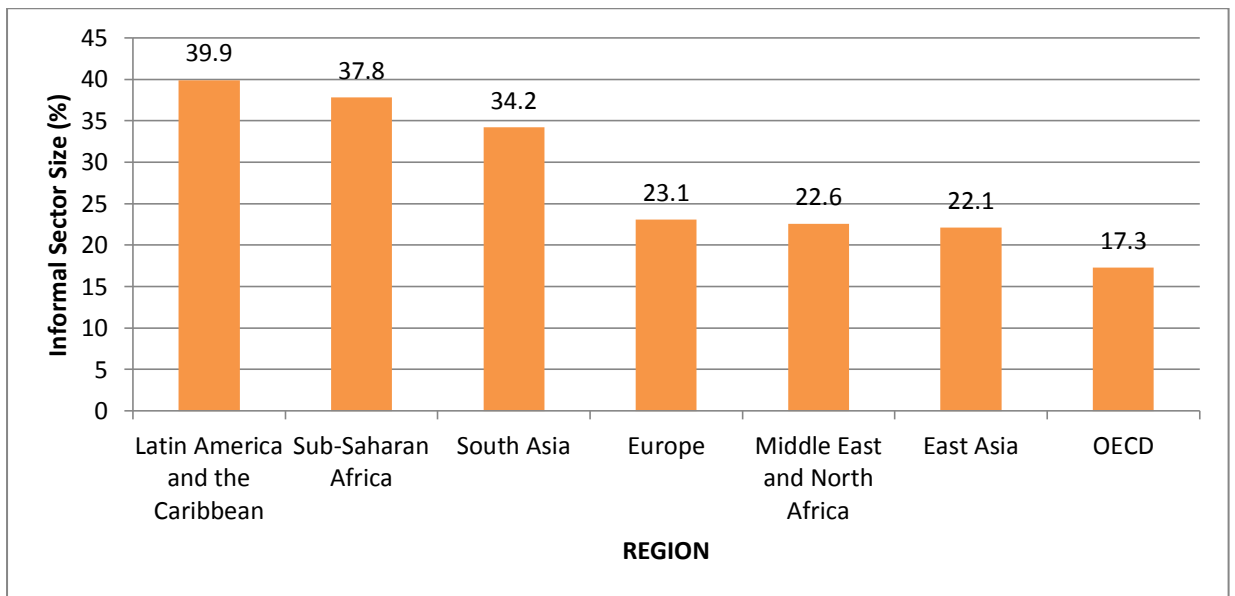


Figure 1.2: The scope of the informal economy for different World regions

Data Source: Schneider & Buehn (2018)

East Asia and OECD have the lowest percent of informal sector standing at 22.1% and 17.3% respectively. Both Latin America and Sub-Saharan Africa constitute the highest number of developing countries. As noted above the rate of informality is high in developing countries. This is because of the high cost of starting a business which includes procedures, cost and time of registration, and high taxation rate. The average

number of procedures involved in registering a business in Sub-Saharan Africa and Latin America are 7.4 and 8.2 respectively. In addition, the average number of days involved in registering a business in Sub-Saharan Africa and Latin America are 23.3 and 28.5 respectively. This is compared to the average number of procedures and the average number of days involved for registering a business in Europe and Central Asia which are 5.2 and 11.9 respectively (World Bank Group, 2018).

1.1.2 The informal economy in Africa

The informal economy in African countries is prevalent. Additionally, the informal economy forms a significant share of all the African countries. The idea regarding the informal economy was first applied in Africa to describe enterprises that generate income outside the formal sector (Hart, 1973). According to ILO (2016), non-agricultural workforces in the informal economy stands at 71.9 percent in Africa. This means that the shadow economy contributes to earnings and livelihood for a large portion of the population in Africa. Figure 1.3 depicts the magnitude of the informal economy for different regions in African.

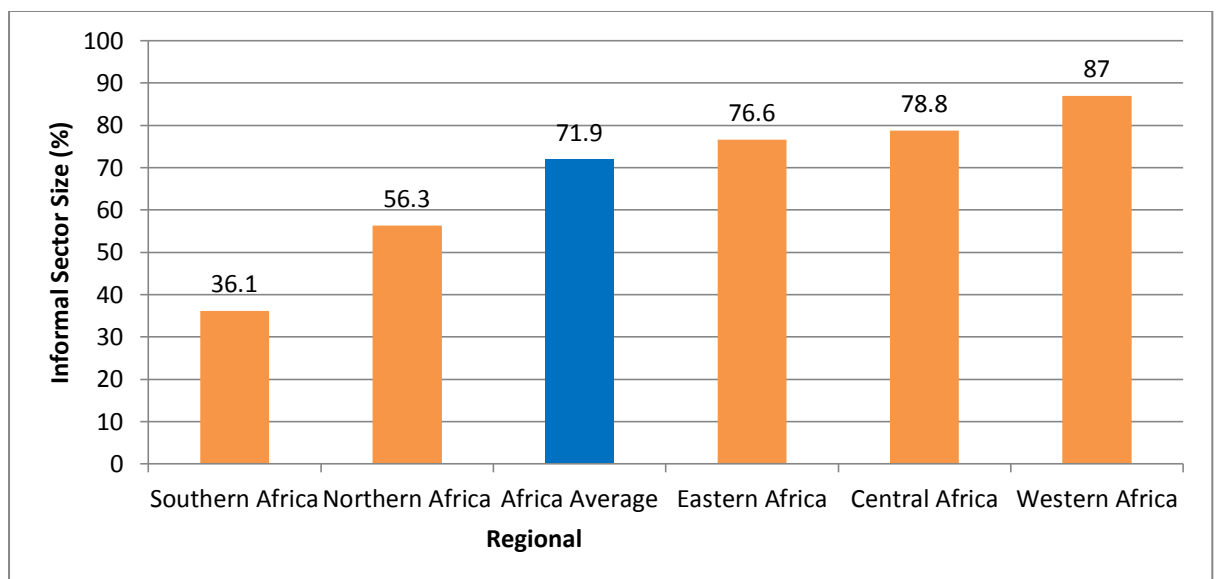


Figure 1.3: The magnitude of the informal economy for different regions in African

Data Source: ILO (2016)

According to Figure 1.3, Southern Africa has the lowest percentage of the informal economy at 36.1%, followed by Northern Africa at 56.3%. The informal economy is most prevalent in Central and Western Africa, at 78.8% and 87%, respectively. With 76.6% of the total economy, the informal sector is only moderately large in Eastern Africa. The World Bank Group (2018) claims that due to the high costs allied with conducting business in the formal sector, the informal sector is more prominent in Eastern, Western, and Central Africa. These include high start-up expenses that involve time-consuming governmental processes, expensive registration procedures, and high tax rates.

As per the ILO (2016), informal employment in Africa contributes more to women's employment and livelihoods than it does to men, where it accounts for 78.6% of women's employment and 67.7% of men's employment. The rate of informal employment rate among the young and old is high standing at 94.9% for persons aged between 15-24 and 96.0% for a person aged 65 and above (ILO, 2016). The level of education is negatively related to the rate of informal employment. Those with no education have the highest rate of informal employment at 94.0 percent and those with tertiary education have the lowest rate of informal employment at 27.0 percent. Informal employment dominates both rural standings at 88.3 percent and urban standing at 76.3 percent areas. Informality in the agricultural sector stands at 97.9 percent, in the industry sector stands at 77.4 percent, and in the service, the sector stands at 70.2 percent (ILO, 2016).

1.1.3 Informal sector overview in Kenya

The informal sector in Kenya is large and dynamic. It makes up the largest percentage of all enterprises in Kenya. The informal economy contributes approximately 82% of the total Kenyan employment (The Republic of Kenya, 2019). Therefore, the informal sector is a vital contributor to economic development in Kenya as a source of employment, earnings, and household livelihoods.

Figure 1.4 presents the trends in the total number of jobs in both the formal and informal sectors in Kenya from 1985 to 2018. The figure displays that the number of both the formal and informal sector jobs has grown since 1985. However, the number of informal jobs has grown exponentially from 254,000 jobs in 1985 to 1,486,000 jobs in 2018 while that of the formal sector has slowly grown from 1,207,000 jobs to 2,917,000 jobs. Figure 1.4 displays that the portion of the informal sector in the total employment has been increasing since 1985 while the portion of the formal sector has been decreasing since then.

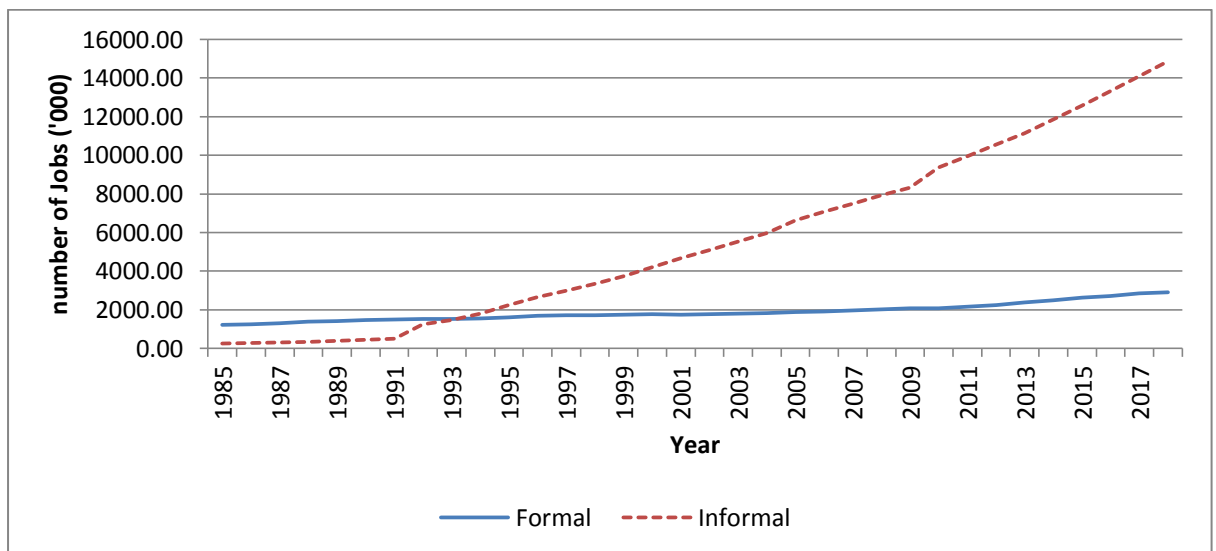


Figure 1.4: Trends in the number of jobs in both formal and informal economy in Kenya.

Data Source: The Republic of Kenya Reports (Various years)

Comparing both the shares of the formal and informal economy, the share of the formal sector has been shrinking while the share of the informal sector has been growing. Figure 1.5 displays the trends in the shares of formal and informal employment in Kenya from 1985 to 2018.

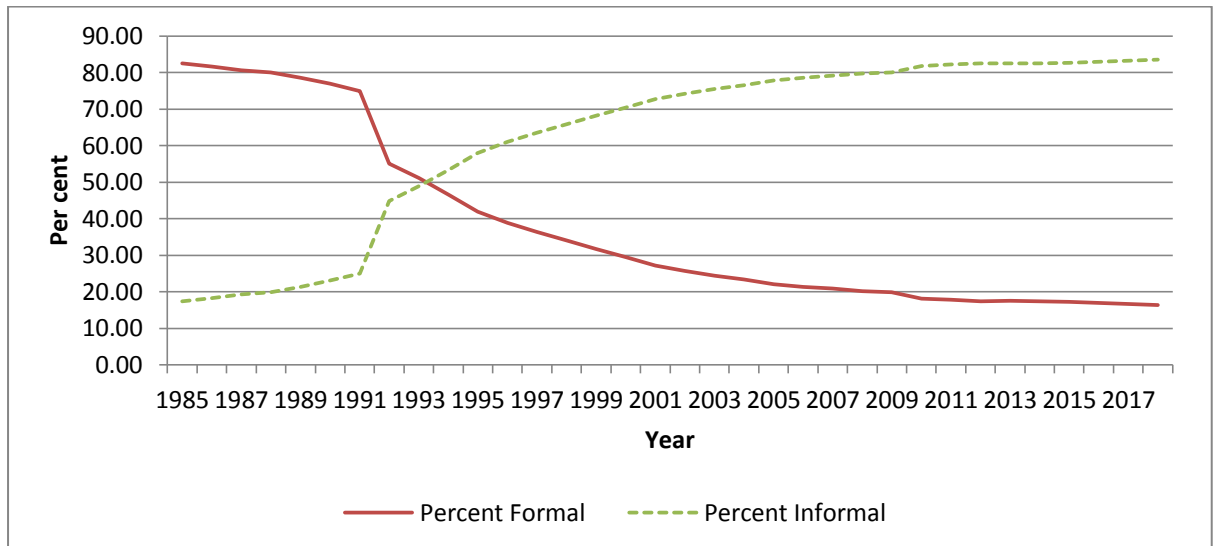


Figure 1.5: Trends in proportion of both formal and informal sector in the total employment in Kenya.

Data Source: The Republic of Kenya Reports (Various years)

Figure 1.5 illustrates that formal employment declined consistently between 1985 and 1991. This is attributed to structural adjustment programs and the Budget Rationalizations Program which called for a reduction in public employment (Republic of Kenya, 1989).

The expansion of the informal sector surpassed the growth in the formal sector in 1993. The increase of the informal sector is largely attributed to increased population, increased urban migration, poor business environment (high start-up cost and high taxation rates), and strict labour regulations. In addition, there is no effective coordination among the institutions responsible for business regulations. This has destabilized the ability to compel these complex reforms (World Bank, 2016).

Figure 1.6 illustrates that the number of new jobs created in the formal and informal economy has been increasingly growing since 1985. The highest number of new jobs created was 1.04 million in 2010 while the lowest was 37100 in 1989.

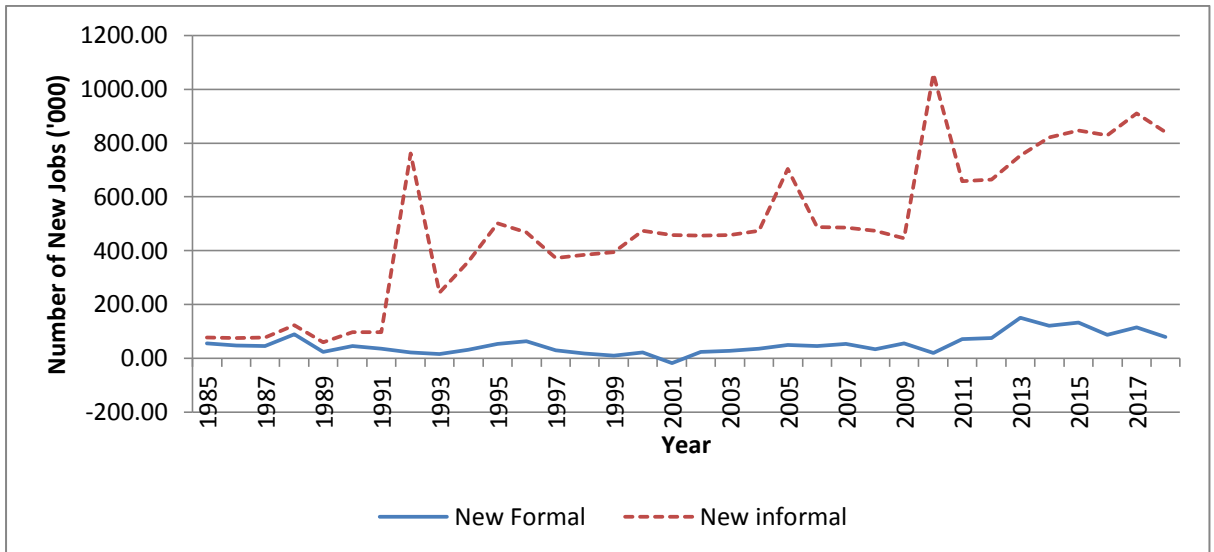


Figure 1.6: Trends in the number of new jobs in both the formal and informal economy in Kenya.

Data Source: The Republic of Kenya Reports (Various Years)

Figure 1.7 compares the growth rate in employment in the formal and the informal sector.

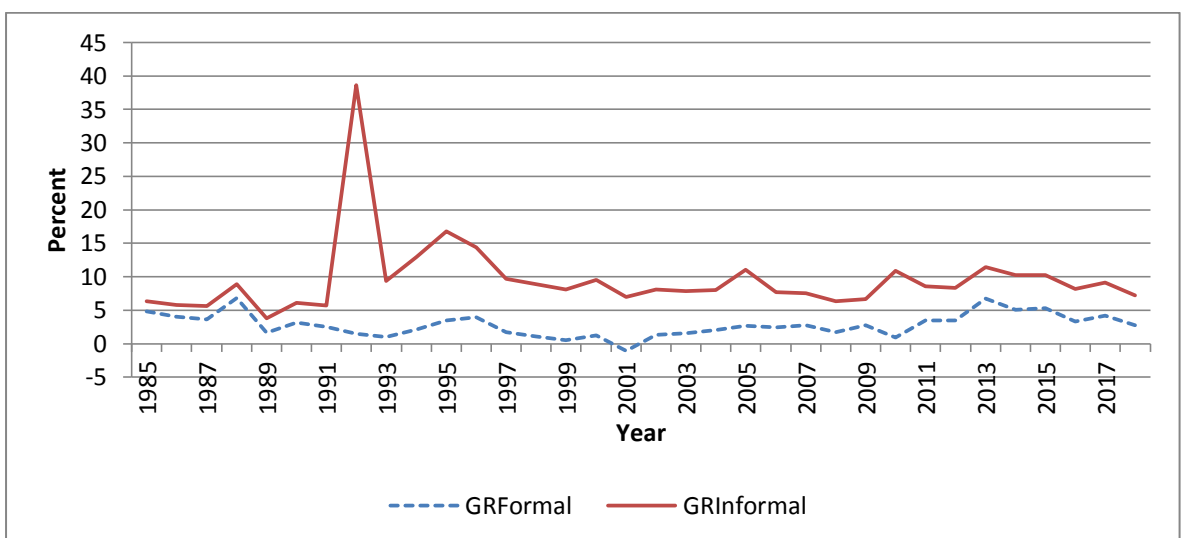


Figure 1.7: Trends in employment growth rate in formal and informal economy in Kenya.

Data Source: The Republic of Kenya Reports (Various Years)

Figure 1.7 demonstrates that since 1985, year-to-year employment growth rate in the informal economy has been higher compared to that of the formal. The growth was highest in 1992 where it peaked at 37.2 percent and can be attributed to modern wage employment becoming scarce, use of cheap technologies, less skilled labour, and little start-up capital (the Republic of Kenya, 1993).

1.1.4 Measures undertaken to encourage formalization in Kenya

Over the years, the Kenyan government has postulated measures to enable enterprise formalization and hence contract the portion of the informal economy. These measures have aimed at making business registration cheaper, easier, and more attractive; rationalizing business licenses, and creating a good business environment. To improve the regulatory framework, several bills and acts have been proposed that are geared towards making the economy more business-friendly. Some of the most recent ones are the Companies Act of 2015 that simplified the registration and running of a company; the Business Registration Service act of 2015 that ensured efficient administration of the incorporation, registration, operation, and management of businesses laws; and the Finance Act of amendments of 2015 (Africa Law Business, 2016). Also, there have been efforts towards improving registration and incorporation, among them being the introduction of e-registry to electronically access licenses and Huduma centers in the major cities and towns to reduce the time taken during registration (Abdalla, Kiragu and Waswa, 2015; Wanjiru, 2018). In addition, the government has introduced the *Kazi Kwa Vijana* initiative which has ensured the youths are getting formal jobs.

However, despite these efforts, the portion of the informal economy is growing enormously. The measures instituted seem not to have borne fruits in regards to

decreasing the magnitude of the informal economy. For instance, according to various Kenya Economic survey reports, the informal economy has been rising over time. For instance, in 2013 the share of those working in the informal economy was 82.49 percent. This figure has continued to rise steadily to 82.73, 82.95, 83.21, 83.45, and 83.6 in 2014, 2015, 2016, 2017, and 2018 respectively (the Republic of Kenya, 2019). Several factors can explain why informality in Kenya has continued to rise but the one most common factor in literature is the increased costs incurred in conducting or doing activities in the formal sector. Some of these start-up costs include high taxes; costly and time-consuming registration process; complicated rules and regulations and burdensome bureaucratic hurdles (Johnson, Kaufmann, and Shleifer, 1997). This study, therefore, investigates the effects of the cost of doing business on the growth of the formal and informal sector in Kenya.

1.1.5 Overview of Cost of doing business in Kenya

From the outgoing section, it is evident that the Kenyan informal economy is large, dynamic, and increasingly growing. Several factors can explain why the informal sector in Kenya has continued to rise but the one most common factor in literature is the costs of doing business under the formal sector. These costs include high taxes, a time-consuming registration process, and bureaucratic hurdles. In addition, other factors that may cause firms to operate informally are low levels of human capital, inadequate monitoring, and enforcement, poor institutional quality, limited access to credit, among others. Kaufmann, McMillan, and Woodruff (2000) and Enste and Schneider (2000) attribute the occurrence of the sector to costly government regulations and taxation administration.

World Bank Group (2018), in its Doing Business Report, identifies 11 indicators that determine the ease of doing business in Kenya. Kenya was ranked at position 80 out of 190 in ease of doing business and scored 65.15 out of 100 in the distance to the frontier tool. The First indicator is starting a business which World Bank Group describes as the procedures, costs, time, and minimum capital necessary to start a company. Kenya ranked poorly at 117 out of 190 and scored 83.20 out of 100 on starting a business. This indicates that Kenya still lags the in the ease of starting a business (World Bank Group, 2018).

The second indicator is dealing with construction permits which World Bank Group describes as the procedures, cost, and time to comply with all legal requirements necessary for warehouse construction, safety mechanisms, and quality control in the construction permitting system. Kenya ranked at 124 and scored 63.16 on dealing with construction permits. The third one is getting electricity which World Bank Group describes as the procedures, time, and cost required to connect with electricity, reliability of the electricity supply, and tariffs transparency. Kenya ranked at 71 and scored 76.68 on getting electricity. The fourth one is registering property which World Bank Group describes as the procedures, cost, and time required to transfer a property. It further elaborates on the land administration system quality. Kenya ranked at 125 and scored 54.9 on the registering property The fifth one paying tax which World Bank Group describes as payments charged, time spent, and total tax required for a business to be compliant with tax regulations and as well as before filing processes. Kenya ranked at 91 and scored 71.7 on paying taxes (World Bank Group, 2018).

Other indicators include safeguarding small investors, gaining access to finance, conducting international trade, upholding contracts, and resolving insolvency. Kenya scored 75.00, 58.33, 67.63, 58.27, and 43.11 in these categories, placing it at positions 29, 62, 90, and 95 overall. The eleventh indicator is labour market regulation (World Bank Group, 2018). This indicates that the cost of doing business (start-up costs in businesses, construction permits, paying taxes, access to electricity, property registration among others) in Kenya is still high. This study, therefore, investigates the effects of the cost of doing business on the expansion of the informal sector within Kenya.

1.2 Statement of the Problem

The informal economy in Kenya is large, dynamic, and is constantly growing. Informal enterprises form the largest portion of all firms in Kenya. Approximately, 83 percent of the total employment in the country is absorbed by the informal sector (the Republic of Kenya, 2019). Therefore, the informal sector is considered a vital contributor to economic development in Kenya as a source of employment, earnings, and household livelihoods. Despite its benefits, a large informal sector is harmful to the economy. The sector is associated with low productivity and indecent lives, low tax collection, job insecurity, lack of social protection (like insurance protection), among others. Thus, it is the interest of any government to promote the growth of the formal sector and thus encourage the formalization of the informal sector to improve the livelihoods of its people.

The Kenyan government has over the years proposed regulatory frameworks, several bills, and acts to enable enterprise formalization. These measures have aimed to make business registration cheaper, easier, and more attractive; rationalize business

licenses, and create a good business environment. For example, the Companies Act of 2015, the Business Registration Service in the year 2015, and Finance Act amendments of 2015 (Africa Law Business, 2016). Also, there have been efforts towards improving licensing and incorporation efficiency through e-registry and Huduma centers (Abdalla, Kiragu, and Waswa, 2015; Wanjiru, 2018). In addition, the government has introduced the *Kazi Kwa Vijana* initiative which has ensured the youths are getting formal jobs.

However, despite these efforts, the informal sector size is growing enormously. In 2013, the proportion of employed individuals in the informal sector was 82.49 percent. This figure has continued to grow rapidly to 82.73, 82.95, 83.21, 83.45, and 83.6 in 2014, 2015, 2016, 2017, and 2018 respectively (the Republic of Kenya, 2019). Several factors can explain why the informal sector in Kenya has continued to rise but the one most discussed factor in the literature, is the costs of doing business. As stated by the World Bank Group (2018), Kenya's cost of doing business (starting a business, construction permits, paying taxes, access to electricity, property registration among others) is still high compared to other countries.

Substantial amount of literature has only explored the determinants of informality (Djankov, Simeon, and Rafael La Porta, 2002; Alm and Torgler, 2006; Loayza, 2007; Taymaz, 2009). In addition, to the familiarity of this research paper, most of the studies carried out in Kenya concentrated on the causes, determinants, and the effect of the informal sector and do not investigate the effect of cost of doing business on the growth of the informal sector (World Bank 2013; Ndaka, 2017). Ndaka, 2017 found that avoiding tax, poor tax administration, and lack of inclusion of the informal sector in the national development as the pivotal factors that leads to the growth of the

informal sector. The study also found that Kenya is losing approximately 40 percent of revenue through this informal sector. To comprehend why the informal sector is still growing, the effects of the cost of doing business on the informal sector growth in Kenya need to be determined. Therefore, this research work aimed to seal this gap.

1.3 Research Questions

- i. What is the sectoral distribution of the formal and informal enterprises in Kenya?
- ii. What is the effect of the high cost of doing business on the probability of a business being in the formal or informal sector in Kenya?

1.4 Objectives of the Study

The main objective of this research was to evaluate the effects of the cost of doing business on the growth of the formal sector in Kenya. The objectives of this research were to:

- i. Find out the sectoral distribution of formal and informal enterprises in Kenya.
- ii. Investigate the effect of the high cost of doing business on the probability of a business being in the formal or informal sector in Kenya

1.5 Significance of the Study

The results of this study will add to previous studies on the influence of business costs on the growth of the formal and informal sectors. The results will also be used by the policy makers and government to provide actual data for the informal sector discussion in Kenya. Finally, it will add to the body of knowledge, data, and information already available about the unorganized sector in Kenya.

1.6 Scope of the Study

The effects of business costs on Kenya's growing formal and informal sectors were the main subject of this study. The study made use of 2016 KNBS data on Micro, Small, and Medium Enterprises. The survey covered the whole country (all 47 counties) and contained a sample of 17, 895 individuals. The decision to use the data was informed by the most recent data, which fully covers the cost of doing business for both official and informal firms.

1.7 Organization of the Study

The next chapters of this research project are structured as follows: Chapter two provides an overview of the theoretical and empirical literature, and chapter three provides a description of the technique applied to achieve the goals of this research article. The research design, model description, theoretical framework, data analysis, and measurement and definition of variables are all included in Chapter three. Chapter four presents the empirical findings and chapter five gives an analysis and conclusion of the study.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The chapter commences by analysing the theoretical literature of informality. The chapter then continues to present the empirical literature on the causes of informality. It concludes by providing the literature summary of the study.

2.2 Theoretical Literature Review

This part reviews the theories of informality. The theories that explain informality are Modernization Theory, Dependency Theory, Structuralism Theory, Neo-liberalism Theory & Institutional Theory.

2.2.1 Modernization Theory

The theory was formulated in early in the 20th century by Talcott Parsons and developed by Rostow (1960) in the theory of national development. The theory had an impact on informality studies in the 20th century. According to this, the informal sector is rapidly disappearing while the modern official sector is rapidly growing. According to Rostow, each nation's outdated social-economic institutions are to blame for the Third World countries' underdevelopment (and by extension, informality). Additionally, he made a policy recommendation that these nations should adopt contemporary norms in order to develop contemporary governmental structures, legal systems, and capitalist economies. Rostow defined modernity as having Western institutions, values, and economy. According to modernization theory, these isolated rural people commonly use pre-capitalist, traditional subsistence methods in the informal sector.

The theorist argued that the participants of the informal sector remained outside the legal framework since they lacked proper education skills and values. Ironically, despite what it connotes, that the informal sector will be absorbed into the formal sector through the growth of industrialization, it is apparent that the informal sector continues to persist and grow especially in developing countries. This theory was criticized by the Hart, (1973) and ILO (1972) through wide and all-inclusive research works that were conducted on Ghana's informal economic activities. These studies came to the conclusion that using the informal sector as a substitute for the reserve army was not the best strategy to promote development and economic growth. This idea also had the drawback of viewing the informal sector as a problem that needed to be fixed rather than as a potential source of economic growth, a talent pool for entrepreneurs, or a development strategy that could be strengthened and supported (Yusuf, 2011). Therefore, this ideology was not appropriate for this study, and thus the propositions from this theory were not used in the current study.

2.2.2 Dependency Theory

This theory originates from Hans Singer and Raúl Prebisch (1949). The theory was further advanced in the informal sector by the pioneering work of Keith Hart among the informal in Ghana in 1973, in conjunction with International Labour Organization (ILO) in 1972 on informality in Kenya when they coined unregulated economic activity as the informal sector. In his research, Hart (1973) disproved the idea that traditional subsistence practices were simply expanded in the informal sector. All those who engage in these illegal activities were destined for marginalization and poverty (Hart, 1973, Portes & Schauffler, 1992). However, some academics who adhered to the conventional dependency paradigm described the informal sector as being poorer and more marginalized than the modern capitalist sector (Schauffler

1993). In addition, scholars of the dependency approach (Tokman,1978; Tokman 1989;) frequently detailed the characteristics of informal sector enterprises as little use technology, use of unpaid family labour, low production, low capital, little profits margins, minimal productivity and competition, and less entry barriers. Further, the dependency approach considers informal sector operations aimed at survival and not profit maximization. Furthermore, the dependency approach often coined that the informal firms were deliberately capitalizing on their ability to avoid taxes and regulations. The theory limitation is that that it views the informal labour operations being an outsider of the exploitative formal relations of production hence it was considered as a substitution of the rural subsistence sector into the urban environment. The theory fails to point out the contribution of high cost in the emergence of informal firms. Because of these shortcomings, the ideologies from this theory were inappropriate for the current study and were not incorporated.

2.2.3 Structuralism Theory

Structuralism was opined by Claude Levi Strauss and Roland Barthes in 1950. Structuralism insists that informality is an alternative form of labour supply, not merely as a result of surplus labour supply or state over-regulation. According to Maloney (2004), informality is not in existence because of some accident or flaws in capitalist development but it exists because of peripheral capitalist accumulation logic. According to Rakowski (1994), informality was non-existent in the 19th century because labour relations protections and benefits were paltry or non-existent. In addition, informality came into existence after a long struggle of the labour unions in their efforts to ensure relatively protected workers in the production process. Structuralism made two significance contributions to the conversation on the informal economy. The informal sector of the economy was initially utilized to help the

capitalist system maintain market competitiveness by lowering costs, particularly labour costs Castells and Portes (1989). As a result, there are many reasons why the informal economy is expanding and why it is essential to growth. These include the fact that the informal economy employs a significantly larger number of people than the formal economy does due to the former's higher income and the latter's rising need for affordable goods and services. Additionally, the capitalist uses informal labour because it is less expensive than formal labour. Globalization in developing countries has contributed to the growing demand for formal labour due to the outsourcing of the informal economy. The theory's limitation is that even if industrial informal labour outsourcing is a major feature in Latin American cities, it is an irrelevant feature in developing countries particularly in Africa where informal activities are for subsistence purposes (Capecchi, 1989). Although this theory was not used as a basis for the current study, it does identify costs of labor and of goods and services as important determinants of informal sector. Thus, the cost variable was captured in this research work.

2.2.4 Institutional theory

The Institution theory was suggested by Williamson (1985) and was further advanced by North (1990). The theory describes institutions as the rubrics which control and shape the conduct. Both formal (regulations) and informal (norms, values, and beliefs) institutions exist in civilizations (Denzau & North, 1994; Helmke & Levitsky, 2004). According to Welter and Pobol (2015) and Williams and Vorley (2015), formal activities take place within formal institutions, informal activities take place within informal institutions; and criminal activity takes place outside of both formal and informal institutions (McElwee & Smith, 2014).

The theory suggests that the institutional environment (macro and microenvironment) determines both individuals' and firms' economic behavior and choices which impact the behaviour and choices of entrepreneurs within the economy. The theory argues that weak formal institutions create higher institutional voids in the emerging economies (developing countries) which lead to a substantial amount of businesses activities in these countries to be undertaken in the informal sector (North, 1990). The theory also argues that the higher the level of incompatibility between formal institutions and informal institutions, the high the level of informality. This theory presumes that institutions play key roles in informality but does not describe how institutions affect the cost of doing business hence leading to the increase in informal institutions. Institutional theory is therefore not appropriate for this research. Therefore, the current study was not based on it.

2.2.5 Neo-liberalism Theory

This theory was advanced into the informal sector by Peruvian economist Hernando De Soto (1989). Neoliberalism ideology focuses on economic liberalism. It emphasizes the economic policies seeking to minimize the state's role and maximize the role of the private business sector. Neo-Liberalism aims to shift the direction of the economy from the public to the private sector with the hope that, it will increase government efficiency and result in the growth of the nation's economy.

De Soto (1989) argues that too much government interference and regulations forces the entrepreneurs to leave the formal economy into the informal economy. In addition, he claims that there are many business owners operating formally in the informal sector since it would be too expensive, time-consuming, and difficult to comply with the regulatory requirements placed on an organization. Additionally, he pointed out

that businesses will continue to function clandestinely as long as formal registration processes are time-consuming and expensive, there are no property rights, and it is challenging to get access to capital and technological resources. According to the author, cumbersome rules and regulations which are unreasonable are stifling the enterprises. Furthermore, enterprises fail to register because of over-regulation, too much state interference and high taxes. The current study adopted this study as it concentrates on the cost of doing business in the informal sector and this theory argues that increased cost of doing business (high taxes, time taken to register a business, lack of access to finance, and cumbersome government rules and regulations) has brought about the growth of the informal sector. Therefore, this study borrows the costs variables (as mentioned in the previous sentence) from this theory.

2.2.6 Discrete Choice Random Utility Model

This theory was developed by McFadden and Domencic in 1975. The theory is used to analyze the behavior of individuals when they are making choices between two or several competing alternatives. The theory assumes that individuals are utility maximizers, and that they are rational, that is, they always chose the option that maximizes their utility (welfare). The choice the individual makes is assumed to depend on the characteristics of the alternatives, features of the individual and a random error term. The theory has been widely used in modelling choices, not only in economics, but also in marketing, planning, transportation and health, among others.

In the current context, the theory assumes that the decision to operate in the informal or formal sector is a function of attributes like the operational cost in each sector, and the characteristics of the entrepreneur. Some of the operational costs that the entrepreneur considers are registration, electricity, wages, rent and other compliance costs. The attributes of the entrepreneur that may affect entrepreneur's decisions are

education, age, gender and experience. If the expected utility from operating in the formal sector is higher than the informal, the entrepreneur operates in it, otherwise he operates in the latter. Some of the authors who have used this model include Moyo (2022) who analyzed the factors that affect the formalization of businesses in Sub-Saharan Africa using a sample of 13 countries. Since this theory rhymed with the current study, it was used in conjunction with the Neo-liberalism theory to analyze the probability of a business operating in the formal or informal sector in Kenya.

2.3 Empirical Literature

Extensive research works have been done in this part to explore the association between the cost of doing business and informality. Thereafter, the evaluation of empirical literature on the effect of the cost of doing business on informality is conducted.

2.3.1 Global Empirical Literature

The Haitian Centre for Free Enterprise and Democracy (CFED) and the Institute for Liberty and Democracy (ILD) (2001) carried out a survey to explain the Haitian extra-legal economy and propose a procedure of legislative and institutional reforms to increase the economic value of dormant extra-legal resources. Some variables applied were financial accessibility, tax, infrastructure, technology, cost of registration, and time of registration. The study used descriptive statistics to analyse the data. The study found that roughly 68 percent of all businesses in Haiti operate extra legally. The study found that inadequate limited liability business form for small businesses, complex tax procedures, and complex legal procedures related to labour and entitlements were the main legislative causes for informality (extra legality). The study also found that lack of efficient integration between government institutions and business registration processes increased the time and cost of registration and in as a

result increased the rate of informality. Moreover, the study also established that the burdensome process of renewing a business license was costly and reduced the annual income of a business by five percent hence contributing to the increase of informality. The study's limitation is that it used descriptive statistics instead of an econometric model. The use of descriptive statistics may not generate strong inferences.

Djankov, Simeon, Rafael La Porta, al. (2002) conducted a study to estimate the entry regulation of start-up businesses in 85 countries. The information on entry regulations was gathered from official government documents, reports from development organizations (such the World Bank and USAID), and official government websites. The factors that were of interest were the formal duration, formal volume, and formal expense necessary for a new business to operate legally. The study used the help hand model, the capture model, and the tollbooth models. The study found that strictness of government regulations to business entry was positively correlated to the magnitude of the informal sector. Particularly, informality was high in countries where the number of procedures required to start a business was large and in countries where time and cost to register a start-up were high. The outcomes of the study also established that there was a positive correlation between the severity of government regulations for business entry and corruption. The limitation of this study is that the data used underestimated the cost and complexity associated with the business entry.

The Institute for Liberty and Democracy (ILD) (2004) performed a survey in Tanzania to determine how the Tanzania extra-legal economy works and how the legal body interrelates with the extra-legal economy. The research used descriptive statistics to analyse the data. The study established that despite the legal framework put by the government to ensure business's participation in the legal economy; the

extra-legal economy rate was high with 98% of all businesses operating extralegally. The study also illustrated that the content of legislative and regulatory requirements was positively related to the extra-legal economy. In addition, the study noted that the cost of compliance with legal requirements highly funded to the growing of the extra-legal economy. The study is limited because it used descriptive statistics instead of an econometric model, and thus no econometric variables were used, apart from the measures of the size of the informal sector. The use of descriptive statistics may not generate strong inferences.

Almand Torgler (2006) did a survey to explore what determines an individual's intrinsic willingness to pay taxes focusing on the United States and Europe. The variables used in the study include measures of tax policy, tax compliance, tax revenue, tax morale and culture measures. The study used simple correlation to estimate several years' data from World Values Survey for various countries. The study discovered a substantial inverse association between the shadow economy and tax compliance (tax morale). That is, a rise in tax revenue will probably result in a decline in the size of the shadow economy. Therefore, the level of tax morale influences real behaviour and may have a noteworthy effect on the size of the shadow economy. The obvious problem with this study is that it only looked at a simple correlation between the variables. This lessens the validity of the study and is overly simplistic.

Seira, Kaplan and Piedra (2007) estimated the effect of the System of Fast-Opening of Firms (SFOF) on business start-ups in Mexico. The study used survey data from three sources: the Mexican Institute of Statistics, Geography, and Informatics (INEGI); the Federal Commission for Regulatory Improvement (COFEMER), and the Mexican

Social Security Institute (IMSS). The key variables used were population; the monthly number of new jobs created by new firms; monthly new firms; non-existing firms; the number of establishments; production and Employment. The study applied Discrete-Time Duration Models and the Cox duration model. System of Fast-Opening of Firms program reduced the time of registration to a maximum of 48 hours. The study found that SFOF caused the number of start-ups registered to increase by 4 percent. The study also found that the number of businesses registered reduced in the municipalities that had a high rate of corruption through bribes because the business wanted to avoid excessive bribes. In addition, the study found that businesses failed to register to evade taxes. Furthermore, the study found that in the municipalities where additional registration procedures were cheaper the number of businesses registered was high than those with expensive additional registration procedures. The study did not include 16 competitive municipalities in Mexico hence the data sampling was not well distributed and did not represent the whole country.

Loayza (2007) conducted a cross-country study in Peru to determine the causes and consequences of informality. The variables used were the Schneider index, self-employment, Heritage Foundation Informal Market index, lack of pension, business regulatory freedom, law, and order, average years of secondary schooling, and socio-demographic features (Population and age). The study applied the Multiple Indicator-Multiple Cause (MIMIC) model and regression analysis to accomplish its objectives. The study used various measures of informality, particularly the study used absence of a pension plan, self-employment, and the Multiple Indicator-Multiple Cause (MIMIC) measure by Schneider (2004) for comparison purposes. The study found that the measures of informality are highly correlated with a weak monitoring and enforcement capacity by the government, extreme regulatory burden, poor public

services, negligible level of education, high-level socio-demographic factors such as the high number of youths, a high number of the rural population and a larger share of agriculture in GDP. The limitation of this study is that it ignored some aspects of informality. These aspects are the size of firms, labor-market segmentation, and capital-labour intensity.

Ingram, Vijaya, and Desai (2007) conducted an analysis to determine why businesses select to be informal or formal in Africa. The study used Enterprise Surveys' data for Uganda, Kenya, Tanzania, South Africa, Zambia, and Senegal. The Enterprise Surveys' data contained both formal and informal businesses records which made it easier to compare the data. The study compared the results achieved regarding investment climate quality for both formal and informal businesses. Specifically, businesses were requested to choose their top three aspects from a collection of 18 aspects, which hinder their performance. These included access to finance, tax, availability of electricity, access to land, corruption, among others. The study then analysed how both formal and informal businesses ranked different aspects of investment climate and estimated how each of aspect affected the probability of a business being informal. The study found that the probability of a business being formal is positively linked with access to financial assistance, the availability of electricity supply, and access to land. Additionally, the study found that the probability of a business being formal negatively linked with tax rate and corruption. The limitation of the study is that firms were only asked to rank the top three constraints, yet firms face many constraints. The use of many responses would have improved the study.

Wahba (2009) estimated the probability of transiting from the informal and semi-formal sectors to the formal sector in Egypt. The study used data from the Egyptian Labour Market Panel Survey 2006. The author applied the probity model to achieve the study's objective. Variables used were education, gender, age, non-agriculture employment, non-agricultural waged employment among others. The study found that the probability of moving from the informal and semi-formal or formal sector is highly segmented along with education and gender in Egypt. The study found that the transition exists only for educated male workers but not for uneducated or female workers. The study effects ignored the unemployed population which could be misleading.

Taymaz (2009) conducted a study to investigate the reasons why informal and formal businesses different productivity in Turkey. The study used the World Bank survey dataset of 2008 where 1000 firms were surveyed. Using the logit model, the study sort to find out the dynamics that affect the choice of informality. The variables used were the age of the business, education, gender, size of the business (determined through the number of employees), experience (determined as the number of years a person has stayed in the same job), and trust in the legal and business institutions. The study found that the size of the firm is negatively related to informality. The larger the firm size the higher the probability to perform formally. The study observed that there was a higher probability of older firms operating informally. Additionally, the analysis further established that the education level and informality are negatively related. The firms with managers who have at least a university diploma were less likely to operate informally. The study was limited because the authors lacked a large panel data hence; they were unable to control the unobserved effects that could be removed by a difference in difference estimator.

Ali (2012) led a survey to explore the effect of the informal sector on the output of micro and small enterprises (M/SEs) in Egypt and to examine the link between the productivity of M/SEs, access to finance, and informality in order to discover the best and the most effective formalization mechanism. The study conducted empirical research grounded on the instrumental variable model. The study used private micro and small enterprises information acquired from the Economic Research Forum. The variables used were labour productivity, total factor productivity, gender discrimination, financial accessibility, tax, infrastructure, technology, cost of registration, and time of registration. The study found that the M/SEs that operated informally had significantly slow productivity. The study also found that the positive effect on which registration had on productivity hugely depended on willingness to register. The study asserted that willingness to register a business was highly dependent on the strictness of the taxation processes; access to good infrastructure, access to training, access to technology, access to labour capital and financial capital. The study suggested that creating a business environment that is sustainable enables efficiency in registration, facilitate good technology and infrastructure, ensure flexible regulations, and enable access to better funding. The study is limited because it did not consider medium-sized enterprises which form a significant part of the informal sector and this could result in biases.

Rothenberg, Gaduh and Burger. (2015) carried out a study to investigate the theory that best explains the persistence of informality in Indonesia. The study used nationally representative survey data on micro, small, and medium-sized firms from Indonesia's National Statistical Agency in 2013 and World Bank Enterprise Survey in 2009. The variables considered included skills, tax, wages, transportation, capital, raw materials, and others. The study used a cost-benefit model using regression analysis

methods to achieve its objectives. The study revealed that many businesses in Indonesia fail to register as they have no aim to expand or seek formal financial services or because they are avoiding tax. The study also found that the cost of registration plays a major role in informality. However, the study found that the steps taken by the Indonesian government to reduce registration costs did not affect the informality rate. The study used simple statistical analysis which may not generate strong inferences.

2.3.2 Empirical literature from Kenya

World Bank (2013) conducted a survey to examine the main impediments faced by informal enterprises, ascertain the trends of productivity and firm dynamics, and better comprehend elements of formalization in Kenyan the cities of Nairobi, Mombasa, and Nakuru, and in the Central and Nyanza region. The survey also investigated why enterprises remain in the informal sector. The variables used were; the size of the firm (in terms of employment and sale size), access to credit, access to electricity, access to water, corruption (bribes), paying tax (time and tax rate), registering requirements (time, fees and procedure), bureaucratic procedure among others. The study applied a summary statistic model. The survey showed that the cost of registration and taxes that a formal business pays were the common reasons businesses opt not to register. Also, the study found that businesses choose to formalize in order to escape the difficulties they face during accessing finance, electricity, water, and other public services, and escape dealing with the corruption and harassment by public officers. If an enterprise does not face these difficulties, then the willingness to register is very low. In addition, the study found that enterprises' willingness to register varied with regions with Nyanza having the highest number of enterprises willing to register and Central having the lowest. The

study was limited because it focused on a few regions in Kenya and the sample was thus not representative enough.

Ndaka (2017) conducted research to explore the causes and effects of the informal sector on revenue collection in Kenya. Secondary data was utilized to attain the survey's objectives. The study used a meta-analysis model to analyse the data. The study found that the following variables were the principal factors that leads to the growth of the informal sector in Kenya: avoiding tax, poor tax administration, and lack of inclusion of the informal sector in the national development. The study also found that Kenya loses approximately 40 percent of revenue through this informal sector. This study is limited because it relied on other studies' findings to accomplish its objectives. In addition, the paper did not specify its own econometric model; the paper just provided a discussion of the informal sector without conducting any analysis.

2.4 Overview of the Literature

The modernization theory assumes that informality is a result of the backward social-economic systems of individual countries and ignores other determinants of informality. Institutional Theory explains how institutions affect an entrepreneur's decisions of being formal or informal. Dependency theory considers the informal sector as an outsider of the exploitative formal relations of production hence it was considered as substitution of the rural subsistence sector into the urban environment. Structuralism theory insists of informality being is another source of labor supply and not as a result of surplus labour supply or state over-regulation. These theories explain why businesses opt to exit the formal sector or why businesses are excluded in the formal sector. However, only the Neo-liberalism theory explicitly links the cost of

doing business and the rate of informality. This study uses Neo-liberalism theory due to the fact that the current study assumes that increases in the cost of doing business have led to the growth of the informal sector in Kenya.

The majority of the empirical literature discussed focused on the determinants of informality. Majority of researches reviewed agree that the cost of doing business affects the rate of informality (for instance, Djankov, Simeon and Rafael La Porta, 2002; Alm and Torgler, 2006; Ali, 2012; Rothenberg, Gaduh and Burger, 2015; World Bank's Informal Enterprise Surveys (IFS), 2013). Most of the studies used survey data (Seira, Kaplan & Piedra, 2007; Ingram, Vijaya and Desai, 2007; Wahba, 2009; Taymaz, 2009). Some of the cost of doing business variables used in these studies were taxes, cost of registration, and time of registration, and financial accessibility (Alm and Torgler, 2006; Ali, 2012; Rothenberg, Gaduh and Burger, 2015; World Bank 2013).

Other variables that were found to affect the formality status, the education level, experience, age of the business, size of the business, labour productivity, total factor productivity, gender discrimination, infrastructure, and technology, among others. Some studies used simple descriptive statistics (ILD and CLED, 2001; ILD, 2005) while others used probity and legit models (Wahba, 2009; Taymaz, 2009) and others used multivariate regression model and Loayza (2007) used Multiple Indicator-Multiple Cause (MIMIC) models. None of the studies reviewed above investigates the effects of the cost of doing business on informality in Kenya. This study, therefore, investigates the effects of the cost of doing business on informality in Kenya using MSME survey data of 2016.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter discusses the research design of the study, the econometric model, theoretical framework, data analysis techniques, and data sources used to accomplish the objectives of the study. In addition, model diagnostics techniques were proposed to ensure that the results achieved were robust.

3.2 Research Design

The research embraced an exploratory research design which is described as investigation into a theoretical or hypothetical idea. Exploratory research design refers to incidence where an investigator has detected a phenomenon and pursues to have a clearer understanding of it (Reiter, 2017). The objective of the exploratory research is to demonstrate the cause-and-effect links among variables and therefore, exploratory research design is appropriate for this research.

3.3 Theoretical Framework

This research work is founded on the both the discrete choice random utility model and the Neo-liberalism theory. It argues that enterprises are excluded or exits from the formal economy due to costs, time, and efforts involved in confronting bureaucratic burdens imposed by the state (De Soto, 1989). Therefore, the theory explains the cost of doing business as a cause of the informal sector. The analysis presents the choice of being in the informal sector as a choice: one's choice of the formal or informal sector is revealed by the sector the one belongs to. Several authors like Hosmer and Lemeshow, 2000; Agresti, 2002; Green, 2003 and Aikaeli and Mkenda, 2014 have

modelled the decision to be in the formal or informal sector as a discrete choice model.

A discrete choice model models the individual's choice to belong to one of the categories (in this case formal or informal sector). The individual is assumed to maximize their utility. Every individual n faces 2 alternatives: be formal or informal. The individual makes a choice to be in either of these if the utility of one exceeds the other (in terms of expected profits). The utility acquired from either choice is unobservable while the choice's features are observable. The choice selected by the individual explains which sector provides the best utility. The observed choice equals one if the individual chooses the formal sector and zero if otherwise. Equation (3.1) relates the individual choice, with the characteristic that makes them choose to operate in the formal or informal sector.

There is $n(n = 1, 2, \dots, N)$ individuals who each face $m(m = 1, 2, \dots, M)$ alternatives. They each make choice i if the utility acquired from selected choice ($u(i)$) is better than $u(j)$. Although the utility acquired from each choice is unobservable, the features associated with each individual choice are observable. Therefore, let x_{nj} represent the observable features of an individual n and y_{nj} represent his unobservable features. In addition, let $u(nj)$ represent the utility individual n acquires from selecting choice j while u_j represent the utility off selecting choice j . The utility derived from choice j is represented by the following model:

$$u(nj) = u_j(x_{nj}, y_{nj}) \quad (3.1)$$

When confronted with either two or more choices, the individual will select the choice that have the maximum utility. Therefore, the selected choice reveals the option that provides the individual the best utility, although the utilities are unobservable.

Therefore, the observed indicator will be one if the utility attained from that category is better than the utility attained from other categories. In this case, the observed indicator will be a binary choice indicating whether a person is either in the formal or informal sector.

The individual will choose alternative G (the informal sector) to alternative H (the formal sector) if:

$$U_G > U_H \text{ (Utility of } G \text{ is better than utility of } H \text{)} \quad (3.2)$$

$$V_G(x_{nj}, \gamma) + \varepsilon_{nG} > V_H(x_{nj}, \gamma) + \varepsilon_{nH} \quad (3.3)$$

Where,

$V_G(x_{nj}, \gamma) + \varepsilon_{nG} = U_G$ is the utility obtained from option G,

$V_H(x_{nj}, \gamma) + \varepsilon_{nH} = U_H$ is the utility obtained from option H,

γ is a vector of unidentified parameters and

ε_{nG} and ε_{nH} are error terms for option G and H respectively.

$$V_G(x_{nj}, \gamma) - V_H(x_{nj}, \gamma) + (\varepsilon_{nG} - \varepsilon_{nH}) > 0 \quad (3.4)$$

Let $g(x_n, \gamma) = V_G(x_{nj}, \gamma) - V_H(x_{nj}, \gamma)$ be a vector of observable characteristics and

$\mu_i = (\varepsilon_{nG} - \varepsilon_{nH})$ be a vector of unobservable characteristics,

Equation $V_G(x_{nj}, \gamma) - V_H(x_{nj}, \gamma) + (\varepsilon_{nG} - \varepsilon_{nH})$ can therefore be written as:

$$g(x_n, \gamma) + \mu_i \quad (3.5)$$

Letting h_i^* represent the latent variable capturing the formal or the informal sector,

equation (3.5) can be written as:

$$h_i^* = f(x_n, \gamma) + \mu_i \quad (3.6)$$

Where h_i^* is a binary variable which captures the informal or the formal sector,

x_n is a vector capturing both the workers and the environmental characteristics

γ is the coefficient vector and

μ_i is a stochastic error term.

Equation (3.1) can also be rewritten as:

$$h_i^* = \gamma_k X_i + \mu_i \quad (3.7)$$

Where γ_k is a row vector of coefficients and

X_i is a vector of explanatory variables.

A latent variable model is then used as it is easy to compare the binary models (logit and probit) and the natural log is taken to bound the probabilities between zero and one. The logit or the probit model can be used, the choice depends on the assumption made on the distribution of the error term. If the assumption of the error term is normally distributed, a probit model is used; otherwise, a logit model is adopted. The current study adopted the logit model.

$$\begin{aligned} \text{Logit}(E(E(Y_i|X_{1,i}, X_{2,i} \dots X_{m,i}))) &= \text{Logit}(p_i) = \text{Ln}\left(\frac{p_i}{1-p_i}\right) = \gamma_{0,k} + \gamma_{1,k}X_{1,i} + \\ &\gamma_{2,k}X_{2,i} + \dots + \gamma_{m,k}X_{m,i} + \mu_{m,i} = \gamma_k X_i + \mu_i \end{aligned} \quad (3.8)$$

The general model for the logit model is thus specified as:

$$pr(y_i = r) = \frac{e^{\gamma_r X_i}}{e^{\gamma_0 X_i} + e^{\gamma_1 X_i}} \text{ where } r = 0 \text{ (formal) and } 1 \text{ (informal)} \quad (3.9)$$

$$pr(y_i = r) = f(r, \gamma_0 X_i, \gamma_1 X_i) \quad (3.10)$$

From equation (3.6), the model can be re-specified as:

$$y = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \dots + \gamma_k X_i + u_i \quad (3.11)$$

Where y is a binary variable (formal or informal sector) and X_i represents the independent variable. The independent (control) variables are the firm and the entrepreneur characteristics. This is because these variables determine the probability that the entrepreneur is engaged in the formal or the informal sector.

3.4 Specification of the model

The empirical model applied in this research work was based on the theoretical literature and the empirical literature discussed. The study also took into the account the data availability in the Medium, Small, and micro enterprises survey data (2016) which is to be used in the current study. The variables affecting the entrepreneur's choice to be formal or informal were categorized into entrepreneur characteristics, business characteristics and business environment variables. The firm's characteristics include the market and ownership, size of business, and age of the business. The entrepreneur's characteristic includes age and gender. The business environment variables captured the different costs that were central to this study and they include the cost of electricity, cost of credit, and cost of licensing. The model is specified as follows:

$$y = \alpha_1 + \alpha_2 Elect + \alpha_3 Credit + \alpha_4 Licen + \alpha_5 Rent + \alpha_6 Taxes + \alpha_7 Gender + \alpha_8 Educ + \alpha_9 Age_bus + \alpha_{10} market + \alpha_{11} size + \alpha_{12} Ownership + u_i \quad (3.7)$$

y is a binary variable representing whether one is formal or informal,

$Elect$ is the cost of electricity,

$Credit$ is the cost of credit,

$Licen$ is the licensing cost,

$Rent$ is the total amount paid for renting a business premise,

$Taxes$ are the amount paid in taxes,

$Gender$ is the sex of the business owner,

$Educ$ is the highest education level attained by the business owner,

Age_bus is the age of the business,

$Market$ is the main market served by the business?

$Size$ is the size of the business and

Ownership is the nature of ownership of the business

3.5 Definition and Measurement of variables

Table 3.1 displays the definition and measurement of variables.

Independent (control) variables include; the firm, business, and entrepreneur characteristics. This is because these variables to determine the probability that the entrepreneur is engaged in the formal or informal sector.

Table 3.1: Definition and measurement of variables

Variables	Description	Measurement
Y	Whether a business is formal or informal.	It was measured as a binary variable: 1 if formal and 0 if informal.
Credit	Amount of money a business borrows from lenders	It was be measured in Ksh
Electricity	Source of energy and lighting for businesses in form of power connectivity	It was be measured in Ksh
License	The amount of money that a business pays annually to be able to operate	It was be measured in Ksh
Rent	The monthly fee that the business pays for using the business premises	It was be measured in Ksh
Taxes	Fees that the business pays to the government for operating	It was be measured in Ksh
Business age	A measure of the period of operation of the business.	It was measured in years
Market	The major customer of the business	It was be a variable dummy where 1 represent the business selling to other businesses and zero otherwise
Business size	A measure of how big a business is	It was is a categorical variable where 1 represents micro-business, 2 represents small businesses, and 3 represents medium businesses.
Ownership	Nature of the business ownership	It was be a dummy variable where

		1 represents sole proprietorship and 0 if otherwise.
Gender	The sex of the firm owner(s)	It was be a dummy variable where 1 represents male and zero otherwise
Education	the peak level of education of the firm owner(s) or managers	Education was be measured as a categorical variable where 0 represents those with no education, 1 represents primary, 2 represents secondary, and 3 represents post-secondary.

Source: author's illustrations

3.6 Data Type and Sources

This research work utilized secondary cross-sectional household survey data gathered by the Kenya National Bureau of Statistics (2016) named the Medium, Small and micro enterprises survey data of 2016. The sample contained 24,164 enterprises. Among this, 17,895 were informal while the rest (6269) were formal Only those aged 18 years and above were considered in this study. The survey targeted enterprises with at most 100 employees.

3.7 Data Analysis

The objective one of this research was to determine the sectoral distribution of formal and informal enterprises in Kenya. This objective was achieved by analysing the 16 sectors getting the distribution of the informal sectors in each sector. The fifteen sectors to be analyzed are presented in Table A.1 in the appendix section. The second objective investigated the effects of high cost doing business on the probability of a business being formal or informal in Kenya .The objective was to be answered using the logit model. To address these objectives equation (3.7) was estimated. The marginal effects of the estimated parameters were calculated and interpreted accordingly.

3.8 Diagnostics Tests

This research work conducted several diagnostic tests to ensure the results are accurate. These included multi collinearity test, the specification test and the Hosmer-Lemeshow goodness of fit test. The multi collinearity test was conducted to investigate whether the variables are highly correlated. The Hosmer-Lemeshow goodness of fit test measures the cause-and-effect link between the dependent variables and the independent variable. A specification test was conducted to investigate whether the specified model was correct.

CHAPTER FOUR

EMPIRICAL FINDINGS

4.1 Introduction

This chapter discusses the results of the research. Before the results were interpreted, several diagnostics were carried out to ascertain that the model was correct and devoid of misspecifications. These include classification tests, Hosmer Lemeshow goodness of fit test and the specifications test. The chapter then presents the descriptive statistics of the data, for both categorical and continuous variables. The results are displayed in line with the objectives posed in chapter one. The results are then compared with other studies that have been undertaken previously in the same area.

4.2 Descriptive Statistics

The study had 24164 observations. Table 4.1 gives the descriptive statistics of the categorical variables utilized in the research.

Table 4.1: Descriptive statistics of the categorical variables

Variable	Count	Share	Variable	Count	Share
Registration Status			Gender of owner		
No	17895	74.05	Male only	10880	46.08
Yes	6269	25.94	Female only	6061	25.67
Main market			Male-Female Partners	5555	23.53
Individual consumers	21002	86.91	Male-Male partners	797	3.38
MSMEs	1994	8.25	Female-Female Partners	320	1.36
Non-MSMEs	713	2.95	Education of the employees		
Others	258	1.06	College and above	6906	33.5
Government	141	0.58	Secondary	6380	30.95
Direct exports	56	0.23	Primary	5429	26.34
Ownership			None	1793	8.7
Individual	21763	91.64	Other (Specify)	105	0.51
County Government	990	4.17			
Community	595	2.51			
National Government	265	1.12			
Others	135	0.56			

Source: Author's calculations

The table 4.1 reveals that majority of the businesses were informal (74 per cent), had individual customers as their main market (86 per cent), were owned by individuals (92 per cent), had 46 per cent of their owners being male only and had 34 and 31 per cent of their employees having college and above, and secondary education, respectively. However, not all this entire sample was used for analysis. Data wrangling was carried out in order to remove negative costs, responses like “Others”, businesses with less than 100 employees (since the survey targeted SMEs, which by KNBS definition have less than 100 employees), the study retained 17614 observations. This is the sample that was used in the estimation of the model specified in Chapter three. The summary statistics of the categorical variables in this dataset

are presented in Table 4.2, while Table 4.3 gives the summary statistics of the continuous variables.

Table 4.2: Descriptive statistics of the categorical variables

Variable	Count	Share	Variable	Count	Share
Registration Status			Gender of owner		
No	13937	79.12	Male only	8020	45.53
Yes	3677	20.88	Female only	4868	27.64
Main market			Male-Female Partners	4007	22.27
Individual consumers	15660	88.91	Male-Male partners	506	2.87
MSMEs	1430	8.12	Female-Female Partners	213	1.21
Non-MSMEs	408	2.32	Education of the employees		
Government	93	0.53	College and above	5434	30.85
Direct exports	23	0.13	Secondary	5604	31.82
			Primary	4955	28.13
Ownership			None	1621	9.20
Individual	16291	92.49	Business Size		
County Government	822	4.67	Micro	16239	92.19
Community	307	1.74	Small	1247	7.08
National Government	194	1.10	Medium	128	0.73

Source: Author's calculations

The findings from Table 4.2 mirror those of Table 1. 79 per cent of the businesses were not registered, meaning they were informal. 89 per cent of the businesses had individual consumers as their major customer, while only 0.13 per cent of the businesses exported their products. 92 per cent of the businesses were owned by the individuals while only 1.1 per cent were owned by the national government. Males owned 46 per cent of the businesses and women owned 28 per cent of them. 32 per cent of the employees had secondary education, while 31 per cent had college and above. Only 9.2 per cent of the individuals did not have any education. When the

businesses were compared in terms of size, 92 per cent of the businesses were micro (meaning that they had less than 10 employees), seven per cent were small and only one per cent were medium. Next, the descriptive statistics of the continuous variables was examined. Table 4.3 outlines the findings.

Table 4.3: Descriptive statistics of the continuous variables

	count	mean	std	min	Median	max
Monthly credit	17614	7313.29	306786	0	0	2E+07
Monthly electricity	17614	597.42	4341.76	0	135	255000
Monthly licenses	17614	864.15	25413.5	0	97.5	1462502
Monthly rent	17614	1680.22	6444.13	0	630	218400
Monthly taxes	17614	2929.39	167636	0	0	2.1E+07
Business age	17614	7.88	7.62	0	6	76

Source: Author's calculations

The table displays the on average, the mean monthly credit cost is Ksh 7312.29. Notably, this was the highest cost among the other higher mean costs. It was closely followed by taxes, which averaged Ksh 2929.39. Rent came third, with the costs averaging Ksh 1680.22 per month. The mean monthly costs of licenses and electricity were Ksh 864.15 and Ksh 597.42, respectively. Finally, the mean age of the business was 7.88 years. A notable feature of the data is that there exists huge discrepancy between the mean and the median of the data. For example, the median monthly business taxes is zero, while the mean is 2929.39. Similarly, the median monthly credit cost is Ksh zero, while the mean is Ksh 7312.29. This clearly indicates that there are outliers in the data. To solve this problem, the survey utilized the logs of these variables, instead of the variables in their level form.

Next, the study sought to examine the distribution of the costs by formality status. This is because the descriptive statistics may give are silent on the distribution of these costs by formality – the costs are likely to be higher for the formal sector compared to the informal one. To verify this, the study calculated the descriptive statistics by formality. The results are presented the results in Figure 4.1.

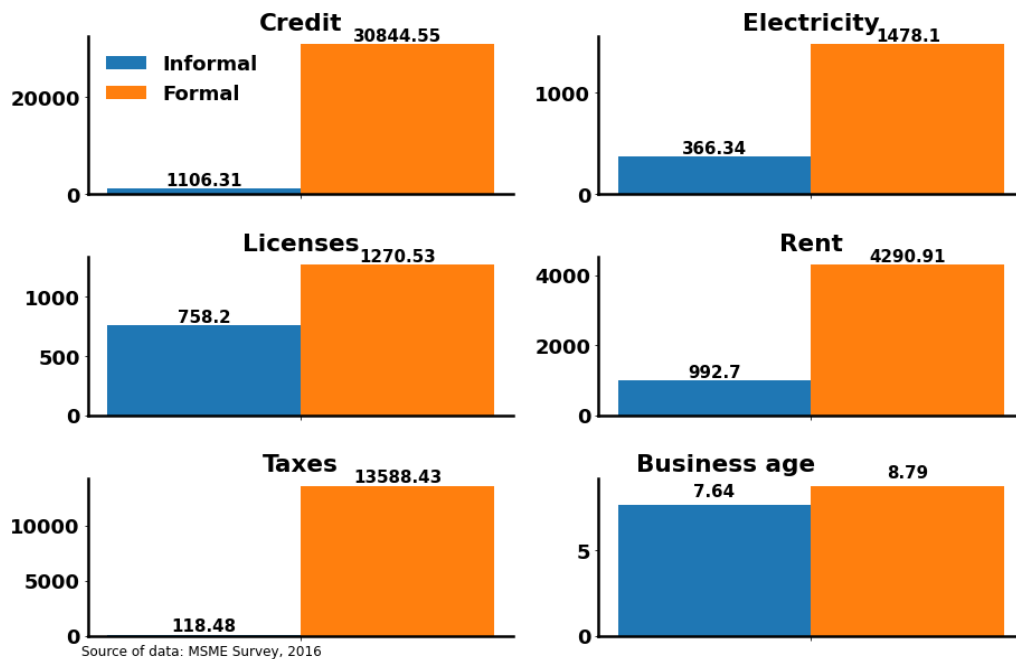


Figure 4.1: Mean monthly cost by formality status

The figure shows that all the costs are extremely high in the formal sector compared to the informal one. For example, the mean credit cost in the formal sector is Ksh 30843.55, while the one in the informal sector is merely Ksh 1105.31. The monthly costs of taxes show the greatest discrepancy between the two sectors: the cost is a paltry Ksh 117.48 in the informal sector, while it is a staggering Ksh 13,587.43 in the formal sector.

4.3 Diagnostic Tests

The diagnostic tests conducted include multi collinearity test, Hosmer Lemeshow test, specification and classification report. The results are presented in the appendix.

4.3.1 Multi collinearity test

A multi collinearity test was conducted to find out if the variables are highly correlated. Multicollinearity above 80 per cent reduces the precision of the estimates and biases the confidence tests. Table 4 presents the results of the multi collinearity test.

Table 4.4: Multi collinearity test results

	Credit	Electricity	License	Rent	Taxes	Business age
Credit	1					
Electricity	0.20	1				
License	0.00	0.03	1			
Rent	0.01	0.08	0.15	1		
Taxes	0.01	0.20	0.00	0.02	1	
Business age	0.02	0.08	0.00	0.02	0.01	1

The table 4.4 shows that all the correlations are very small, far below the 80 per cent threshold. For this reason, estimations were conducted without the fear of high collinearity among the variables.

4.3.2 Classification report

A good model has high classification accuracy. The outcomes are exhibited in Table A3 in the appendix section shows the classification results. The findings illustrate that the model attained a classification accuracy of 83 per cent $((13570 + 1011)/17614)$. This is far much better than the baseline classification accuracy of the random classifier of 50 per cent. The results of the model are corroborated with those of the area under curve graph which shows that the area under curve is 75.16 (a random classifier has an area of 50 per cent).

The area under the curve (AUC), which serves as a summary of the ROC curve, is a measurement of a classifier's capacity to distinguish between classes. The model performs better at differentiating between positive and negative classes when the AUC is higher. This means that the classifier model used is better than the random classifier. The further up from the 45 degrees line, the better the model.

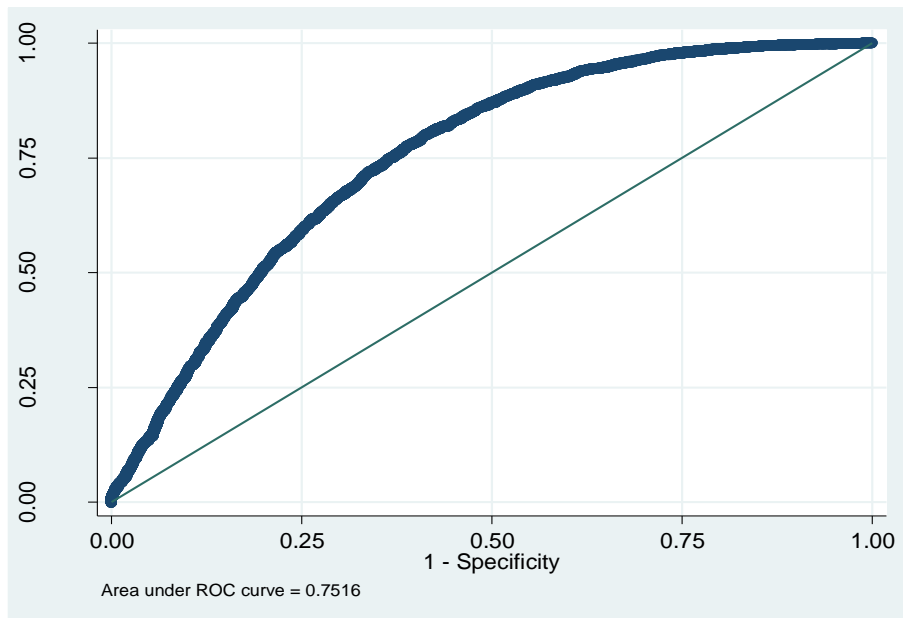


Figure 4.2: AUC curve

4.3.3 Hosmer Lemeshow Test

This is a goodness of fit test for logistic and probit regressions. The fitted probability of several subgroups within the actual values of the outcome variable are compared in this test. Any of the two random groups meeting the null hypothesis would be equal.

The table 8 captured in the appendix section displays that the Pearson Chi2 has a value of 18183 with a probability of 0.2345. This implies that the null hypothesis is not rejected: any two groups are equal, and hence the model is well specified. The outcomes are exhibited in Table A4 in the appendix section.

4.3.4 Specification Test

The estimation procedure assumed that the logistic model was the best for modelling the relationship between the formality status and the various explanatory variables. In addition, it was assumed that majority of the variables that affect formality status had been captured, and thus there were no omitted variables. The link test was used to test for this assertion. The outcomes are exhibited in Table A5 in the appendix section.

The table displays that the probabilities of both χ^2 and χ^2 squared are greater than 0.05. This indicates that the model is well specified and that there are no omitted variables.

4.4 Empirical Findings

4.4.1 Sectoral distribution of formal and informal enterprises in Kenya

The objective one of this research was to explore the sectoral distribution of formal and informal enterprises in Kenya. Descriptive analysis was used to analyse the 16 sectors that were surveyed by KNBS. Table 4.5 shows the results.

Table 4.5: Sectoral distribution of formal and informal enterprises in Kenya

Sector	Informal	Formal
Wholesale, retail and motor vehicle repair	77.10	67.15
Accommodation and food service activities	6.21	8.02
Agriculture, forestry and fishing	4.03	2.96
Manufacturing	3.86	3.94
Other service activities	3.42	4.30
Financial and insurance activities	2.27	3.67
Transportation and storage	1.15	1.33
Real estate activities	0.43	0.79
Education	0.40	2.31
Human health and social work activities	0.24	0.71
Administrative and support service activities	0.22	1.52
Information and communication	0.19	0.44
Arts, entertainment and recreation	0.12	0.44
Construction	0.12	0.38
Mining and quarrying	0.11	0.03
Professional, scientific and technical activities	0.10	1.69
Water supply; sewerage, waste management and remediation activities	0.02	0.16
Public administration and defence; compulsory social security	0.01	None
Electricity, gas, steam and air conditioning supply	None	0.08
Activities of households as employers; undifferentiated goods and services producing activities of households for own use	None	0.08
Total	100	100

Source: Author's calculations

The table 4.5 reveals that majority of the businesses in Kenya are in wholesale, retail and motor vehicle repair sector. The sector formed 67.15 and 77.10 per cent of all the businesses in the formal and informal sector, respectively. Accommodation and food service activities came second at 8.02 and 6.21 per cent in the two sectors, respectively. The third sector with most businesses was agriculture, forestry and fishing, with 2.96 and 4.03 of the businesses in the formal and informal sector belonging to the sector. Figure 4.3 shows a graphical distribution of the top 10 sectors in the two business types (formal and informal).

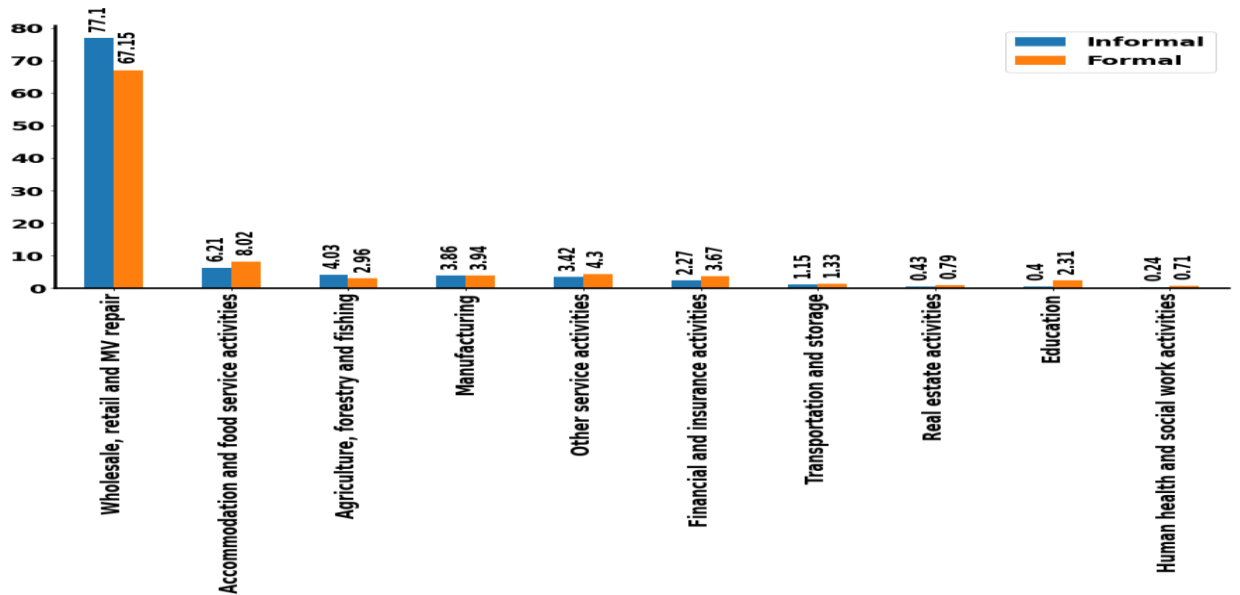


Figure 4.3: Distribution of top 10 sectors by formality status

The figure 4.3 shows that wholesale, retail and motor vehicle repair dominate the other sectors. 77 and 67 per cent of informal and formal businesses in the informal and formal businesses belong to this sector. Having presented the diagnostic tests, Table 4.6 display the outcomes of the logistic regression.

Table 4.6: Logistic regression results

Dep var: Formality status	Log of odds	St.Err.	t-value	p-value
Log of taxes	1.17	0.01	17.68***	0.00
Log of rent	1.04	0.01	5.22***	0.00
Log of licenses	1.37	0.03	17.59***	0.00
Log of electricity	1.05	0.01	5.70***	0.00
Log of credit	0.98	0.01	-2.31**	0.02
Business age	1.01	0.00	3.82***	0.00
Business size				
Small	1.90	0.14	8.62***	0.00
Medium	2.43	0.57	3.80***	0.00
Education				
Primary	0.96	0.09	-0.41	0.68
Secondary	1.25	0.11	2.51**	0.01
College and above	2.39	0.21	9.93 ***	0.00
Gender				
Female only	0.83	0.04	-3.54***	0.00
Male-Female partners	0.96	0.05	-0.72	0.47

Male-male partners	1.68	0.19	4.73***	0.00
Female-female partners	1.41	0.24	2.01**	0.04
Ownership				
County government	1.14	0.12	1.23	0.22
Community	1.30	0.19	1.75*	0.08
National government	1.70	0.32	2.80**	0.01
Main buyer				
MSME	0.90	0.07	-1.36	0.18
Non MSME	1.82	0.22	4.84***	0.00
Government	7.85	2.36	6.85***	0.00
Direct Exports	5.00	2.26	3.55***	0.00
Constant	0.02	0.00	-32.49***	0.00
Mean dependent var				
Pseudo r-squared		0.1622		
Chi-square		2999.57		0

Source: Author's calculations

Note: *** and ** indicate significance at one and five per cent, respectively. The base category for the base size is micro, none for education level, male only for gender, individuals for ownership, and individual consumers for main buyer. Formality status is one if formal and zero otherwise.

The coefficients of the logit model can be interpreted as the log of odds. For example, the log of odds of credit is 0.98. This implies that the log of odds in favour of being formal increase by 0.98 when the credit increases by one per cent. For the categorical variables, for example the value of college education of 2.39, it implies that the logs off odds for a business being formal are 2.39 higher for the college and above educated employees, compared to those with no education. To interpret the results as probabilities and hence make the results more intuitive, the marginal effects were conducted. Table 6 shows the marginal effects results.

Table 4.7: Marginal Effects results

Dep var: Formality status	dy/dx	Std.	Err.	Z value
Log of taxes	0.021***	0.00	18.26	0.00
Log of rent	0.005***	0.00	5.23	0.00
Log of licenses	0.042***	0.00	17.96	0.00
Log of electricity	0.007***	0.00	5.71	0.00
Log of credit	-0.002**	0.00	-2.32	0.02
Business age	0.001***	0.00	3.82	0.00
Business size				
Small	0.099***	0.01	7.62	0.00
Medium	0.144***	0.04	3.25	0.00
Education				
Primary	-0.004	0.01	-0.41	0.69
Secondary	0.027***	0.01	2.61	0.01
College and above	0.125***	0.01	11.38	0.00
Gender				
Female only	-0.025***	0.01	-3.59	0.00
Male-Female partners	-0.005	0.01	-0.73	0.47
Male-male partners	0.080***	0.02	4.32	0.00
Female-female partners	0.051*	0.03	1.88	0.06
Ownership				
County government	0.017	0.01	1.20	0.23
Community	0.036	0.02	1.65	0.10
National government	0.080**	0.03	2.52	0.01
Main buyer				
MSME	-0.013	0.01	-1.38	0.17
Non MSME	0.091***	0.02	4.33	0.00
Government	0.385***	0.06	6.14	0.00
Direct Exports	0.290***	0.10	3.02	0.00

Source: Author's calculations

Note: *** and ** indicate significance at one and five per cent, respectively.

Discussion of the results

The Table 4.7 displays that with the exception of the log of credit, all the other costs are positively associated with formality. This is contrary to what was expected. An increase in the cost of credit by one per cent reduces the probability that the business is likely to be formal by 0.002 per cent, ceteris paribus. However, an increase in the cost of taxes, rent, licenses and electricity by one per cent increase the probability that the business becomes formal by 0.021, 0.005, 0.042 and 0.007, respectively. Since these results seemed contrary to expectations, the study dug deeper to find a possible

explanation. A possible explanation would be that formal sector has higher revenues, and to get these revenues, businesses have to formalize and formalization comes with these given costs. Therefore, businesses that choose to be formal face higher costs, but at the same time get higher revenues. Table 4.8 shows the mean and median sales revenues for the businesses in the formal and informal sector.

Table 4.8: Sales revenues by formality

Formality Status	Median revenue	Mean Revenue
Informal	9750	32274.4
Formal	25350	344452

Source: Author's calculations

The table 4.8 reveals that the mean sales revenue for the informal sector is 32274, while that of the formal sector is 344452. The respective median values for the two sectors are 9750 and 25350. Therefore, the assertion that there are higher revenues in the formal sector and higher costs is validated.

Small and medium businesses have higher probability of becoming formal, compared to the micro ones. The respective probabilities are 0.099 and 0.144, respectively. Business with individuals with secondary and college and above education tend to have higher probabilities of being formal compared to those with employees with no education. These probabilities are 0.027 and 0.125, respectively. Female only businesses have a probability of 0.025 lower of being formal compared to the males only businesses. The respective probabilities for male-male partners and female-female partners are 0.08 and 0.051, respectively.

Businesses owned by national government has a probability of 0.08 higher compared with those owned by individuals. The coefficients of other forms of ownership were not statistically significant. Businesses whose main buyer are non-MSMEs,

government and direct exports have higher probability of being formal compared to the ones owned by individuals. The respective probabilities for these categories are 0.091, 0.385 and 0.29, respectively. Notably, businesses whose main buyer is government have the highest probability of being formal which is expected – one has to have a registered business in order to do business with the government.

The study found that credit and the probability of a business being formal were negatively related: the higher the cost of credit, the lower the likelihood of a business being formal. The findings are corroborated by those of Ingram, Vijaya, and Desai (2007) who did a similar study using a sample of six African countries. The authors found a positive relationship between the likelihood of a business being formal and its ease of access to credit. Ali (2012) also found that inaccessibility to credit hinders the growth of businesses in Egypt. However, unlike the current study which found a positive relationship between the cost of electricity and formality, the authors found a negative relationship between the two. Moreover, the authors also found a negative relationship between formality and corruption and tax, but a positive one with access to land.

The finding that education and gender are key determinants of informality are also in line with those of Wahba (2009) who found that highly educated workers and males were more likely to have formal businesses. A study by Taymaz (2009) in Turkey also mirrors those of the current study. The authors found that bigger firms had the highest probability of becoming formal and businesses with more educated workers were also likely to operate formally. However, the findings by the author that old business were likely to operate informally contradicts those of the current study which

found that the older the business, the higher was the likelihood of the business operating formally.

The finding that higher costs of electricity, rent, taxes and licenses are positively related with formality, though unintuitive, has been documented in literature by other authors as well. A study by World Bank (2013) in Kenya found that businesses chose to formalize in order escape the hardships they face in harassment by public officers, corruption and accessing water, electricity, finance and other public resources.

CHAPTER FIVE: SUMMARY, CONCLUSION AND POLICY IMPLICATION

5.1 Introduction

This chapter concludes the overall paper. It then draws policy conclusions and recommendations grounded on the results and gives the other possible areas of potential future studies.

5.2 Summary

This research aimed at finding the effect of the cost of doing business on the formalization of businesses in Kenya. The research was motivated by the fact that informal sector has been enormously growing in the country, despite the various governments effort to contain it. For instance, its share to total employment rose from 18 to 84 per cent between 1986 and 2018, and currently averages 80 per cent (as at 2021). The study had two specific objectives: to find out the sectoral distribution of the formal and informal enterprises in Kenya and to investigate the effect of the high cost of electricity, licenses, and taxes on the cost of doing business on the growth of the formal and informal sector in Kenya. To answer its objectives, the study utilized secondary cross-sectional household survey data gathered by the Kenya National Bureau of Statistics in 2016.

This study differed from the previous ones conducted in this area in two main aspects, and hence contributes positively to knowledge. First, the study examined the effects of the various costs of doing business on the growth of the formal and informal businesses in Kenya. Previous studies never explored this area. Moreover, the paper examined the sectoral distribution of businesses by their formality status. The previous studies only concentrated on the causes, determinants and the effects of large

informal sector. Second, the study used the most recent data (KNBS MSME household survey data of 2016).

To determine the sectoral distribution of the formal and informal sector businesses in Kenya, the survey utilized descriptive analysis on the 16 sectors that were surveyed by the KNBS. It was found that the wholesale, retail and motor vehicle sector forms the largest share of businesses in Kenya, with 77.1 and 67.15 of the businesses being in the informal and formal sector, respectively. The shares of the other sectors trailed by far, with their values being below 10 per cent. For instance, the shares of the accommodation and food service activities in the two sectors were 6.21 and 8.02 per cent, respectively.

The second objective was answered using a logistic regression. Since the results of the logistic regression can only be interpreted as odds ratio, the study calculated the marginal effects. With the exception of the coefficient of the log of credit, all the other variables were found to be positively associated with formality (or negatively related with informality). For example, a rise in the cost of credit by one per cent lowered the probability that the business is likely to be formal by 0.002 per cent, *ceteris paribus*. However, an increase in the cost of taxes, rent, licenses and electricity by one per cent increase the probability that the business becomes formal by 0.021, 0.005, 0.042 and 0.007, respectively. These results appeared to be rather contradictory. A further analysis was done to explore a possible explanation. It was found that formal businesses have higher revenues, and to get these revenues, they incur high costs of formalization. Thus, these businesses get higher revenues at the expense of high costs.

5.3 Conclusion

The main conclusion of this research is that the most prevalent business in both the formal and the informal sectors is the wholesale, retail and motor vehicle sector. It is followed by far by the accommodation and food services. Second, only credit was found to negatively affect the formation of formal businesses (or, alternatively, increase the formation of informal businesses).

5.4 Policy Implications

This paper provides several insights to policy makers that will enable their decision making on the various issues regarding the formal and informal sector in Kenya. First, the paper has found that the cost of credit is a great hindrance to the formation of formal businesses in the country. Thus, the policy makers and the relevant institutions should strive to offer affordable credit to ensure that individuals can access cheaper credit. However, caution should be exercised so that the measures instituted do not negatively affect the economy. For example, the government measures to lower the interest rates in order to encourage lending to the MSMEs in 2016 was unsuccessful – leading to its scrapping three years later. An alternative would be to set aside a given amount for lending and encourage small businesses to borrow. The two main sectors of focus should be wholesale, retail and motor vehicle sector and the accommodation and food services (since they form the highest proportion of small businesses).

Second, it was found that formal businesses make profit at the expense of the high costs. Thus, there are high costs that prevent informal businesses from formalizing and recouping the high revenues in the formal sector. The government should therefore improve business climate by ensuring that these costs are lowered, for example, enabling cheaper access to electricity and licenses, and lowering taxes.

Business size was found to significantly affect formality status: Medium businesses

were found to have a higher likelihood of becoming formal (compared to micro businesses), while small businesses were found to have a higher likelihood of formalizing (compared to the micro-ones). To this end, the government should enhance the growth of small businesses so that they are able to graduate from micro and small businesses to medium.

Third, it was found that entrepreneurs with workers with college and above have a higher probability of becoming formal compared to those with no education. This implies that the government should keep enhancing higher education opportunities for its citizens (for example, the continued effort by government to open more universities and college extend HELB and bursaries to the students is highly welcome).

Finally, businesses that sell to government, export markets or non-MSMEs are more likely to formalize (compared to those that sell directly to consumers). This is because they are more competitive. Therefore, the government should ensure training opportunities to the small and micro businesses owners to educate them and thus improve their competitiveness so that they can compete at a higher level. Moreover, it should remove the barriers that prevent them from selling at a higher level, more so the costs that were previously outlined.

5.5 Scope for Further Research

This paper has shed light on the sectoral distribution of various business in the formal and informal sectors. Moreover, it has enumerated the factors that are associated with the high probability of becoming informal. These include high credit costs, low education and business size. It was also found that the formal businesses are profitable at the expense of high costs. Nevertheless, there are other areas that can be considered

for future analysis. These include the following: (i) A similar study but at a panel level, say covering the last four surveys (in the last two decades) and examine the dynamics of change (ii) a similar study but at a sectoral level, and then compare the results (say manufacturing and accommodation) (iii) A related study, but at the county level and (iv) a gender disaggregated study to enable different conclusion and policy recommendations for males and females.

REFERENCES

- Aikaeli, J., & Mkenda, B. K. (2014). Determinants of informal employment: a case of Tanzania's construction industry. *Botswana Journal of Economics*, 12(2), 51-73.
- AGRESTI, A., (2002), *Categorical Data Analysis*, New York, Wiley-Interscience.
- Abdalla, A. G., Kiragu, J. K., Waswa, F. A., Ono, F. T., Kariuki, J. W., & Ikua, D. M. (2015). Effect of Hudumacenters (one stop shops) in service delivery—A case study of Mombasa Huduma Centre. *International Journal of Academic Research in Business and Social Sciences*, 5(6), 102-117.
- Africa Law Business. (2016). Legal developments that have changed the Kenyan business landscape Retrieved from <https://www.africanlawbusiness.com/news/6568-legal-developments-that-have-changed-the-kenyan-business-landscape>
- Ali, N. M. (2017). *Towards a better integration of the informal sector: three empirical essays on the interaction between formal and informal firms in Egypt and beyond* (Doctoral dissertation).
- Buehn, A., & Schneider, F. (2008). MIMIC models, cointegration and error correction: An application to the French shadow economy.
- Alm, J. and B. Torgler (2006). Culture Differences and Tax Morale in the United States and Europe, *Journal of Economic Psychology*. 27: 224-246.

- Capecchi, V. (1989) 'The Informal Economy and the Development of Flexible Specialization in Emilia Romagna', in A. Portes (eds) *The Informal Economy*, pp. 189-215. Baltimore: Johns Hopkins University Press.
- Castells, M. and A. Portes (1989) 'World Underneath: the Origins, Dynamics, and Effects of the Informal Economy', in A. Portes . (eds) *The Informal Economy*, pp. 11-37. Baltimore: Johns Hopkins University Press.
- Dada, J. (1995). Harnessing the potentials of the informal sector for sustainable development—Lessons from Nigeria. *United Nations Public Administration Network*.
- Denzau AT and North D (1994) Shared mental models: ideologies and institutions. *Kyklos* 47(1): 3-30.
- Medina, L., Jonelis, M. A. W., & Cangul, M. (2017). *The informal economy in Sub-Saharan Africa: Size and determinants*. International Monetary Fund.
- De Soto. (1989). *The other Path. The invisibly revolution in the Third World*. London. New - York: Harper and Row.
- Dell'Anno, Roberto. 2003. "Estimating the Shadow Economy in Italy: A Structural Equation Approach." Discussion Paper, Department of Economics and Statistics, University of Salerno.
- Dell'Anno, Roberto and Friedrich Schneider. 2004. "The Shadow Economy of Italy and other OECD Countries: What Do We Know?" Linz: University of Linz, Department of Economics. Discussion Paper.

Del'Anno (2003) and Del'Anno and Schneider (2004). Estimating the Size of the Shadow Economy.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2002. "The Regulation of Entry," *Quarterly Journal of Economics*, 117, 1: 1-37.

Enste, Dominik, and Friedrich Schneider. 2000. "Shadow Economies: Size, Causes, and Consequences," *Journal of Economic Literature*, 38, 1: 77-114.

Farazi, S. (2014). Informal firms and financial inclusion: status and determinants. *Journal of International Commerce, Economics and Policy*, 5(03), 1440011.

Farrell, D. (2004). The hidden dangers of the informal economy. *Mckinsey Quarterly*, 3, 27-37

Garcia-Bolivar, O. E. (2006). Informal economy: is it a problem, a solution or both? The perspective of the informal business. *Bepress Legal Series*, 1065.

GREENE, W. H., (2003), *Econometric Analysis*, Prentice Hall.

Haitian Center for Free Enterprise and Democracy (CLED) and the Institute for Liberty and Democracy (ILD), (2001). *The Capitalization Of The Poor In Haiti: Urban Real Estate and Business Formalization*, Urban Real Estate and Business Formalization Program. *United States Agency for International Development (USAID/Haiti)*.

Hart, K. (1973) .Informal Income Opportunities and Urban Employment in Ghana.
Journal of Modern African Studies, 11 (1) 62-89.

Helmke G and Levitsky S (2004). Informal institutions and comparative politics: a
research agenda. Perspectives on Politics 2(6): 725-740.

HOSMER, W. AND S. LEMESHOW, (2000), Applied Logistic Regression, Wiley.

Ronald Hope Sr, K. (2014). Informal economic activity in Kenya: benefits and
drawbacks. *African Geographical Review*, 33(1), 67-80.

Hsieh, Chang-tai and Peter J. Klenow, 2009, —Misallocation and Manufacturing TFP
in China and India, || Quarterly Journal of Economics, 124(4): 1403-48.

International Labor Organization. 2012. “Statistical update on Employment in the
Informal Economy.” Department of Statistics

International Labour Organization.(2016). *Women and men in the informal economy.*
A statistical picture. International Labour Organization.

ILO, 2012, Global estimate of forced labour. Available at:

[https://www.ilo.org/global/topics/forced-
labour/publications/WCMS_181953/lang--en/index.htm](https://www.ilo.org/global/topics/forced-labour/publications/WCMS_181953/lang--en/index.htm)

Institute for Liberty and Democracy (ILD), (2005). Program to Formalize the Assets
of the Poor of Tanzania and Strengthen the Rule of Law, *Institute for Liberty
and Democracy (ILD)*, Lima, Peru.

Johnson, Simon, Daniel Kaufmann, and Andrei Shleifer, 1997, “The Unofficial Economy in Transition,” *Brookings Papers on Economic Activity*, Washington, D.C.

Johnson, Simon, Daniel Kaufmann, John McMillan and Christopher Woodruff. 2000. “Why do firms hide? Bribes and unofficial activity after communism,” *Journal of Public Economics*, 76, 3: 495-520.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues". World Bank Policy Research Working Paper No. 5430 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1682130)

La Porta and Shleifer (2008) *The unofficial economy and economic development*. *Brookings Papers on Economic Activity* 47(1): 123-135.

La Porta, R., & Shleifer, A. (2014). *Informality and development*. *Journal of Economic Perspectives*, 28, 109–126.

Levy, Santiago, 2008, *Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico*, Washington: Brookings Institution Press

Loayza, N. (1999). *The economics of the informal sector: a simple model and some empirical evidence from Latin America*. The World Bank.

Loayza, Norman, 2007. —*The Causes and Consequences of Informality In Peru*,|| Banco de Reserva del Perú DT. N° 2007-018

- Loayza, Norman, Luis Servén and Naotaka Sugawara. 2009. "Informality in Latin America and the Caribbean." Policy Research Working Paper 4888, World Bank, Washington DC.
- Maizels, A., Palaskas, T. B., & Crowe, T. (1998). The Prebisch-Singer Hypothesis Revisited. In *Development economics and policy* (pp. 63-85). Palgrave Macmillan, London.
- McFadden, Daniel. "Econometric models of probabilistic choice." *Structural analysis of discrete data with econometric applications* 198272 (1981).
- Maloney, W. F. (2004). Informality revisited. *World development*, 32(7), 1159-1178.
- Moyo, B. (2022). Factors affecting the probability of formalizing informal sector activities in Sub Saharan Africa: evidence from World Bank enterprise surveys. *African Journal of Economic and Management Studies*, (ahead-of-print).
- McElwee G and Smith R (2014) Exploring Criminal and Illegal Enterprise. Bingley: Emerald.
- Ndaka, A. K. (2017). Informal Sector and Taxation in Kenya: Causes and Effects. *International Journal of Law, Humanities & Social Science*, 1(4), 77-86
- North DC (1990) Institutions, institutional change and economic performance. Cambridge: Cambridge University Press.

- Portes, A. Schauffler, R (1992). The informal Economy in Latin America: Definition, Measurement and policies. Working paper No 5. Programme in Comparative International Development. Department of Sociology, John Hopkins University. Baltimore MD 21218 USA.
- Prebisch, R. (1949). Growth, disequilibrium and disparities: interpretation of the process of economic development. *En: Economic survey of Latin America, 1949-E/CN. 12/164/Rev. 1-1951-p. 3-85.*
- Republic of Kenya.(1988). *Economic survey*. Nairobi: Government Printer.
- Republic of Kenya.(1992). *Economic survey*. Nairobi: Government Printer.
- Republic of Kenya.(1998). *Economic survey*. Nairobi: Government Printer
- Republic of Kenya.(2001). *Economic survey*. Nairobi: Government Printer.
- Republic of Kenya.(2006). *Economic survey*. Nairobi: Government Printer.
- Republic of Kenya.(2010). *Economic survey*. Nairobi: Government Printer.
- Republic of Kenya.(2014). *Economic survey*. Nairobi: Government Printer.
- Republic of Kenya.(2019). *Economic survey*. Nairobi: Government Printer.
- Rakowski, C. A. 1994a. Convergence & divergence in the informal sector debate: A focus on Latin-American. *World Development* 22 (4): 501-516.
- Ronald Hope Sr, K. (2014). Informal economic activity in Kenya: benefits and drawbacks. *African Geographical Review*, 33(1), 67-80.

- Rostow, W. W. 1960. The five stages of growth: A summary. The stages of economic growth: A non-communist manifesto. pp. 4-16. Cambridge, Massachusetts: Cambridge University Press.
- Rothenberg, A. D., Gaduh, A., Burger, N. E., Chazali, C., Tjandraningsih, I., Radikun, R., & Weiland, S. (2016). Rethinking Indonesia's informal sector. *World Development*, 80, 96-113.
- Schneider, F., Buehn, A., & Montenegro, C. E. (2010). Shadow Economies all over the World: New Estimates for 162 Countries from 1999 to 2007. (Policy Research Working Paper Series No. 5356). Washington, DC: The World Bank.
- Schneider, F., & Buehn, A. (2018). Shadow Economy: Estimation Methods, Problems, Results and Open Questions. *Open Economics*, 1(1), 1-29.
- Seira, E., Kaplan, D. S., & Piedra, E. (2007). Entry regulation and business start-ups: Evidence from Mexico. *Available at SSRN 978863*.
- Smith, Philip. 1994. "Assessing the Size of the Underground Economy: The Statistics Canada Perspectives." *Canadian Economic Observer*, Catalogue No.: 13-604- MIB1994028.
- Taymaz, E. (2009). *Informality and productivity: Productivity differentials between formal and informal firms in Turkey* (No. 0901). ERC-Economic Research Center, Middle East Technical University.
- Tokman, V. (1978). Informal – Formal Sector Relationships. *CEPAL Review*. First half of 1978, 99-134

- Tokman V. (1989) Economic development and Labour market segmentation in Latin American Periphery . Journal of Inter- America Studies
- Wanjiru, M. J. (2018). *Determinants of Utilization of Government Online Services in Kenya: Case of National Transport and Safety Authority Nairobi* (Doctoral dissertation, Kenyatta University).
- Welter F, Smallbone D and Pobol A (2015) Entrepreneurial activity in the informal economy: a missing piece of the jigsaw puzzle”, Entrepreneurship and Regional Development 27(5-6): 292-306.
- Williams N and Vorley T (2015) Institutional asymmetry: How formal and informal institutions affect entrepreneurship in Bulgaria. International Small Business Journal 33(8): 840-861.
- Wahba, J. (2009). Informality in Egypt: a stepping stone or dead end?, (ERF Working Paper No. 456). Cairo, Egypt: Economic Research Forum.
- World Bank. 2016. Informal Enterprises in Kenya. World Bank, Washington, DC. © World Bank. <https://openknowledge.worldbank.org/handle/10986/24973> License: CC BY 3.0 IGO.
- World Bank. (2016). Kenya: Country economic memorandum; From economic growth to jobs and shared prosperity. World Bank, Washington, DC. © World Bank.
- World Bank Group. (2018). *Doing Business 2018: Reforming to Create Jobs: Comparing Business Regulation for Domestic Firms in 190 Economies*. International Bank for Reconstruction and Development/World Bank.

Yusuff, O. S. (2011). A theoretical analysis of the concept of informal economy and informality in developing countries. *European Journal of Social Sciences*, 20(4), 624-636.

APPENDICES

APPENDIX A

Table A.1: Sectors to be Covered in the Study

	Sector
1	Wholesale and retail
2	Manufacturing
3	Accommodation and Food
4	Arts and recreation
5	Finance and insurance
6	Education
7	Human health
8	Construction
9	Transport and storage
10	Professional services
11	Information and communication
12	Agriculture and fishery
13	Real estate
14	Mining and quarrying
15	Water and sewage
16	Electricity and gas

Source: Author's calculation

Table A.2: Manufacturing Sub-sectors

	Manufacturing sub sectors
1	Manufacture of furniture
2	Manufacture of wearing apparel
3	Manufacture of food products
4	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
5	Manufacture of fabricated metal products, except machinery and equipment
6	Manufacture of textiles
7	Manufacture of leather and related products
8	Repair and installation of machinery and equipment
9	Printing and reproduction of recorded media
10	Other manufacturing
11	Manufacture of basic metals
12	Manufacture of chemicals and chemical products
13	Manufacture of beverages

Source: Author's calculation

DIAGNOSTIC TESTS

Table A3 : Classification results

Classified	D	~D	Total
+	13570	2666	16236
-	367	1011	1378
Total	13937	3677	17614

Source: Author's calculation

Table A4: Hosmer-Lemeshow Goodness of fit test

Label	Value
Number of observations	17614
Number of Covariates pattern	16058
Pearson Chi2	18183
Prob > chi2	0.2345

Source: author's calculation

Table A5: Specification Test

Label	Value
yhat	0.1464
yhat_squared	0.4531
Constant	0.642

Source: author's calculation