MODELLING POVERTY IN NAMIBIA USING BETA DISTRIBUTION

by

Ndubano Mafale

220008833

A thesis submitted in fulfilment of the requirements
for the degree of Master of Science in Applied Statistics

in the

Department of Mathematics and Statistics
Faculty of Health and Applied Sciences
Namibia University of Science and Technology

Supervisor:
Dr Jacob Ong'ala

Co-Supervisor:
Dr Dismas Ntirampeba

17 May 2022

DECLARATION

I, Ndubano Mafale, hereby declare that this study, *"Modelling poverty in Namibia using beta distribution"*, is my own work and is a genuine impression of my research, and that this work, or any part thereof, has not been presented or submitted for any degree award at any other institution. No piece of this report might be imitated, put away in any recovery framework, or communicated in any structure or using any and all means (for example electronic, mechanical, copying, recording or something else) without prior approval of the author or Namibia University of Science and Technology (NUST) in that behalf.

Mr. Ndubano Mafale
Student Number: 220008833
Author/Researcher
Signature: .... Date: 24/05/2022 ....

CERTIFICATION

This is to certify that this thesis titled **"MODELLING POVERTY IN NAMIBIA USING BETA DISTRIBUTION"** was undertaken by Ndubano Mafale at the department of Mathematics and Statistics, Namibia University of Science and Technology in partial fulfillment for the requirements of the award of the Master of Science degree in Applied Statistics.

Dr. J. Ong'ala
Supervisor
Signature: ....
Date: ...24/5/22...

Dr. D. Ntirampeba
Co-supervisor
Signature: ....
Date: ...24/05/22...

Mr. B.E. Obabueki
Head of Department
Signature: ....
Date: ...30/5/2022...

Prof. L. N. Kazembe
External Examiner
Signature: ....
Date: ...31/05/2022...

ABSTRACT

Modelling poverty is important as it helps to pinpoint human development areas that are most affected by poverty. Also, modelling poverty helps in understanding the patterns and levels of poverty, which helps policy makers to plan and make targeted interventions to reduce poverty. The traditional methods of estimating poverty such as the cost of basic needs approach or the poverty line approach are surrounded by a lot of controversies as they are said to underestimate or overestimate poverty. These methods are uni-dimensional as they only estimate poverty in one dimension (e.g consumption, income and expenditure) which neglects the humanistic needs side of poverty such as access to health or education. On the other hand, methods that include the Alkire and Santos (2011) method measure poverty in more than one dimension (e.g living standards, health, and education) but are faced with prejudice as the weighting method used is based on experts' opinion or the consensus of different stakeholders. Thus, this type of weighting method may result in biased weights and consequently result in inaccurate estimates of Multidimensional Poverty Index (MPI) values. This study focused on developing a multidimensional poverty model using beta distribution capable of estimating poverty for Namibia on regional and national levels. In addition, the study aimed at assessing the impact of weighting methods on MPI.

The first specific objective was to develop a multidimensional poverty model using beta distribution that could be used to model poverty for Namibia. The developed model showed that the MPI is equivalent to the expected value of the left-truncated beta distribution. The uncertainty surrounding the MPI was measured through the specification of the variance.

The second specific objective was to assess the impact of weighting methods on MPI. Two weighting methods (equal and entropy weighting) were adopted and their effect was assessed on the MPI obtained using the Alkire and Santos (2011) and the beta distribution approaches. The results revealed that the MPI values obtained when using entropy were

slightly bigger than the MPI values obtained using equal weighting under the Alkire and Santos (2011) approach compared to the beta distribution approach where the MPI values obtained when using equal weighting were bigger than the ones obtained using the entropy weighting method. Moreover, the entropy weighting method was found to be better than equal weighting as it is a mathematical based approach and is not affected by a change in the number of indicators compared to equal weighting which is subjective and sensible to the number of indicators.

The third specific objective was to identify more potential indicators that could be used in computing MPI which were not used in the initial computation of MPI by fitting a beta regression model. Using the NHIES 2015/2016 data, we fitted a beta regression model and identified the indicators that were left out in the initial computation of MPI.

In conclusion, the results revealed that the beta distribution model can be used to estimate regional and national poverty. The results also revealed that the entropy weighting method is useful in allocating weights when computing MPI as it eliminates the bias that comes with allocating weights. Moreover, the model can be used to identify areas that are highly affected by poverty and thus helping to come up with ways to alleviate the poverty. Finally, the beta regression model can help to identify indicators to be included in computing MPI.

ACKNOWLEDGEMENTS

First of all, I would like to thank the Almighty God for giving me the opportunity, knowledge, wisdom, and the strength to complete my studies. Further, I would like to thank my supervisors, Dr. Jacob Ong'ala and Dr. Dismas Ntirampeba. for their guidance and support throughout this study. I would also like to thank my family and friends for their support during my studies.

Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

BMI	**B**ody **M**ass **I**ndex
CBN	**C**ost of **B**asic **N**eeds
CBS	**C**entral **B**ureau of **S**tatistics
CI	**C**onfidence **I**nterval
CRITIC	**Cri**teria **I**mportance **T**hrough **I**nter-**C**riteria
DHS	**D**emographic and **H**ealth **S**urvey
FEI	**F**ood **E**nergy **I**ntake
FFA	**F**lood **F**requency **A**nalysis
GB2	**G**eneralised **B**eta distribution of the second order
GDP	**G**ross **D**omestic **P**roduct
GLMs	**G**eneralised **L**inear **M**odels
HDI	**H**uman **D**evelopment **I**ndex
ICT	**I**nformation and **C**ommunications **T**echnology
IID	**I**ndependent and **I**dentically **D**istributed
MDGs	**M**illennium **D**evelopment **G**oals
MLE	**M**aximum **L**ikelihood **E**stimator
MOM	**M**ethod of **M**oments
MPI	**M**ultidimensional **P**overty **I**ndex
MW	**M**ean **W**eight
NDP5	Fifth **N**ational **D**evelopment **P**lan
NHIES	**N**amibia **H**ousehold **I**ncome and **E**xpenditure **S**urvey
NPC	**N**ational **P**lanning **C**ommission
NSA	**N**amibia **S**tatistics **A**gency
NSF	**N**ational **S**ampling **F**rame
OPHI	**O**xford **P**overty and **H**uman **D**evelopment **I**nitiative
OR	**O**dds **R**atio

PDF	**P**robability **M**ass **F**unction
POME	**P**rinciple **o**f **M**aximum **E**ntropy
PPP	**P**urchasing **P**ower **P**arity
PSUs	**P**rimary **S**ampling **U**nits
SDGs	**S**ustainable **D**evelopment **G**oals
SDSN	**S**ustainable **D**evelopment **S**olutions **N**etwork
SMART	**S**imple **M**ulti-**A**ttribute **R**ating **T**echnique
SPSS	**S**tatistical **P**ackage for the **S**ocial **S**ciences
TV	**T**elevision
UNDP	**U**nited **N**ations **D**evelopment **P**rogramme
UNECE	**U**nited **N**ations **E**conomic **C**ommission for **E**urope
USAID	**U**nited **S**tates **A**gency for **I**nternational **D**evelopment
USD	**U**nited **S**tates **D**ollar
WB	**W**orld **B**ank

CHAPTER 1

INTRODUCTION

This chapter gives the background of the study, and it introduces the research topic by outlining the poverty situation in Namibia, Africa, and the world at large.

1.1 Poverty situation

1.1.1 Poverty situation worldwide

There is a great deal of exertion worldwide towards defeating worldwide poverty. As indicated by Peer (2018), under 8.6% of the world is assessed to be affected by extreme poverty. Extreme poverty is a complex and multi-dimensional marvel, ordinarily estimated quantitatively utilising a proxy consumption-based measure characterised as living under $1.90 every day (The World Bank, 2022). The United States Agency for International Development (USAID) characterises extreme poverty as the powerlessness to meet fundamental consumption needs on a manageable premise. Individuals who live in extreme poverty need both pay and resources and ordinarily experience the ill effects of interrelated and constant hardships, including hunger and ailing health, unforeseen weakness, constrained education, and underestimation or rejection (United States Agency for International Development, 2015).

Kenton (2021) defines the international poverty line as a financial limit under which a person is viewed as living in poverty. It is determined by taking the poverty threshold from every nation given the estimation of the merchandise needed to support one grown-up and changing it into dollars. The World Bank utilises median poverty lines in various financial settings to derive an estimation for extreme poverty comprehensively (Peer, 2018).

These median poverty lines are resolved using the information gathered by every nation through household surveys. Indisputably, the international poverty line is set at USD 1.90/day for low-income nations, USD 3.20/day for lower-middle income states, and USD 5.50/day for middle income countries (Peer, 2018).

Satapathy and Jaiswal (2018) define poverty as a condition where an individual needs fundamental conveniences, both money related and non-budgetary, to fulfil essential human needs. Satapathy and Jaiswal (2018) outline that in order to estimate poverty, a standard is first set which is the base use required to buy a container of products and ventures which are important to fulfil fundamental human needs, which is the poverty line. Each one of those individuals falling over the line are viewed as non-poor and those underneath the line are considered to be poor, and all the endeavours for social considerations and poverty mitigation are coordinated towards these individuals for their upliftment (Satapathy & Jaiswal, 2018). The poverty line partitions the whole population to be concentrated into two sections, specifically poor and non-poor.

The global Multidimensional Poverty Index (MPI) is a globally similar proportion of acute poverty that catches the different deprivations that individuals experience in regard to three dimensions; education, health, and living standards (Alkire & Jahan, 2018). The MPI looks past financial hardships to perceive how people experience poverty in various and concurrent habits. It perceives how people are left behind across the dimensions (Oxford Poverty and Human Development Initiative, 2019a). The three dimensions comprise of ten indicators which are used to compute the MPI. Out of the ten indicators, two from education, two are from health, and six from living standards as summed up in Table 1.1 (Alkire et al., 2011). In the present study, only nine indicators were considered as the data set used did not have information on child mortality.

TABLE 1.1: The MPI dimensions, indicators, deprivation cut-offs and weights

Dimension	**Indicators**	**Deprivation thresholds**	**Weights**
Health	Child mortality	Deprived if the family has lost any child to death	$\frac{1}{6}$
	Nutrition	Deprived in the event that any grown-up or kid, for whom there is nutrition data, is underweight	$\frac{1}{6}$
Education	Years of schooling	Deprived in the event that no household member has finished six years of school	$\frac{1}{6}$
	School attendance	Deprived in the event that any school-matured kid isn't going to class up to class 8	$\frac{1}{6}$
Living Standard	Cooking fuel	Deprived on the off chance that the family cooks with animal dung, wood or charcoal	$\frac{1}{18}$
	Sanitation	Deprived in the event that the household's sanitation facility isn't improved [according to MDG (Millennium Development Goals) guidelines], or it is improved but shared with other households	$\frac{1}{18}$
	Drinking water	Deprived on the off chance that the household doesn't have access to safe drinking water (as per MDG guidelines) or safe drinking water is in excess of a 30-minute stroll from home roundtrip	$\frac{1}{18}$
	Electricity	Deprived in the event that the household has no access to electricity	$\frac{1}{18}$
	Flooring	Deprived on the off chance that the household has a soil floor	$\frac{1}{18}$
	Assets	Deprived in the event that the household doesn't claim more than one of these resources: radio, TV, phone, animal cart, PC, bike, motorbike or fridge and doesn't possess a vehicle or truck	$\frac{1}{18}$

Alkire et al. (2011) state that the indicators inside each dimension are weighted equally such that the indicators in health and education get a $\frac{1}{6}$ weight and indicators within living standards gets a $\frac{1}{18}$ weight ($\frac{1}{3} \div 6$). Additionally, every individual is dispensed a deprivation score as per their deprivations in the component indicators. This is determined by taking a weighted total of deprivations experienced, with the target that the deprivation score for each individual lies around the interval of 0 and 1. The score increments as the measure of deprivations of the individual augmentations and appears at its imperative of 1 when the individual is deprived in all of the ten indicators, and 0 if a person isn't deprived in any indicator (Alkire et al., 2011). Alkire and Jahan (2018) add that the global MPI utilises the cross-dimensional poverty thresholds of 33%, recognising every individual as poor if their weighted deprivations add up to 33% or more.

A paper by Satapathy and Jaiswal (2018) explains that the MPI indicates the deprivation over indistinguishable three dimensions from the Human Development Index (HDI) and reveal the quantity of individuals who are enduring deprivation in 33% or a greater amount of the weighted indicators (multidimensional poor) and the quantity of weighted deprivation with which deprived households are ordinarily worried about.

1.1.2 Poverty profile in Africa

The average deprivation rate for sub-Saharan Africa stays at around 41%, and of the world's 28 most deprived countries, 27 are all in sub-Saharan Africa with a deprivation rate of more than 30% (Patel, 2018). Hamel et al. (2019) estimated that by 2030, 87% of the world's extreme poor will be living in delicate states, explicitly in African nations. In particular, Sub-Saharan Africa devastatingly represents 64.3% of the global population who are multidimensionally poor. Again, thinking about both outrageous and multidimensional poverty, they found that Africa represents over 70% of the world's least fortunate individuals in the two settings.

The United Nations expresses that despite being a resource-rich continent with the world's third fastest growing region, Africa contains most of the world's least developed nations (Cameron, 2011). According to Packtor (2017), about 70% of the world's poorest countries are situated in Africa. A total of 414 million people were living in outrageous poverty across sub-Saharan Africa in 2010. Also, the Central African Republic was positioned the poorest on the planet with a GDP per capita of $656 in 2016. Around one of every three individuals living in sub-Saharan Africa are malnourished and about 589 million people live without electrical power. Thus, a stunning 80% of the population is depended on biomass items, such as wood, charcoal and animal dung for them to cook. Moreover, 37% of 738 million individuals in need of access to clean water are living in sub-Saharan Africa. In addition, poverty in Africa brings about an excess of 500 million people who experience waterborne infections (Packtor, 2017).

Poverty in most African nations is estimated using the international line specified by the World Bank, which is currently at $1.90 per day with the different nations falling into their individual financial classes. In spite of the fact that it's non-exclusive, it is important to note that the majority of African nations are least developed hence most nations are estimated using the international poverty line for low income nations at USD 1.90/day, USD3.20/day for lower middle income nations (like India, Egypt and the Philippines) and USD 5.50/day for upper - middle income countries (like South Africa, Jamaica and Brazil) (Silver & Gharib, 2017).

This estimation of poverty to a great extent relies upon household surveys that examine and show household expenditure information. In spite of the fact that poverty insights are especially hard to create in Africa because of the absence of funds, intrigue or potentially ineptitude, there are different strategies that give hints about poverty and these are not really proportions of poverty, for example, resource indicators among others (Klasen, 2014). In fact, the resource indicators give an idea regarding poverty by the accumulation of assets, and this data created from cell phone records, social networks and web traffic is alluded to as bulky data.

However, these methodologies are shockingly restricted as they produce one-sided results, for example, the accumulation of assets may not really mean rich but this might be because of affordability. Similarly, bulky data avoids the poor when information concerning poverty is being gathered as most poor individuals have no access to these devices (Klasen, 2014).

1.1.3 Poverty status in Namibia

Poverty is perceived worldwide as one of the difficulties confronting a number of nations, particularly in Africa, and Namibia is one of the African nations confronted with poverty (Namibia Statistics Agency, 2012). Despite it being considered a middle-income nation, Namibia has one of the most inconsistent conveyances of riches and income on the planet (Korn, 2017). According the Errata on 2015/2016 NHIES Report of the Namibia Statistics Agency (2021a), the Gini coefficient for Namibia stands at 0.576, which is an improvement from the previous years. In particular, since independence the Gini coefficient has improved as follows: 0.701 (in 1993/1994), 0.600 (in 2003/2004), 0.597 (in 2009/2010) and 0.576 (in 2015/2016) (Namibia Statistics Agency, 2021a). On the other hand, the poverty rate in Namibia stands at 26.9% and it is more intense in the northern areas of Kavango, Zambezi, Kunene, Ohangwena and Oshikoto, where up to 33% of the population lives in poverty (English, 2016).

According to the World Bank Group (2020), the greater part of the Namibian population lives beneath the national poverty line which is currently at 18%. It was additionally uncovered that the national poverty line moved from 69.3% in 1993/94 to 28.7% in 2009/10, showing a decrease of 40.6% in the quantity of individuals who live in poverty, that is, those falling beneath the poverty line concerning the total population and a further 17.4% in 2015/16. Along these lines, the poverty levels have diminished fundamentally since independence. A household that spends over 60% of its expenditure on food in Namibia is considered to be 'poor' and the one that burns through 80% or more on food is viewed as 'seriously poor' (Schmidt, 2009).

For 1993/1994, 2003/2004 and 2009/2010, two poverty lines were set up for individuals living in poverty where the levels of consumption per grown-up equal were less than N$145.88, N$262.45 and N$377.96; and N$106.78, N$184.56 and N$277.54 for seriously poor for the years 1993/1994, 2003/2004 and 2009/2010, individually. Utilising these definitions, the rate of poor and seriously poor people were assessed at 28.7 and 15.3 percent, separately (Namibia Statistics Agency, 2012). For the 2015/2016 period, new poverty lines were set up with the food poverty line of N$ 293.1, the lower bound poverty line of N$ 389.3, and the upper bound poverty line of N$ 520.8 (Namibia Statistics Agency, 2016). According to Namibia Statistics Agency (2016), the general poverty level has diminished fundamentally with 10.7 percent (from 28.7 to 18.0), while the imbalance in income appropriation has stayed high with a slight decrease of 2.5 percent points from the past survey of 2009/2010 to the survey of 2015/2016.

In another study led by the National Planning Commission (NPC) with respect to poverty planning, the methodology of Elbers *et al.* was utilised (National Planning Commission, 2016). The technique includes, initially, picking a lot of household attributes found in the given two datasets, trailed by the utilization of the smaller dataset that has consumption information. A connection between the picked set of household qualities and family consumption is then inferred and used to anticipate the normal degree of use for every household in the census. Poverty rates are for the most part used to compile statistics in Namibia and the information is used for arranging and looking for help from different nations.

In 2008, the Central Bureau of Statistics (CBS) introduced a new poverty line following the purported Cost of Basic Needs (CBN) approach, which evaluates the expense of a heap of basic food and non-food products. As indicated by the National Planning Comission (2008), the way towards setting the new poverty line was partly into two significant stages. To begin with, utilising the Namibia Household Income and Expenditure Survey (NHIES) information for households with low consumption expenditure, a food container is resolved dependent on real consumption patterns of low pay households. Second, considering non-food prerequisites notwithstanding food needs, two poverty lines were built for "poor" and "seriously poor" household units where levels of consumption per grown-up equivalent were below N$ 262.45 and N$ 184.56, separately. Schmidt (2009) outlines that the CBN technique for estimating poverty is without a doubt a major advance forward for Namibia, aligning the nation with best practices globally. It is along these lines that other data diagnostic tools are utilised in contrast with the poverty line approach used by the Namibia Statistics Agency (NSA), and beta distribution is one of those.

The beta distribution represents continuous probability distributions and it is an exceptionally flexible approach to represent results like extents or probabilities, and a few uncertainties might be well represented by them (Johnson et al., 1994; Mazucheli & Menezes, 2019). Mazucheli and Menezes (2019) notes that beta distribution in its standard form is the most significant likelihood distribution for data analysis characterised on the continuum somewhere in the range of 0 and 1. In particular, there are two parameters which cooperate to decide whether the distribution has a mode in the inside of the unit interval and whether it is symmetrical. The beta distribution can be used to depict the variety observed across individuals; however, it can likewise portray subjective a degree of beliefs (from a Bayesian perspective) and one advantage of it is that it can take on a wide range of shapes (Johnson et al., 1994). In addition, the beta distribution of the first kind can be utilised to display the appropriation of estimations whose qualities all lie somewhere in the range of zero and one. It can likewise be used to model the distribution for the probability of an event of some discrete occasion (Johnson, 2013). The data used in poverty measure lies in the range of zero and one, hence, the beta distribution is the choice used in our present research study to model multidimensional poverty for Namibia. Modelling was done by beta distribution, and estimation was done using the maximum likelihood estimation method to estimate poverty in Namibia which employed the Newton-Raphson method.

1.2 Statement of the problem

Poverty is one of the major challenges facing quite a number of countries in the world, especially in Africa (World Bank Group, 2020). The National Planning Commission of Namibia (NPC) developed an index of multiple deprivation which was based on the 2001 and 2011 housing and population census using five main deprivation domains (material, employment, health, education, and living environment) to estimate poverty (National Planning Commission, 2015).

The method used by NPC modified the global recommended indicators for MPI computation without further scrutiny of their indicator inclusion, that is, they modified these indicators without justification of why they didn't use the recommended indicators and why the modification was done (National Planning Commission, 2015). Therefore, the deprivation domains still need more scrutiny to determine their significance in the MPI. On the other hand, the Namibia Statistics Agency (NSA) uses the absolute poverty lines which are based on the Cost of Basic Needs (CBN) approach which is an income-based approach to estimate poverty in Namibia (Namibia Statistics Agency, 2016). However, there are controversies which surround this current method of computing the poverty indices. The discussion is still expected to continue more so in political freedom environments. This method of poverty measure may underestimate or overestimate poverty and it only focuses more on consumption rather than other poverty humanistic needs. Also, when using the CBN approach, poverty reduction is just understood as income and economic growth, which is too simplistic and is unable to reflect poverty's humanistic needs such as education and health attainment. Moreover, a comparison with other countries becomes a challenge when this approach is used. There is therefore a need to use more comprehensive approaches to estimate poverty such as the MPI which gives a more inclusive definition of poverty and that incorporates global recommendations, country dynamics and random variations. Methods such as the Alkire and Santos (2011) approach of computing MPI are faced with prejudice as the type of weighting used is based on experts' opinions and thus may result in biased weights and consequently resulting in inaccurate MPI estimates. Also, the Alkire and Santos (2011) method does not quantify the uncertainty around the computed MPI values. The beta distribution fits the type of data that was used since it is also characterised in the same interval and it is the most significant likelihood distribution for data analysis in its standard form as suggested by Mazucheli and Menezes (2019). Our study therefore sought to develop a multidimensional poverty model for Namibia with improved accuracy and inclusivity in poverty measures using the beta distribution.

1.3 Objectives of the study

The main objective of our study was to develop a multidimensional poverty model for Namibia with improved accuracy and inclusivity in poverty measures using beta distribution.

Specific objectives

(i) To develop a multidimensional poverty model in Namibia using beta distribution;

(ii) To perform sensitivity analysis of weighting methods on the multidimensional poverty index in Namibia;

(iii) To fit beta regression model capable of determining other potential significant indicators of the multidimensional poverty index in Namibia.

1.4 Significance of the study

The development of a basic multidimensional poverty model for Namibia will bring a new and different perspective on poverty in addressing the issue of poverty in Namibia. This research study also helps to shed some light on the use of the beta distribution method with regards to estimating poverty rates. The results of this study provide the necessary evidence for planners and policy makers in order to design appropriate poverty interventions across the country through the identification of areas in severe multidimensional poverty conditions. In particular, the results will contribute immensely to the goals and targets of the 2030 agenda for sustainable development in Namibia. Moreover, the determination of deprivation scores and decomposition of MPI by region will help policy makers in making targeted interventions at regional level which will help reach the Sustainable Development Goals such as Goal 1 (End poverty), Goal 2 (End hunger), Goal 3 (Good health and well-being), Goal 4 (Quality Education), Goal 6 (Clean Water and Sanitation), Goal 7 (Affordable clean energy), Goal 9 (Industry, Innovation and Infrastructure) and Goal 11 (Sustainable cities and Communities). In Addition, our research study provides accuracy and more clarity on the issue of poverty in Namibia when using multidimensional poverty measures which capture poverty conditions in different areas like education, health, and living standards compared to the uni-dimensional approach which only captures poverty conditions in one area. Furthermore, the findings of this study can be used to compare multidimensional poverty levels among different regions in Namibia and also to compare Namibia with other countries.

1.5 Thesis outline

Chapter One introduced the poverty situation in Namibia, Africa and worldwide. Also, it provided a short review of beta distribution, the statement of the problem, the objectives of the study, and the significance of the study. In Chapter Two, we discuss the methods used in measuring poverty, namely, one dimensional poverty measure and the multidimensional poverty index measure. Poverty dimensions, indicators, weights, deprivation thresholds, and deprivation scores are discussed. Also, the beta distribution of modelling poverty and other approaches used to estimate poverty are reviewed. Furthermore, beta distribution and beta regression, and their applications were reviewed. This chapter helps the reader to understand the methods which are used in poverty measures and a general understand of the beta distribution and its application.

In Chapter Three, we give a brief description of the data used in this study. We also discuss how the data processing was carried out. The delimitation of the study and the ethical issues are presented in this chapter. It provides steps on how the deprivation scores and weights were determined. Furthermore, it discusses how the computation of MPI was carried out, how modelling with beta distribution and estimating of its parameters was done, and how beta regression was used to identify potential indicators.

Chapter Four presents the results, as well as the interpretation and discussion of the study findings.

Chapter Five reviews and evaluates the objectives of the study to determine if they have been met and it finally provides conclusions and recommendations for future research studies.

An appendix of all the R-programme codes used in this study and a reference list of all the references used in this study are shown at the end of Chapter Five.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter discusses the work done on the measures of poverty by scrutinising the available literature, assessing challenges and identifying loopholes and suggesting possible improvements.

2.2 Poverty measures

There are different ways that governments, practitioners, international organisations, and policy makers use to measure poverty. The measurement of poverty can be considered as absolute or relative depending on their computations. The absolute approach of poverty deals with set guidelines that are reliable over the long haul and between nations. On the other hand, the relative poverty approach deals with poverty in a social context. Both absolute and relative poverty approaches are based mostly on personal annual income and in most cases they do not take the total wealth into account. This ignores important components of economic well-being since poverty usually involves being deprived in several indicators (Boundless, 2014). Therefore, this approach may be considered as a one dimensional approach.

Apart from the wealth aspect, access to fundamental necessities may be employed in measuring poverty, since it has been set up that individuals who may have sufficient income do not use it wisely. Similarly, in the event that adequately solid informal organizations or social assistance frameworks are in place, then the extremely poor people may be deprived

(Boundless, 2014). The focus of the economic measure is on material needs, including necessary needs like food, shelter, clothing, or clean drinking water. This aspect of poverty in this sense might be perceived as a condition where an individual or community lacks basic needs to attain minimum well-being, resulting from a diligent lack of income. The social measure of poverty includes lack of social needs such as education, access to information, political power or health care ; inequitable social relations, dependency, social exclusion, and/or diminished capacity to participate in society. All these can help in understanding the multidimensional poverty measure (Pogge, 2015).

2.2.1 One dimensional poverty measure

The one dimensional poverty measure, also referred to as the uni-dimensional method, is used at the point when an obvious single-dimensional asset variable, like income, is chosen as the reason for the poverty assessment measure. In this method, a solitary dimensional variable and a solitary threshold are utilized. It could either be a solitary asset variable, for example, income or total expenditure accumulated across various classes announced in an expenditure survey, or maybe separated from consumption surveys that expect respondents to review their amounts and costs (Alkire & Santos, 2011). The construction of a single (uni-dimensional) index may not be enough or of much use when it comes to sound development policy making (Ravallion, 2011).

Ravallion (2011) suggests that efforts and resources should focus more on developing a unique and best possible measure with different dimensions of poverty which are deemed relevant aiming for a credible index rather than a single dimensional index. A consensus exists that a uni-dimensional measure of poverty is not credible in capturing a realistic aspect of poverty (Alkire & Santos, 2011; Ravallion, 2011).

Wong (2012) points out that the uni-dimensional poverty measure has been adopted in a lengthy time by policy makers around the world, which places a massive priority on the income perspective. Wong (2012) notes that the approach affiliates poverty with income by defining being poor as a deprivation of financial resources, which in turn leads to understanding poverty reduction as growth in financial resources.

Wong (2012) further indicates that the basic needs approach which is commonly applied as a uni-dimensional poverty measure, is materialistic and identifies a package of basic consumption needs such as shelter, sanitation, clean water, food amongst others. The population's access is then evaluated to know whether the population has fair access to the package. A person with inadequate access to these resources may be considered to be poor. The basic needs approach is easy to implement in the sense that patterns of poverty can be adequately attained by comparing the access to basic needs at different time intervals. However, this approach emphasises more on materialistic deprivations and it does not address the issue of different groups or individuals who require different needs at different times.

Watson and Derrill (2014) notes that the basic needs approach is more interested in the poverty experienced in the present and not much on the poverty experienced in the long run. Watson and Derrill (2014) also notes that the basic needs approach's main foundation is that whether people can meet their basic needs or an individual who is unable to meet their basic needs is considered poor. While the basic needs approach can be supported as a measure of poverty, Watson and Derrill (2014) further argues that there is no preceding ground to say that a person's basic needs might not be relative to the needs of the other members in the society. In addition, a great possibility exists that needs might relatively increase as a consequence of the average income rising and the needs taking on increasingly social dimensions.

According to the United Nations Economic Commission for Europe (2017), poverty can be measured using approaches such as the uni-dimensional and multidimensional measures. Uni-dimensional uses the monetary and food energy intake (FEI) approaches to measure poverty. The monetary approach is income and expenditure based, it looks at both the relative and absolute poverty line for both income and expenditure. The absolute poverty line describes the threshold which is the subsistence minimum for the cost of basic needs as the cut-off point considering a person as poor if they are below the subsistence minimum. Using the thresholds, a person is severely poor if they have an income/expenditure below the Purchasing Power Parity (PPP) of $1.90/day and just poor if the person has an income/expenditure below the PPP of $3.1/day.

Purchasing Power Parity (PPP) is defined as the amount of units of a country's money expected to buy a comparable proportion of items and adventures in the domestic market as one dollar would buy in the US. This method of PPP permits the measure of what trade between two monetary forms is expected to communicate the exact purchasing power of the two monetary standards in the particular nations (Jose, 2015).

Multidimensional poverty considers deprivations as an indicator dashboard and indices of multiple deprivations, including material deprivation; multidimensional poverty estimates which are internationally comparable look at the multidimensional poverty index thresholds for different dimensions; and the official national multidimensional which looks at the severely and moderately poor (United Nations Economic Commission for Europe, 2017).

2.2.2 Multidimensional poverty index (MPI)

MPI uses education (Years of schooling and School Attendance), health (Nutrition and Child mortality), and standard of living (Cooking fuel, Sanitation, Water, Electricity, Floor, and Assets) dimensions to determine the intensity and incidence of the poverty experienced by a population. Two versions of the MPI exist. A regional or global MPI, which is comparable across countries like the global income measure of PPP$1.90/day. Like national monetary poverty measures, national MPIs echo national specific priorities that are not necessarily comparable across countries (United Nations Economic Commission for Europe, 2017).

The MPI is designed to measure the non-monetary based dimensions of poverty (acute poverty), that provide a full assessment of the magnitude of poverty and deprivations (Sustainable Development Solutions Network, 2020). This is achieved by combining two significant pieces of information: the poverty incidence or the proportion of people (within a specific population) experiencing multiple deprivations denoted by H, and their deprivation intensity (the average proportion of their experienced deprivation) denoted by A. The relevance of both the intensity and the incidence of these deprivations is very high (Alkire & Santos, 2011). The MPI is then computed as the product of both the head count ratio and the average deprivation score of multidimensional poor people (Alkire et al., 2011).

MPI can be decomposed into different sub-population groups by breaking it using dimensions to reflect the difference in the composition of poverty between regions or other groups (Oxford Poverty and Human Development Initiative, 2017).

The MPI can be an exceptionally useful tool in evaluating how countries meet the Sustainable Development Goals (SDGs) (United Nations Economic Commission for Europe, 2017). The SDG (goal 2.1) focuses on multidimensional poverty and eliminating poverty in all its dimensions, which is the first goal (United Nations Economic Commission for Europe, 2017). This goal targets decreasing the proportion of children, women and men of any age living in poverty in the entirety of its dimensions at least by half by the year 2030 according to national definitions (Sustainable Development Solutions Network, 2020). Countries can use either regional/global (comparable) or national MPIs to measure the achievement of the second goal of the SDG (United Nations Economic Commission for Europe, 2017).

There is quite a number of advantages associated with the MPI, that is: it takes into account comparisons across nations or regions of the world, and inside a country, correlations between regions, rural and urban areas, ethnic groups, and other key household attributes in the community can be done. The comparison can be achieved due to its robustness and direct measure of acute poverty. It further enables the analysis of poverty patterns overtime and a quantification of the contribution of each indicator and dimension to the overall poverty (Alkire & Santos, 2011). In the study on the multidimensional poverty index, Alkire et al. (2011) explained that, any available information for household members can be used to construct the MPI in order to identify whether a person is deprived or not.

2.2.3 Multidimensional poverty dimensions and indicators

Alkire et al. (2011) used ten indicators to compute the MPI which are grouped in three dimensions namely, education, health, and living standards. There are two indicators for education (child school attendance and years of schooling), two for health (child mortality and nutrition), and six for living standards (electricity, safe drinking water, improved sanitation, cooking fuel, flooring, and asset ownership). In computing MPI, data must be recent, available and having all the relevant indicators, and all this data should come from the same survey (Alkire et al., 2011). There is a possibility of expanding the dimensions and indicators for country specifics.

The education dimension has two complementing indicators that is; completed schooling years of household members, and if children are going to school. Years of schooling goes about as an intermediary for the degree of knowledge and comprehension of the household. But it can be noted that both school attendance and years of schooling are defective intermediaries because, for example, the degree of knowledge attained, the quality of schooling or abilities are not captured by these proxies (Alkire et al., 2011).

The health dimension is the most hard to quantify in light of the fact that its tantamount indicators of health are generally missing from the household survey and this puts a restraint on this dimension in terms of data availability. Even though the two indicators used in this dimension are related, they depart from the standard health indicators (life expectancy at birth, the proportion of immunization inclusion, relieved tuberculosis cases, baby death rate, or of met family arranging needs, maternal mortality proportion, extent of tuberculosis patients forsaking treatment, rate of incidence of AIDS, etc.) significantly and depend mostly on the survey utilized and the demographic make-up of that household. Nutritional status of the household members is measured in children with malnutrition. People with nutritional issues are more prone to other health complications such as less ability to learn, less ability to concentrate and perform work in the best way (Alkire et al., 2011).

According to Alkire and Santos (2011), the nutritional information indicator in children looks at the aspect of being under-weight (weight-for-age). A child is malnourished on the off chance that she/he is at least two standard deviations beneath the middle of the reference populace, whereas the Body Mass Index (BMI) is used as the nutritional indicator in adults considering an adult as undernourished in the event that he/she has a BMI less than 18.5.

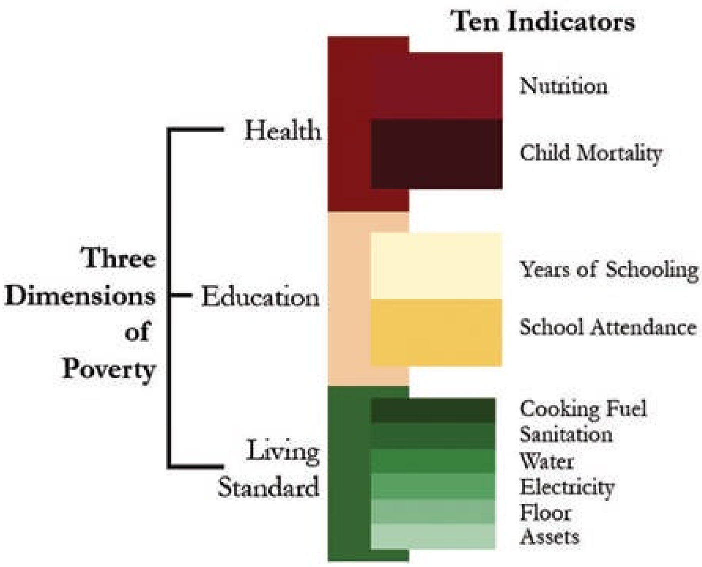
Three out of the six indicators grouped under the living standard dimension are in tandem with three standard MDG pointers that are identified with living standards and health, and which affect women especially, for example, the utilization of clean cooking fuel, further developed sanitation, and clean drinking water. The two indicators used to assess the quality of the housing (electricity and flooring material) are non-MDG indicators. While the final indicator covers asset ownership of some consumer goods such as: television, radio, bicycle, telephone, motorbike, truck , car and refrigerator (Alkire & Santos, 2011).

The selected deprivations thresholds for every indicator (with the exception of the one identifying with assets) are upheld by the universal accord as they follow the MDG indicators as intently as information granted (Alkire & Santos, 2011).

The living standard indicators measure the means of living rather than ends. And these living standard indicators have two strengths. Firstly, unlike income, these means are closely linked to the end they are supposed to measure. Secondly, most of these indicators are closely related to the Millennium Development Goals (MDGs), which provide strong evidence to include them in the MPI (Alkire & Santos, 2011).

The theoretical framework of our study to model multidimensional poverty in Namibia hinged on the framework outlined by Alkire and Santos (2011). This framework outlines the MPI with ten indicators (two for education, two for health, and six for living standards) which were chosen after an exhaustive counsel measure including specialists in all the three dimensions. Moreover, the chosen indicators had to meet the states of information accessibility and cross-country examination.

Furthermore, the final ten chosen indicators had to be nearly the main arrangement of indicators that could be utilised to compare around 100 nations. The composition of the MPI as outlined by Alkire and Santos (2011) is shown in Figure 2.1.



Source: Alkire and Santos (2011)
FIGURE 2.1: The MPI framework - indicators and dimensions

In this MPI, the unit of measurement was the household rather than every individual since when utilising every individual as the unit of measurement certain factors were not watched for all household members. Hence, accepting the household as the unit of assessment doesn't reveal intra-nuclear family irregularities, yet it is regular and expects to be shared positive (or adverse consequences) of achieving (or not achieving) certain results which is ideal for estimating multidimensional poverty (Alkire & Santos, 2011). The theoretical framework of our study hinges on this framework while using the household as the unit of measurement.

2.2.4 Estimating weights in a standard MPI

When constructing MPI, the indicators and domains/dimensions are combined using weights. The issue of how to set pertinent weights across dimensions is one of the complex and exceptionally disputable issues arising in the multidimensional setting of well being research (Cavapozzi et al., 2015). Alkire et al. (2011) showed that the three dimensions in the MPI are equally weighted, such that each dimension receives a 1/3 weight. The indicators in each dimension are also weighted equally as shown in Table 2.1. The child mortality in Table 2.1 was not considered because our data set did not have any information on child mortality.

TABLE 2.1: Equal weights of Dimensions and Indicators of MPI

Domain	**Weight**	**Indicator**	**Weight**	**Final Indicator weights**
Health	$\frac{1}{3}$	a	$\frac{1}{2}$	$\frac{1}{6} = ((1 \div 3) \times (1 \div 2))$
		b	$\frac{1}{2}$	$\frac{1}{6}$
Education	$\frac{1}{3}$	a	$\frac{1}{2}$	$\frac{1}{6}$
		b	$\frac{1}{2}$	$\frac{1}{6}$
Living Standard	$\frac{1}{3}$	a	$\frac{1}{6}$	$\frac{1}{18}$
		b	$\frac{1}{6}$	$\frac{1}{18}$
		c	$\frac{1}{6}$	$\frac{1}{18}$
		d	$\frac{1}{6}$	$\frac{1}{18}$
		e	$\frac{1}{6}$	$\frac{1}{18}$
		f	$\frac{1}{6}$	$\frac{1}{18}$

Gao and Sun (2020) adopted the entropy weight method to measure the weight of each index and dimension in the MPI instead of using the equal-weight method. They noted that the entropy method is a widely used method which considers giving objective weight to each indicator. Furthermore, when finding a solution to a problem, index weighting can be done by the entropy method in accordance with the degree of the variation that exists in the index value. In addition, they maintain that the entropy weight only illustrates the index degree of discrimination for the evaluated target but itself is not an important coefficient of the index. Gao and Sun (2020) suggest that the entropy weighting method should be used when there is a variability between indicators.

2.2.5 Poverty deprivation thresholds

In the MPI model, the minimum level of deprivation is also referred to as the deprivation threshold. To arrive at these cut-offs, a consultation process involving professionals and experts in all three dimensions is always held to determine these (Alkire & Santos, 2011).

The cut-off point within the education dimension requires that somewhere around one individual in the family has finished five years of schooling and that all children of school age are attending classes 1 to 8 of school in terms of deprivation thresholds (Alkire & Santos, 2011). Our research study improved the cut-off points to at least two people in the household that have completed five years of schooling and that all school age children have attended grades 1 to 8 of school.

According to Alkire and Santos (2011), a person is deprived in nutrition if someone in their household is malnourished. On the other hand, child mortality indicators are considered as child deaths that are caused by infectious diseases or child malnutrition. The household is deprived if at least one child death has been observed in the household. There is room for improving this dimension in terms of child mortality. Our study did not use this indicator as the information needed was not present in the data set we used.

A household is said to have access to clean drinking water if the water source is any of the accompanying kinds: public tap, piped water, secured spring or rainwater, borehole or siphon, ensured well, and it is inside a distance of 30 minutes' walk (round trip). The household is deprived in water if it fails to satisfy these conditions (Alkire & Jahan, 2018).

A household is deprived if it does not have access to improved sanitation and in the event that it doesn't have some sort of flush latrine or lavatory, or ventilated further developed pit or composting latrine given that they are not shared. Deprivation in electricity occurs n the event that a household does not have access to electrical power (Alkire et al., 2011).

To consider a household as deprived in flooring if the ground surface material of the household is made of earth, dung manure or sand; and a household that cooks with compost, wood or charcoal is considered to be deprived in cooking fuel; whereas, on the off chance that a family unit doesn't claim more than one TV, radio, phone, motorbike or fridge, bicycle, and doesn't possess a farm hauler or a vehicle, it is considered as deprived (Alkire & Santos, 2011). Our study improved the assets indicator by dividing it into two indicators: (i) access to information/ICT (radio, TV...), (ii) access to transport (motorbike ...) which were used in the beta regression model to see if they are statistically significant to be included in the computation of MPI.

According to Alkire et al. (2011), a deprivation score is assigned to each person in accordance to his/her deprivation in the part indicators. Different people are faced with a number of different deprivations. Each individual's deprivation score lies between 0 and 1, which is obtained by taking a weighted total of the experienced deprivations. The score increments as the quantity of deprivations of the individual increments and when the individual is deprived in every one of the ten indicators the deprivation score reaches its maximum of 1. A score equal to zero is given to an individual who is not deprived in any indicator i.e.:

$$C_i = W_1 I_1 + W_2 I_2 + ... + W_d I_d, \tag{2.1}$$

where each individual's deprivation score is $I_i = \begin{cases} 1 & if\ household\ is\ deprived \\ 0 & otherwise \end{cases}$ and W_i is the weight attached to indicator i with $\sum_{i-1}^{d} W_i = 1$, and d is the total number of weight and indicator in a given dimension.

The global MPI utilises the cross-dimensional poverty cut-off of 33%, distinguishing every individual as poor if their weighted deprivations percentage is 33% or more. Two other poverty cut-offs are additionally utilised: serious poverty (the level of individuals deprived at any rate half of the weighted indicators) and weakness/vulnerability (the extent of individuals deprived in 20 to 33% of weighted indicators) (Alkire & Jahan, 2018). The poverty threshold is characterized as the portion of (weighted) deprivations a family should have to be viewed as multidimensionally poor.

If a person’s deprivation score is equal to or more than the deprivation threshold, then that person is considered multi-dimensionally poor. The person’s score is replaced by a ‘0’ if their deprivation score is less than the poverty threshold, even if it is non-zero, and any current deprivations are not considered in the ’censored headcounts’. This step is alluded to as censoring the deprivations of the non-poor (Alkire et al., 2011).

2.3 Modelling of the MPI

There has been a sluggish development in the literature on the modelling of multidimensional poverty. According to the Oxford Poverty and Human Development Initiative (2019b) the MPI joins two parts of poverty:

$$MPI = H \times A \tag{2.2}$$

The incidence (H) $\sim$ the level of individuals who are poor, or the poverty rate or headcount ratio and the intensity (A) $\sim$ the normal level of measurements wherein poor individuals are deprived, or the normal deprivation score of helpless people; of poverty to come up with the MPI.

Massuanganhe (2005) expanded the study of multidimensional poverty utilising a fuzzy set to present a blend of disintegration analysis. This model yielded the most significant components of poverty (health, education, and so on) and most applicable sub-groups (areas, gender, and so forth) so as to recognise the principle powers that add to the general measure of the condition of poverty. The analysis of these outcomes is valuable for chiefs that think about financial approaches for poverty decrease. This approach has a shortcoming as it takes the approach of uni-dimensional first.

Alkire et al. (2015) proposed regression models for modelling MPI, which can represent the impact or the ’size’ of determinants of multidimensional poverty, which would not be conceivable with an absolutely descriptive analysis. Such analyses are routinely performed for monetary poverty using what is termed ’micro’ or ’macro’ regressions.

The term 'micro' alludes to examinations in which the unit of investigation is an individual or household; the term 'macro' alludes to analysis in which the unit of examination is a subgroup, for example, a region, a state, a province, or a nation.

Generally, in micro regressions, the central variable to be modelled might be a double factor indicating an individual's status as poor (or non-poor) or a variable representing the deprivation score allocated to poor people. In macro regression, the central variable to demonstrate is a subgroup poverty measure like the poverty headcount ratio or some other multidimensional poverty measure (Alkire et al., 2015). Similarly as with the regression that models the monetary headcount ratio or poverty, macro regression with dependent variables must regard their tendency as cardinally important qualities going from zero to one. In these cases, a classic linear regression isn't the suitable model. The basic presumptions of the classic linear regression model miss the mark in light of the fact that the scope of the reliant variable is limited and may not be continuous or follow a normal distribution that is regularly accepted in linear regression models (Alkire et al., 2015).

Alkire et al. (2015) further noted that Generalized Linear Models (GLMs), on the other hand, are favoured as the information logical procedure since they represent the limited and discrete nature of the AlkireFoster-type dependent variable. In particular, GLMs stretch out classic linear regression to a group of regression models where the dependent variable might be ordinarily disseminated or may follow a distribution inside the exponential family, for example, the gamma distribution, beta distribution, Bernoulli distribution, or binomial distribution. Moreover, GLMs incorporate models for quantitative and qualitative dependent variables, for example, linear regression models, logit and probit models, and models for partial information (Alkire et al., 2015). Subsequently they offer an overall structure for our investigation of useful connections.

It is for the above reasons that our research study employed beta distribution to model multidimensional poverty for Namibia. The data used to measure poverty is continuous and lies in the range of zero and one which qualifies the beta distribution model as a reasonable way to model poverty as it is versatile/flexible and can be applied in a wide range of data analysis.

2.4 Modelling with beta distribution

Beta distribution is a group of continuous probability distributions characterized on the range between 0 and 1, parameterised by two positive shape parameters (α and β). The beta density function is a useful way to represent outcomes such as proportions or probabilities (Johnson et al., 1994).

2.4.1 Two parameter beta distribution model

The standard beta distribution provides the probability density of a value y on the interval (0, 1):

$$Beta(\alpha, \beta)(y) = prob(y; \alpha, \beta) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha, \beta)}, \tag{2.3}$$

where B is the beta function

$$B(\alpha, \beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1} dy \tag{2.4}$$

with mean;

$$\mu = \frac{\alpha}{\alpha + \beta} \tag{2.5}$$

and variance;

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{2.6}$$

The beta distribution is successfully used in modelling data measured continuously on the open interval (0, 1) such as estimating poverty rates, but it cannot be used in the data that contains zeros and/or ones (Ospina & Ferrari, 2010). When modelling the distribution for the probability of occurrence of some discrete event, beta distribution can be used (Johnson, 2013). A beta-2 distribution was used to model income distributions for developing countries in Asia, and the parameters were estimated using the method of moments procedure which was applied to grouped data. The estimated parameters of these distributions were then used to calculate measures of poverty, inequality, and pro-poor growth (Chotikapanich et al., 2014).

Estimation of parameters: for two-parameter beta

The maximum likelihood estimator: The maximum likelihood approach is a common method of estimating parameters. Owen (2008) defines the likelihood function for an independent and identically distributed sample $Y_1, ..., Y_n$ from a population with pdf $f(y \mid \theta_1, ..., \theta_k)$ as $\ell(\theta_1, ..., \theta_k \mid y_1, \ldots\ldots\ldots, y_n) = \prod_{i=1}^{n} f(y_i \mid \theta_1, ..., \theta_k)$. The maximum likelihood estimator (MLE) is the parameter value for which the observed sample is the best bet. Possible MLEs are solutions to $\frac{\partial}{\partial \theta_i} \log \ell(\theta \mid Y) = 0, i = 1, ..., k$. It may be verified that the points found are maxima, as opposed to minima using the first derivative. Furthermore, MLEs are desirable estimators due to their being consistent and asymptotically efficient; that is, they combine in probability to the parameter they are assessing and accomplish the lower bound on fluctuation.

Singh et al. (2014) observed that it is impossible to drive the Maximum Likelihood Estimator's (MLE's) precise distribution as the MLE is not obtained in closed form. Using asymptotic distribution of the MLE was proposed in constructing an approximate confidence interval. According to Schröder and Rahmann (2017), estimating parameters using maximum likelihood approach with beta distribution has a problem due to singularities in the log-likelihood function if some of the observations take the values 1 or 0. The approach of Combining latent variables with the method of moments was proposed for estimating parameters to mitigate the problem faced with the MLE that doesn't present the same problems as MLE (Schröder & Rahmann, 2017).

In the maximum likelihood estimation method, when there are no explanatory solutions, numerical methods are utilised to discover the root to the first derivative of the maximum likelihood estimation. This is where the Newton-Raphson method is used. By and large, the Newton-Raphson method is anything but difficult to actualise and to be dependable (Löffler & Posch, 2011). Since in the maximum likelihood estimation there is no closed-form solution to the type of system of equations, we employ the Newton-Raphson method to solve the system of equations obtained by the MLE method.

Newton-Raphson method:

Newton's methods (likewise recognised as the Newton–Raphson method), named after Isaac Newton and Joseph Raphson, is a method for judgment consecutively better approximations than the extraction (or zeroes) of a genuine esteemed function (Akram & Ann, 2015).

In this study, we are keen on computing maximum likelihood estimates (MLEs) for the beta dispersion. Such estimates are frequently very convoluted nonlinear elements of the observed data. Thus, closed form articulations for the MLEs will commonly not exist for these kinds of models (Quinn, 2001).

In this case, the Newton Raphson method is an iterative method that we can use to compute MLEs. The essential thought behind the computation is the accompanying. To begin with, build a quadratic estimation to the function of interest around some underlying boundary esteem (ideally near the MLE). Next, change the boundary incentive to that which expands the quadratic estimate. This method is iterated until the boundary esteems balance out (Quinn, 2001).

2.4.2 Generalised beta distribution with four-parameter model

According to Zaninetti (2013), a random variable Y with values for y in the interval [a, b] is a generalised beta distribution with the Probability Mass Function (PDF):

$$f_{ab}(y; a, b, \alpha, \beta) = \frac{(b-a)(y-a)^{\alpha-1}(b-y)^{\beta-1}}{b^{\alpha+\beta-1} b(\frac{b-a}{b})^{\alpha+\beta} B(\alpha, \beta)}, \tag{2.7}$$

where b greater than 0 is the scale parameter; a, α, and β are positive shape parameters. The parameter a represents the overall shape, α governs the left tail, and β the right tale.
The expected mean of the generalised beta distribution is:

$$(y; a, b, \alpha, \beta)_{ab} = \frac{\alpha b + a\beta}{(\alpha+\beta)} \tag{2.8}$$

and its Variance is:

$$\sigma(y; a, b, \alpha, \beta)^2_{ab} = \frac{(a-b)^2 \alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} \tag{2.9}$$

The generalised beta distribution is a four parameter distribution and it can be estimated by:

$$\bar{\alpha} = -\frac{(-\bar{y} + \bar{a})(-\bar{y}\bar{b} + \bar{b}\bar{a} + \bar{y}^2 + S^2 - \bar{a}\bar{y})}{S^2(\bar{a} - \bar{b})} \tag{2.10}$$

$$\bar{\beta} = -\frac{(\bar{b} - \bar{y})(-\bar{y}\bar{b} + \bar{b}\bar{a} + \bar{y}^2 + S^2 - \bar{a}\bar{y})}{S^2(\bar{a} - \bar{b})} \tag{2.11}$$

where $\bar{a}$ =Minimum of sample and $\bar{b}$ =Maximum of sample
The generalised beta distribution of the second order (GB2) can be a versatile and flexible distribution providing a good description to various types of data such as increasing, decreasing, uni-modal, uniantimodal, or bath-tub shape distribution depending on the parameter values (Ng et al., 2018). Furthermore, if the wrong distribution is used in estimation, problems of under estimation may arise.

Chen and Singh (2017) introduced generalised beta distribution of the second kind in flood frequency analysis (FFA). They employed the principle of the maximum entropy (POME) method to estimate the GB2 parameters. The result showed that the GB2 is an appealing distribution for FFA, since it has four parameters estimated by the POME method which are found reasonable. Our model took the form of beta regression since MPI is in the interval of (0, 1).

2.5 Modelling with beta regression

Regression models are typically used in dissecting information that is seen to be related to other variables (Ferrari & Cribari-Neto, 2004). Also, the Linear regression model has been utilized in numerous applications to inspect how certain components sway a steady factor that takes on qualities on the real line; such models require stringed expectation to be applied. Ferrari and Cribari-Neto (2004) proposed beta regression modelling and inferential techniques with the reaction conveyance being essential for the exponential family. In the proposed model, the response variable is the beta distribution using a parametrisation of the beta law that is ordered by mean and dispersion parameters.

The regression model presented by Ferrari and Cribari-Neto (2004) is portrayed by only two boundaries, and it is sufficiently versatile to manage a wide extent of uses. This model is versatile for showing data on the standard unit interval. Similarly, since its PDF can have exceptionally different shapes depending upon the assessments of the two parameters that rundown the regression, as it is eminent, the beta distribution is altogether versatile for modelling proportions.

The proposed model is valuable for circumstances where the variable of interest is nonstop and confined to the span (0; 1) and is identified with different variables through a regression structure. The regression parameters of the beta regression model are interpretable regarding the mean of the reaction and, when the logit connect is used, of an odds ratio, not at all like the parameters of a direct regression that utilises a changed reaction (Ferrari & Cribari-Neto, 2004). Also, Parameter estimation of this model is performed by maximum likelihood.

The PDF of beta indexed by α and β is given in Equation 3.10. It is ordinarily progressively helpful in regression analysis to model the mean of the reaction and to characterise the model with the goal that it contains a precision (or scattering) parameter. So as to get a regression structure for the mean of the response alongside a precision parameter, distinctive parametrisation of the beta density was used.

Letting $\mu = \frac{\alpha}{\alpha+\beta}$ and $\phi = \alpha + \beta$ i.e $\alpha = \mu\phi$ and $\beta = (1-\mu)\phi$. It follows from equations 2.5 and 2.6 that;

$$E(y) = \mu \tag{2.12}$$

and

$$Var(y) = \frac{V(\mu)}{1+\phi}, \tag{2.13}$$

where $V(\mu) = \mu(1-\mu)$ such that the mean of the response variable is μ, and ϕ can be deduced as a precision parameter in the sense that, for a fixed μ, as μ increases, the variance of y decreases.

In the new parametrisation, now the density function of beta distribution can be composed as;

$$f(y;\mu,\phi)=\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1},0<y<1 \tag{2.14}$$

where $0<\mu<1$ and $\phi>0$.

In order to fit the model, we first consider $y_1,..........,y_n$ to be independent random variables, with each $y_i, i=1,........,n$ following the density in equation 2.14 above which has a mean μ_t and an unknown precision parameter ϕ. Following Ferrari and Cribari-Neto (2004), the model was then achieved by assuming that the mean of y_i can be expressed as;

$$g(\mu_t)=\sum_{i=1}^{k}x_{ti}\beta_i=\eta_t, \tag{2.15}$$

where $\beta=(\beta_1,...........,\beta_k)^\tau$ is the vector of unknown regression parameters ($\beta\in\mathbb{R}^k$), $x_{t1},.......,x_{tk}$ are fitted k covariates and $g(\cdot)$ is a link function that maps (0; 1) into $\mathbb{R}$. It can be noted that the variance of y_i is a function of μ_t and, as a result of the covariate values (Ferrari & Cribari-Neto, 2004).

Ferrari and Cribari-Neto (2004) highlight that few potential decisions for the link function $g(\cdot)$ exist. For example, one can use the logit specification $g(\mu)=log\{\frac{\mu}{(1-\mu)}\}$; the probit function $g(\mu)=\Phi^{-1}(\mu)$, where $\Phi(\cdot)$ is the aggregate distribution function of a standard normal random variable; the complementary log-log interface $g(\mu)=log\{-log(1-\mu)\}$; and the log-log link $g(\mu)=-log\{-log(\mu)\}$.

Bonat et al. (2019) proposed another class of regression models for continuous restrained data. In their study, they emphasised that regardless of the distinctions in the mean and variance connection between the simplex and beta distributions, such a relationship can all around be modelled by a basic function of the expected values.

This reality propels an indication of a regression model by using just second-moment presumptions. Therefore, they considered a cross-sectional data-set, $(y_i, x_i), i = 1, ..., n$, where y_is are independent and identically distributed (iid) realisations of Y_i as indicated by an unknown appropriation, whose expectation and variance are given by,

$$E(Y_i) = \mu_i = g^{-1}(x_i^\tau \beta), \tag{2.16}$$

and

$$Var(Y_i) = \phi \mu_i^p (1 - \mu_i)^p, \tag{2.17}$$

where x_i and β are $(q \times 1)$ vectors of known covariates and regression parameters, respectively. Additionally, g is a standard link function. Our study utilised this type of model that was proposed by Ferrari and Cribari-Neto (2004) to help us identify covariates of our MPI model, More specifically, to help us identify possible significant indicators that may have been left out in the computation of MPI.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This clarifies how the research study was carried and how methods were utilized to address the objectives. This section further portrays the information sources, statistical data analysis procedures, and information assortment strategies that were utilized to deliver the conclusive results.

3.2 Data description

Our research study used secondary data on Namibia Household Income and Expenditure Survey (NHIES) for 2015/2016 which was obtained from the Namibia Statistics Agency (NSA) to attain the desired results as it was the best data which had most of the variables of interest. The data covered the country's 14 regions. A sample of 864 Primary Sampling Units (PSUs) consisting of 10 386 household units was chosen over the 14 regions through two-stage sampling method as indicated in Table 3.1. However, only 10090 households were used in this study as they had complete information.

A stratified two-stage cluster sample where the unit for first stage was the geographical areas assigned as PSUs which was chosen from the National Sampling Frame (NSF) of PSUs, and the second stage unit was the households which were chosen from a list made in the field during the listing procedure done before interviewing household units inside the PSU was used (Namibia Statistics Agency, 2015).

Source: Namibia Statistics Agency (2015)
TABLE 3.1: Distribution of sample PSUs and households by region

Region	**PSU**	**Sample Households per PSU**	**Total sample households**
!Karas	48	12	576
Erongo	72	12	864
Hardap	48	12	576
Kavango East	48	12	576
Kavango West	48	12	576
Khomas	96	12	1152
Kunene	48	12	576
Ohangwena	72	12	864
Omaheke	48	12	576
Omusati	72	12	864
Oshana	72	12	864
Oshikoto	72	12	864
Otjozondjupa	72	12	864
Zambezi	48	12	576
Total	**864**		**10 368**

There were two sets of data in this survey. The Household data set which covered the frequent and infrequent transactions of durable goods and services acquired and consumed and also the household income, and individual data set which covered demographic, education, health, labour force information and income, etc. The data contained more parameters most of which were not used in this study. The parameters that we used from the two data sets included: the education and literacy domain, health domain, and the living standards domain (Assets, housing type, housing costs, and access to services). The variables in the health domain (child mortality and nutrition) data for these variables have not been collected in our data sets the way our MPI theoretical framework requires, therefore, we used proxies from the household data set such as food adequacy in the past 7 days (q02_61) for nutrition. However, in our data set, there was no data for child mortality and we didn't use any. In the individual data set, we identified the variables for education domain (years of schooling and school attendance). For the years of schooling we identified the variables from the individual data set that we used: Age in years (q01_06_y), and years of study past enrollee (q03_05). Where as for the school attendance variable we used: Age in years (q01_06_y), and currently attending school yn (q03_03). The variables for the living standards domain (cooking fuel, drinking water, sanitation, electricity, flooring, and assets) were obtained from the household data set.

For cooking fuel, the variable we used is the main source of energy used for cooking (q02_07) from the household data set. For sanitation we used the following from the household data set: types of toilet facility used by the household members (q02_16), and do you share the toilet facility with other households (q02_17). For the drinking water variable we used: the main source of drinking water (q02_10), and do you treat water in anyway to make it safe for drinking (q02_14). For electricity we used: The main source of energy used for lighting (q02_09). For flooring we used: The main material used for the floor (q02_04). For the assets we used the following variable from the household data set: Number of owned motor car, station wagon (q05_02_01), number of owned motor cycles/scooters (q05_02_04), number of owned bicycles (q05_02_05), number of owned refrigerators (q05_02_09), number of owned radios (q05_02_13), number of owned television (q05_02_15), number of owned cellular telephones (q05_02_19), number of owned tractors (q05_02_29). All these were the variables we identified and used in our study as shown in Table 3.2. Other potential variables included in this study were identified. Variables like health status/condition, access to social services, consumption and social demographic characteristics were all identified from the data sets.

3.3 Data processing

After identifying the variables of interest from the two data sets, data cleaning was done to identify any outliers or missing data and the appropriate methods in dealing with them such as; approaching live data transformation omission/deletion were used. Variables that we were not interested in or going to be used in this study were deleted from the two data sets, leaving only the variables of interest. The two data sets were in Statistical Package for the Social Sciences (SPSS). Our unit of analysis was the household. The individual dataset was merged in SPSS with the household dataset to connect them to their respective household information. The merged dataset only included the potential variables and the variables that were used in the study. The indicators that were not available in the dataset as defined in the MPI requirements were estimated or thin proxy variables used for the construction of the model.

TABLE 3.2: How MPI variables are Identified in the data sets

MPI indicators	How it is identified	Individual data set	Household data set
Nutrition	Food adequacy in the past 7 days (q02_61)		X
Years of schooling	Age in years (q01_06_y), and years of study past enrollee (q03_05)	X	
School attendance	Age in years (q01_06_y), and currently attending school yn (q03_03)	X	
Cooking fuel	Main source of energy used for cooking (q02_07)		X
Sanitation	Types of toilet facility used by the household members (q02_16), and do you share the toilet facility with other households (q02_17)		X
Drinking water	The main source of drinking water (q02_10), and do you treat water in anyway to make it safe for drinking (q02_14)		X
Electricity	The main source of energy used for lighting (q02_09)		X
Floor	The main material used for the floor (q02_04)		X
Assets	Number of owned motor car, station wagon (q05_02_01), number of owned motor cycles/scooters (q05_02_04), number of owned bicycles (q05_02_05), number of owned refrigerators (q05_02_09), number of owned radios (q05_02_13), number of owned television (q05_02_15), number of owned cellular telephones (q05_02_19), number of owned tractors (q05_02_29)		X

3.4 Delimitation of the study

This study only used NSA data from the Namibia Household Income and Expenditure Survey (NHIES) for 2015/2016, and not any other years or other type of sources such as census or Demographic and Health Survey (DHS) data. The NHIES data was used to achieve the ideal outcomes as it was the best data which has most of the variables we desired. In addition, it made it easier to compare other countries due to the variables included in this dataset. Also, this data set is more recent as compared to all other possible datasets such as the 2011 population and housing census and demographic and health survey DHS (2013).

3.5 Ethical issues

Chetty (2016) states that the researcher needs to hold fast to advance the points of the research granting legitimate information, truth and anticipation of mistakes while the inquiries with respect to the moral system ought to likewise be settled. In fact, the researcher ought to guarantee welfare and dignity of the subjects and safeguarding the gathered data. The researcher conducted himself in a professional manner. The data and the responses remained confidential.

3.6 Determination of deprivation scores

In order to determine the deprivation scores, we first needed to determine the weights which were then used to compute the deprivation scores for each household. In this study, we used two weighting methods namely the equal weighting method which was proposed by Alkire and Santos (2011), and the entropy weighting method. We denoted the deprivation scores of a specific household by y_i where i is the number of the household.

3.6.1 Determination of weights

According to Kadapa (2020), determining weights can be classified as subjective, objective or integrated. The equal weighting is an example of Subjective weighting. In Subjective weighting, the role of allocating weights solely lies on the decision maker who is an expert in the field and the allocation is done based on previous experience. This brings constraints of designer preferences (Kadapa, 2020).

On the other hand, in the Objective weighting, the decision maker has no role in the determination of weights and it is more useful when the decision maker does not exist (Kadapa, 2020). Our research study used both Subjective (equal weighting) and Objective (Entropy) weighting methods. These methods were used to allocate weights to indicators and hence, used the weights obtained in both methods to compute the deprivation scores. The weighting methods utilised in our study are discussed in detail below.

(i) Equal-weighting Method

The three domains of MPI are equally weighted such that each domain receives a weight of $\frac{1}{3}$. The weights W_h were determined using equal weighting allocation with $\sum_{h=1}^{m} W_h = 1$ where m is the total number of indicators as proposed by Alkire and Santos (2011). Within each domain, we considered indicators of deprivations; two for education, one for health, and six for living standards. These indicators are also equally weighted such that each indicator in health receives a weight of $\frac{1}{2}$, each indicator in education also receives a weight of $\frac{1}{2}$, and indicators in the living standard domain were given a weight of $\frac{1}{6}$ each. This results now in the final weights for each indicator which are obtained by multiplying the domain weight with the indicator weight, such that the indicators in the health domain receive an equal weight of $\frac{1}{3}$ and the indicators in the education domain receive equal weights of $\frac{1}{6}$ each, the indicators in the living standard domain also receive equal weights of $\frac{1}{18}$ as shown in Table 3.3.

TABLE 3.3: Equal weights allocation of Dimensions and Indicators of MPI

Domain	Domain Weight	Indicator	Indicator Weight	Final Indicator weight
Health	$\frac{1}{3}$	Nutrition	1	$\frac{1}{3} \times 1 = \frac{1}{3}$
Education	$\frac{1}{3}$	Years of Schooling	$\frac{1}{2}$	$\frac{1}{6}$
		School Attendance	$\frac{1}{2}$	$\frac{1}{6}$
Living Standard	$\frac{1}{3}$	Cooking Fuel	$\frac{1}{6}$	$\frac{1}{18}$
		Sanitation	$\frac{1}{6}$	$\frac{1}{18}$
		Drinking Water	$\frac{1}{6}$	$\frac{1}{18}$
		Electricity	$\frac{1}{6}$	$\frac{1}{18}$
		Floor	$\frac{1}{6}$	$\frac{1}{18}$
		Assets	$\frac{1}{6}$	$\frac{1}{18}$

(ii) Entropy-weighting method

The weights W_h^* were determined using the entropy weight method proposed by Gao and Son (2020). In their study, certain steps are followed in the weighting process. Assuming that there are k domains/dimensions to be evaluated and g evaluation index/indicators, and I_j is the j^{th} indicator where j runs from 1 up to h, and h is the total number indicators of the i^{th} household, the basic data matrix $M = (I)_{k\times h}$ where k is the number of households in the domains, which can then be expressed as follows;

$$M = \begin{pmatrix} I_{1,1} & I_{1,2} & \cdots & I_{1,h} \\ I_{2,1} & I_{2,2} & \cdots & I_{2,h} \\ \vdots & \vdots & \ddots & \vdots \\ I_{k,1} & I_{k,2} & \cdots & I_{k,h} \end{pmatrix}_{k\times h}$$

The weight of the index value for the j^{th} indicator is represented by P_{ij}, the entropy value of the j^{th} indicator is expressed by e_j, and the entropy weight of the indicator is represented by W_j^*. The calculation steps are illustrated as follows;

$$P_{ij} = \frac{I_j}{\sum_{i=1}^{m} I_j} \tag{3.1}$$

$$e_j = -z \sum_{i=1}^{k} P_{ij} \cdot \ln p_{ij} \ , \tag{3.2}$$

where $z = 1/\ln(k)$

$$W_j^* = \frac{(1 - e_j)}{\sum_{j=1}^{h} (1 - e_j)} \tag{3.3}$$

Therefore, the deprivation scores were also calculated using the entropy weight method as follows;

$$y_i^* = \sum_{j=1}^{h} W_j^* I_j \tag{3.4}$$

(Gao & Sun, 2020).

3.6.2 Computation of deprivation scores

Once we determined the weights, the next step was to compute the deprivation scores of each household.

The deprivation scores were computed as a function of indicators as follows. Let y_i be the deprivation score for each household, I_j be the j^{th} indicator for i^{th} individual, and W_j are the weights of indicator I_j. Then,

$$y_i = \sum_{i=1}^{h} W_j I_j, \ i = 1, 2, 3, ..., k \ and \ j = 1, 2, 3.........., h \ , \tag{3.5}$$

where k is the total number of households and h is the total number of indicators.

A household is considered to be deprived if $y_i \geq p$ (where $p \in (0,\ 1)$). Otherwise, a household with a deprivation score below the threshold is considered non-deprived and its deprivation score is equated to zero.

This whole process is referred to as censoring (Santos & Alkire, 2011). We denote the censored deprivation scores by $y_i(p)$. That is,

$$y_i(p) = \begin{cases} y_i \; if \; y_i \geq p \\ 0 \; otherwise \end{cases} \tag{3.6}$$

We created a new variable in the merged data set called $y_i(p)$ which provides the deprivation values of households. The merged dataset that was generated was used to carry out some descriptive analysis, fit the distribution(beta), fit a distribution of $y_i(p)$ using kernel density and check if it was similar to beta. We also estimated the parameters of the beta distribution. This approach of computation of deprivation scores was done both for equal weights and entropy weights.

3.7 Computation of MPI

After computing the deprivation scores of individuals, allocating weights, determining deprivation cut-offs, then we we computed the MPI. As stated in the literature review, MPI combines two key components,namely the incidence of people with multiple deprivations and the intensity of their deprivation. The incidence is referred to as the multidimensional headcount ratio, denoted by H and calculated as follows;

$$H = \frac{q}{n}, \tag{3.7}$$

where q is the number of people who are multidimensionally poor and n is the total population. The second component is referred to as the breadth of poverty, denoted by A. This was computed as the average deprivation score of the household who were multidimensionally poor and this was computed as follows;

$$A = \frac{\sum_{i=1}^{n} y_i(p)}{q}, \tag{3.8}$$

where $y_i(p)$ is the censored deprivation score of individual i and q is the number of households who are multidimensionally poor (Alkire & Santos, 2011).

Therefore, MPI is the product of the incidence (H) and the intensity (A):

$$MPI = H \times A \tag{3.9}$$

In this study, we considered two approaches of computing MPI: one is when we used the equal weight method which we denoted by MPI; and the other one is when we used the entropy weight method which we denoted by MPI^*.

3.8 Modelling deprivation scores using beta distribution

Let y_i be the multidimensional deprivation score for each household. We assume $y_i \sim B(\alpha, \beta)$ because it takes values in the interval (0, 1). According to Johnson et al. (1994), the probability density function of beta distribution with parameters $\alpha > 0$ and $\beta > 0$ is given by:

$$f(y_i; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} y_i^{\alpha-1}(1-y_i)^{\beta-1}, 0 < y_i < 1 \tag{3.10}$$

where $B(\alpha, \beta)$ is the beta function defined in terms of the gamma function as $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$; α and β are the shape parameters. With

$$\mu = \frac{\alpha}{\alpha + \beta} \tag{3.11}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{3.12}$$

To get the mean μ for beta, we derive the first moment about the origin, and to obtain the variance σ^2 for beta, we find the second moment about the origin which was used together with the first moment to compute the variance. This was done as follows;

$$\begin{aligned} \mu_r &= \frac{1}{B(\alpha, \beta)} \int_0^1 y^{\alpha-1}(1-y)^{\beta-1} y^r dy \\ &= \frac{1}{B(\alpha, \beta)} \int_0^1 y^{(\alpha+r)-1}(1-y)^{\beta-1} dy \end{aligned} \tag{3.13}$$

Simplifying μ_r in terms of beta gives;

$$\mu_r = \frac{B(\alpha + r, \beta)}{B(\alpha, \beta)} \tag{3.14}$$

rewriting μ_r in terms of gamma gives;

$$\begin{aligned}
\mu_r &= \frac{\Gamma(\alpha + r)\Gamma(\beta)}{\Gamma(\alpha + \beta + r)} \times \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
&= \frac{\Gamma(\alpha + r)}{\Gamma(\alpha + \beta + r)} \times \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)}
\end{aligned}$$

$$\begin{aligned}
\text{Setting r} &= 1, 2 \\
\mu_1 &= \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha + \beta + 1)} \times \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \\
&= \frac{\alpha}{\alpha + \beta} \\
\text{using}\Gamma k &= (k-1)\Gamma(k-1) \\
\mu_2 &= \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + \beta + 1)} \times \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \\
E(Y^2) &= \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)}
\end{aligned} \tag{3.15}$$

To derive the variance, we use the fact that;

$$\begin{aligned}
\sigma^2 &= \mu_2 - (\mu_1)^2 \\
\sigma^2 &= \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} - (\frac{\alpha}{\alpha + \beta})^2 \\
&= \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}
\end{aligned} \tag{3.16}$$

Therefore, the mean and variance is what we have in 2.5 and 2.6.

3.8.1 Estimation of parameters (α, β)

(a) Maximum likelihood estimation

To obtain the MLE of the beta distribution, we find the log likelihood of the pdf of the distribution in equation 3.10;

$$\ell(\alpha, \beta) = \prod_{i=1}^{n} f(y_i; \alpha, \beta) \tag{3.17}$$

where y_i are independent variables for $i = 1, 2, \ldots .n$ and n is the total number of households considered in this study. It then follows that;

$$\begin{aligned}
\ell(\alpha, \beta) &= \log \ell(\alpha, \beta) \\
\ell(\alpha, \beta) &= \log \prod_{i=1}^{n} f(y_i; \alpha, \beta) \\
&= \sum_{i=1}^{n} \log f(y_i; \alpha, \beta) \\
&= \sum_{i=1}^{n} \log \frac{1}{B(\alpha, \beta)} y_i^{\alpha-1} (1 - y_i)^{\beta-1}
\end{aligned} \tag{3.18}$$

Rewriting $\frac{1}{B(\alpha,\beta)}$ as $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ gives;

$$\begin{aligned}
&= \sum_{i=1}^{n} \log \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y_i^{\alpha-1} (1 - y_i)^{\beta-1} \\
&= \sum_{i=1}^{n} [\log \Gamma(\alpha+\beta) - \log \Gamma(\alpha) - \log \Gamma(\beta) + (\alpha - 1) \log y_i + (\beta - 1) \log (1 - y_i)] \\
&= n \log (\Gamma(\alpha+\beta)) - n \log (\Gamma(\alpha)) - n \log (\Gamma(\beta)) + (\alpha - 1) \sum_{i=1}^{n} \log y_i + (\beta - 1) \sum_{i=1}^{n} \log (1 - y_i)
\end{aligned} \tag{3.19}$$

The MLEs of α and β are obtained by taking the partial derivatives of the log likelihood with respect to each parameter and then equating them to zero and then solving for $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$. This is illustrated below;

$$\frac{\partial}{\partial \alpha} \log \ell(\alpha, \beta) = \frac{n\Gamma'(\alpha+\beta)}{\Gamma(\alpha+\beta)} - \frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n} \log (y_i) = 0 \tag{3.20}$$

$$\frac{\partial}{\partial \beta} \log \ell(\alpha, \beta) = \frac{n\Gamma'(\alpha+\beta)}{\Gamma(\alpha+\beta)} - \frac{n\Gamma'(\beta)}{\Gamma(\beta)} + \sum_{i=1}^{n} \log (1 - y_i) = 0 \tag{3.21}$$

Noting that the digamma function is given by $\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} = \psi(\alpha)$ and dividing equations 3.20 and 3.21 with n , this can be shown in 3.22 and 3.23 below;

$$\frac{\partial}{\partial \alpha} \log \ell(\alpha, \beta) = \psi(\alpha+\beta) - \psi(\alpha) + \frac{1}{n} \sum_{i=1}^{n} \log (y_i) = 0 \tag{3.22}$$

$$\frac{\partial}{\partial \beta} \log \ell(\alpha, \beta) = \psi(\alpha+\beta) - \psi(\beta) + \frac{1}{n} \sum_{i=1}^{n} \log (1 - y_i) = 0 \tag{3.23}$$

Owen (2008)) notes that "There is no closed-form solution to this type of system of equations." We solve for $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$ using the Newton-Raphson method. The system of equations we get from the MLE is solved using the Newton-Raphson method.

(b) Newton-Raphson method

The Newton-Raphson method is one of the widely-used methods for solving a non-linear univariate function $f(y_i)$ on the interval [a, b]. In this case, estimation of $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ is done iteratively;

$$\hat{\theta}_{k+1} = \hat{\theta}_k - J^{-1} f \tag{3.24}$$

where $\hat{\theta}_k$ is the initial approximation, k is the number of iterations and f is the vector of normal equations we solved

$$f = \begin{bmatrix} \psi(\alpha+\beta) - \psi(\alpha) + \frac{1}{n}\sum_{i=1}^{n} \log(y_i) & \psi(\alpha+\beta) - \psi(\beta) + \frac{1}{n}\sum_{i=1}^{n} \log(1-y_i) \end{bmatrix}$$

and J is the matrix of second derivatives;

$$J = \begin{bmatrix} \psi'(\alpha+\beta) - \psi'(\alpha) & \psi'(\alpha+\beta) \\ \psi'(\alpha+\beta) & \psi'(\alpha+\beta) - \psi'(\beta) \end{bmatrix}$$

where $\psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ and $\psi'(\alpha) = \frac{\Gamma''(\alpha)}{\Gamma(\alpha)} - \frac{\Gamma'(\alpha)^2}{\Gamma(\alpha)^2}$ are the di- and tri-gamma functions.
It follows that;

$$J^{-1} = \frac{1}{Z} \begin{bmatrix} \psi'(\alpha+\beta) - \psi'(\beta) & -\psi'(\alpha+\beta) \\ -\psi'(\alpha+\beta) & \psi'(\alpha+\beta) - \psi'(\alpha) \end{bmatrix}$$

$$J^{-1} f = \frac{1}{Z} \begin{bmatrix} \psi'(\alpha+\beta) - \psi'(\beta) & -\psi'(\alpha+\beta) \\ -\psi'(\alpha+\beta) & \psi'(\alpha+\beta) - \psi'(\alpha) \end{bmatrix} \begin{bmatrix} H & L \end{bmatrix}$$

where Z(the determinant of J) $= (\psi'(\alpha+\beta) - \psi'(\alpha))(\psi'(\alpha+\beta) - \psi'(\beta)) - (\psi'(\alpha+\beta))(\psi'(\alpha+\beta))$;
$H = \psi(\alpha+\beta) - \psi(\alpha) + \frac{1}{n}\sum_{i=1}^{n} \log(y_i)$; and $L = \psi(\alpha+\beta) - \psi(\beta) + \frac{1}{n}\sum_{i=1}^{n} \log(1-y_i)$
The Newton-Raphson algorithm converges as the estimates of α and β change by less than

any acceptable amount with each consecutive iteration, to $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$. For Newton-Raphson, we will need an initialisation and we use the method of momemts to get the initialisation.

(c) Method of moments

We used the method of moments to generate a starting point of the Newton-Raphson optimisation.

Method of moment estimates of α and β were obtained by setting the sample mean $\bar{Y}$ and variance S^2 equal to the population mean and variance in 3.11 and 3.12 respectively. We then solved for α and β in terms of $\bar{Y}$ and S^2. That is,

$$\frac{\alpha}{\alpha+\beta} = \bar{Y} \tag{3.25}$$

$$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = S^2 \tag{3.26}$$

From Equation 3.25,

$$\begin{aligned}(\alpha+\beta)\bar{Y} &= \alpha \\ \Rightarrow \beta\bar{Y} &= \alpha - \alpha\bar{Y} \\ \Rightarrow \beta &= \frac{\alpha}{\bar{Y}} - \alpha\end{aligned} \tag{3.27}$$

Now we substitute Equation 3.27 into Equation 3.26 and solve for α.

$$\begin{aligned}
\alpha\beta &= (\alpha+\beta)^2(\alpha+\beta+1)S^2 \\
\Rightarrow \frac{\alpha^2}{\bar{Y}} - \alpha^2 &= (\alpha + \frac{\alpha}{\bar{Y}} - \alpha)^2(\alpha + \frac{\alpha}{\bar{Y}} - \alpha + 1)S^2 \\
\Rightarrow \frac{\alpha^2}{\bar{Y}} - \alpha^2 &= (\frac{\alpha}{\bar{Y}})^2(\frac{\alpha}{\bar{Y}} + 1)S^2 \\
\Rightarrow \alpha^2(\frac{1}{\bar{Y}} - 1) &= \alpha^2(\frac{1}{\bar{Y}^2})(\frac{\alpha}{\bar{Y}} + 1)S^2 \\
\Rightarrow (\frac{1}{\bar{Y}} - 1) &= (\frac{1}{\bar{Y}^2})(\frac{\alpha}{\bar{Y}} + 1)S^2 \\
\Rightarrow (\frac{1}{\bar{Y}} - 1)(\frac{\bar{Y}^2}{S^2}) &= \frac{\alpha}{\bar{Y}} + 1 \\
\Rightarrow (\frac{\bar{Y}(1-\bar{Y})}{S^2}) &= \frac{\alpha}{\bar{Y}} + 1 \\
\alpha &= \bar{Y}(\frac{\bar{Y}(1-\bar{Y})}{S^2} - 1)
\end{aligned} \tag{3.28}$$

Expressing β in terms of $\bar{Y}$ and S^2 yields:

$$\begin{aligned}
\beta &= \frac{\alpha}{\bar{Y}} - \alpha \\
\Rightarrow \beta &= \alpha(\frac{1-\bar{Y}}{\bar{Y}}) \\
\Rightarrow \beta &= (\frac{1-\bar{Y}}{\bar{Y}})\bar{Y}(\frac{\bar{Y}(1-\bar{Y})}{S^2} - 1) \\
\Rightarrow \beta &= (1-\bar{Y})(\frac{\bar{Y}(1-\bar{Y})}{S^2} - 1)
\end{aligned} \tag{3.29}$$

Therefore, the method of moments (MOM) estimates of α and β are given by;

$$\alpha_{MOM} = \bar{Y}(\frac{\bar{Y}(1-\bar{Y})}{S^2} - 1)$$

$$\beta_{MOM} = (1-\bar{Y})(\frac{\bar{Y}(1-\bar{Y})}{S^2} - 1)$$

According to Glen (2015), the method of moments are frequently used because they often involve basic calculations, unlike the maximum likelihood method which can be cumbersome. However, the parameter estimates in this method might be inaccurate. Also, this method may not bring about sufficient statistics. Hence, we used the method of moments as the initialisation of the Newton-Raphson.

Once we obtained the estimates of α and β, then we computed the estimate MPI for Namibia as

$$\hat{\mu} = \frac{\hat{\alpha}}{(\hat{\alpha} + \hat{\beta})} \tag{3.30}$$

With estimated variance given by

$$\hat{\sigma}^2 = \frac{\hat{\alpha}\hat{\beta}}{(\hat{\alpha} + \hat{\beta})^2(\hat{\alpha} + \hat{\beta} + 1)} \tag{3.31}$$

3.9 General beta distribution

Consider now a general beta distribution. According to Zaninetti (2013), if Y is a random variable with values for y_i in the interval [a, b], then the PDF of Y becomes

$$f_{ab}(y_i(p); a, b, \alpha, \beta) = \frac{(b-a)(y_i(p)-a)^{\alpha-1}(b-y_i(p))^{\beta-1}}{b^{\alpha+\beta-1} b(\frac{b-a}{b})^{\alpha+\beta} B(\alpha, \beta)}; a < y_i(p) < b, \tag{3.32}$$

where α and β are the shape parameters, a and b are the lower and upper bounds. Its expected mean is given as

$$E(Y; a, b, \alpha, \beta)_{ab} = \frac{\alpha b + a\beta}{\alpha + \beta}, \tag{3.33}$$

and its variance is given by

$$\sigma^2(Y; a, b, \alpha, \beta)_{ab} = \frac{(a-b)^2 \alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}. \tag{3.34}$$

It is possible to achieve 3.10 from 3.32 by replacing a=0 and b=1. In poverty modelling using MPI, we consider the deprivation scores from {p to 1}. The individuals with deprivation scores <P are considered non-poor. Let Y be the deprivation scores and Y take values between [P, 1]. From 3.32 we can generate the PDF of Y as shown below. This is called the left truncated beta distribution.

$$f_p(y_i(p); p, \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(y_i(p) - p)^{\alpha-1}(1 - y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta-1}}; p \le y_i(p) < 1 \tag{3.35}$$

The expected value of Y is given as

$$E(Y; p, \alpha, \beta) = \frac{\alpha + p\beta}{\alpha + \beta}, \tag{3.36}$$

and its variance is given by

$$\sigma^2(Y;p,\alpha,\beta) = \frac{p^2\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}. \tag{3.37}$$

where p is the left limit which is the multidimensional deprivation threshold.
We used the PDF of the left truncated beta distribution in equation 3.35 to model MPI.

3.10 Modelling MPI using beta distribution

Considering deprivation scores (y_i) as random variable realisations; $0 < y_i < 1$, and if the value of (y_i) is less than p (the threshold value) we equate the values of (y_i) to zeros. Thus, y_i can be modelled using a left truncated beta distribution.
From Equation 3.32, we obtain the left truncated beta distribution in Equation 3.38 by replacing a with p and b with 1, where p is the deprivation threshold value, α and β are positive shape parameters.

$$\begin{aligned} f(y_i(p);p,1,\alpha,\beta) &= \frac{(1-p)(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta}B(\alpha,\beta)} \\ &= \frac{1}{B(\alpha,\beta)}\frac{(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta-1}} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta-1}}; p \leq y_i(p) < 1 \end{aligned} \tag{3.38}$$

rewriting Equation 3.38 we obtain Equation 3.39

$$\begin{aligned} f(y_i(p)) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta-1}} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha-1}(1-p)^{\beta-1}(1-p)^{1}} \\ &= \frac{1}{(1-p)} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{(y_i(p)-p)^{\alpha-1}}{(1-p)^{\alpha-1}} \cdot \frac{(1-y_i(p))^{\beta-1}}{(1-p)^{\beta-1}} \end{aligned} \tag{3.39}$$

The expected value of the left truncated beta distribution can be derived from Equation 3.38 as follows;

$$E(Y) = \int_p^1 y_i(p) f(y_i(p)) dy_i(p) \tag{3.40}$$

we first transform Equation 3.40 by using Equation 3.39 and obtain the transformation below;

$$
\begin{aligned}
x_i &= \frac{y_i(p) - p}{1-p} \Rightarrow \text{if } y_i(p) = p \Rightarrow x_i = 0 & \quad 1 - x_i &= 1 - \frac{y_i(p) - p}{1-p} \\
y_i(p) &= p + x_i(1-p) & &= 1 - \frac{(1-p) - (y_i(p) - p)}{1-p} \\
dy_i(p) &= (1-p)dx_i & &= \frac{1 - y_i(p)}{1-p}
\end{aligned}
$$

Now we can easily compute the expected value of the left-truncated beta distribution as follows;

$$
\begin{aligned}
E(Y) &= \int_p^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot y_i(p) \cdot \frac{(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta-1}} dy_i(p) \\
&= \int_0^1 \frac{1}{1-p} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (p + x_i(1-p)) x_i^{\alpha-1}(1-x_i)^{\beta-1}(1-p)dx_i \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (p + x_i(1-p)) x_i^{\alpha-1}(1-x_i)^{\beta-1} dx_i \qquad (3.41) \\
&= p \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1}(1-x_i)^{\beta-1} dx_i \\
&\quad + (1-p) \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{(\alpha+1)-1}(1-x_i)^{\beta-1} dx_i
\end{aligned}
$$

We note that;

$$
f(x_i) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1}(1-x_i)^{\beta-1}
$$

is a general beta distribution in the (0, 1) interval, and hence; (3.42)

$$
f(x_i) = \int_0^1 f x_i dx_i = 1
$$

Substituting Equation 3.42 into Equation 3.41 gives;

$$\begin{aligned}
&= p\int_0^1 f(x_i)dx_i + (1-p)\int_0^1 x_i f(x_i)dx_i \\
&= p + (1-p)\cdot\frac{\alpha}{\alpha+\beta} \\
&= p + \frac{\alpha(1-p)}{\alpha+\beta} \\
&= \frac{\alpha(1-p)+p(\alpha+\beta)}{\alpha+\beta} \\
&= \frac{\alpha - p\alpha + p\alpha + p\beta}{\alpha+\beta} \\
&= \frac{\alpha+p\beta}{\alpha+\beta}
\end{aligned} \tag{3.43}$$

To obtain the variance we first compute $E(Y^2)$ as follows;

$$\begin{aligned}
E(Y^2) &= \int_p^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot y_i(p)^2\cdot\frac{(y_i(p)-p)^{\alpha-1}(1-y_i(p))^{\beta-1}}{(1-p)^{\alpha+\beta-1}}dy_i(p) \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot(p+x_i(1-p))^2\cdot x_i^{\alpha-1}(1-x_i)^{\beta-1}dx_i \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot(p^2+2p(1-p)x_i+(1-p)^2x_i^2) \\
&\cdot x_i^{\alpha-1}(1-x_i)^{\beta-1}dx_i \\
&= p^2\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot x_i^{\alpha-1}(1-x_i)^{\beta-1}dx_i \\
&+ 2p(1-p)\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot x_i^{(\alpha+1)-1}(1-x_i)^{\beta-1}dx_i \\
&+ (1-p)^2\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot x_i^{(\alpha+2)-1}(1-x_i)^{\beta-1}dx_i \\
&= p^2 + 2p(1-p)E(X) + (1-p)^2E(X^2) \\
E(Y^2) &= p^2 + 2p(1-p)\frac{\alpha}{\alpha+\beta} + (1-p)^2\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}
\end{aligned} \tag{3.44}$$

From Equations 3.44 and 3.43, its variance is obtained as follows;

$$
\begin{aligned}
Var(Y) &= E(Y^2) - E(Y)^2 \\
&= p^2 + 2p(1-p)\frac{\alpha}{\alpha+\beta} + (1-p)^2\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \\
&- p^2 + (1-p)^2 \cdot \frac{\alpha^2}{(\alpha+\beta)^2} \\
&= (1-p)^2\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{(1-p)^2\alpha^2}{(\alpha+\beta)^2} \\
&= \frac{(1-p)^2\alpha}{(\alpha+\beta)}\left[\frac{\alpha(\alpha+1)}{(\alpha+\beta+1)} - \frac{\alpha}{(\alpha+\beta)}\right] \\
\sigma^2(Y) &= \frac{(p-1)^2\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}
\end{aligned}
\tag{3.45}
$$

Therefore, its Expected value is given by;

$$E(Y) = \frac{\alpha + p\beta}{\alpha+\beta}, \tag{3.46}$$

and its variance is given by;

$$\sigma^2(Y) = \frac{(p-1)^2\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} \tag{3.47}$$

Modelling the deprivation scores y_i using left truncated beta distribution, we obtain MPI which is equivalent to the expected value of Y, hence the following proposition.

3.11 Identification of potential indicators for MPI using beta regression

To identify potential indicators to include in MPI, we used the beta regression model that was proposed by Ferrari and Cribari-Neto (2004).

In this study, we modelled deprivation scores (y_i) using the beta regression through the function of the mean of y_i expressed as follows:

$$g(\mu_t) = \sum_{i=1}^{k} \beta_i x_{ti} = \eta_t \tag{3.48}$$

with x_{ti} being the value of the i^{th} indicator and β_i is the regression parameter corresponding to indicator i.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter presents the data analysis and the discussions of the study findings. The main objective of the study was to develop a multidimensional poverty model for Namibia with improved accuracy and inclusivity in the poverty measure using beta distribution.

4.2 Descriptive statistics

This section presents the descriptive statistics through graphs and tables which give an overview of the analysis done and variables that were considered in this study.

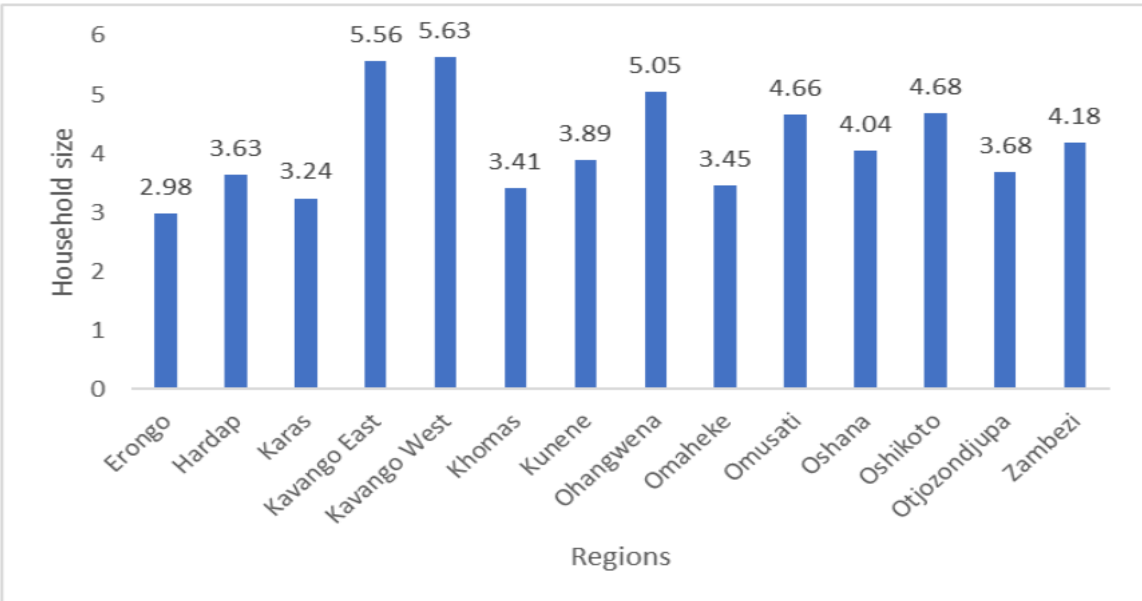


FIGURE 4.1: Average household size by region

Figure 4.1 shows the average household size by region. It can be observed that Kavango West had the highest average household size and Erongo had the lowest average household size.

TABLE 4.1: Average household size by rural-urban classification

Rural-urban classification	Mean	N	Std. deviation
Rural	4.43	5530	3.045
Semi-urban	3.97	285	2.862
urban	3.73	4275	2.901

From Table 4.1, it can be noted that most households that were considered were from rural (5523 households) while few households were from semi-urban (285 households), and urban had the second highest number of households (4275 households). The highest average household size was 4.43 from rural and the lowest average household size was 3.37 from urban.

TABLE 4.2: Average deprivation scores by region using entropy and equal weighting methods

Region	Average deprivation scores using entropy weights	Average deprivation scores using equal weights
Erongo	0.245210164	0.370169085
Hardap	0.330619369	0.452762929
Karas	0.287241796	0.408666273
Kavango East	0.460944693	0.409245902
Kavango West	0.574556305	0.452464808
Khomas	0.282635634	0.383507589
Kunene	0.466720128	0.449220286
Ohangwena	0.532677967	0.507936522
Omaheke	0.457868176	0.491222832
Omusati	0.53887755	0.530510032
Oshana	0.38329368	0.430719735
Oshikoto	0.506247905	0.495239972
Otjozondjupa	0.350513691	0.425925933
Zambezi	0.475348183	0.401060085
Average deprivation score for all the regions	0.415397044	0.4434809

We observed that the region with the lowest average deprivation scores when using both the entropy and equal weighting methods was the Erongo with an average deprivation score of 0.245210164 and 0.37076908, respectively (Table 4.2). On the other hand, Kavango West had the highest average deprivation score of 0.5745563054 when using entropy weighting while Omusati had the highest average deprivation score of 0.53051003 when using equal weighting. This means that Erongo region is the least deprived region when using both the entropy and equal weighting. However, Kavango West is the most deprived when using entropy and Omusati became the most deprived when using equal weighting method. Using the entropy method, we can note that the deprivation is consistent with the average household size as in Figure 4.1 for Erongo and Kavango West but not for all regions.

TABLE 4.3: Average deprivation scores by rural-urban classification using entropy and equal weighting methods

Rural-urban classification	Average deprivation scores using entropy weighting	Average deprivation scores using equal weighting
Rural	0.526309	0.493761
Semi-urban	0.437882	0.416959
urban	0.270426	0.380208
Average deprivation score for all rural-urban classification	0.415397	0.443481

The rural category had the highest average deprivation score of 0.526309 for entropy and 0.493753 for equal weighting, while urban category had the lowest average deprivation score of 0.270426 for entropy and 0.380094 for equal weighting (Table 4.3). This means that on average the poorest households were from rural areas and the less poor were from urban areas.

From Figure 4.2, it can be noted that there exists a difference in the deprivation scores when using the entropy and equal weighting. In particular, the entropy method resulted in low deprivation scores compared to the equal weighting method.

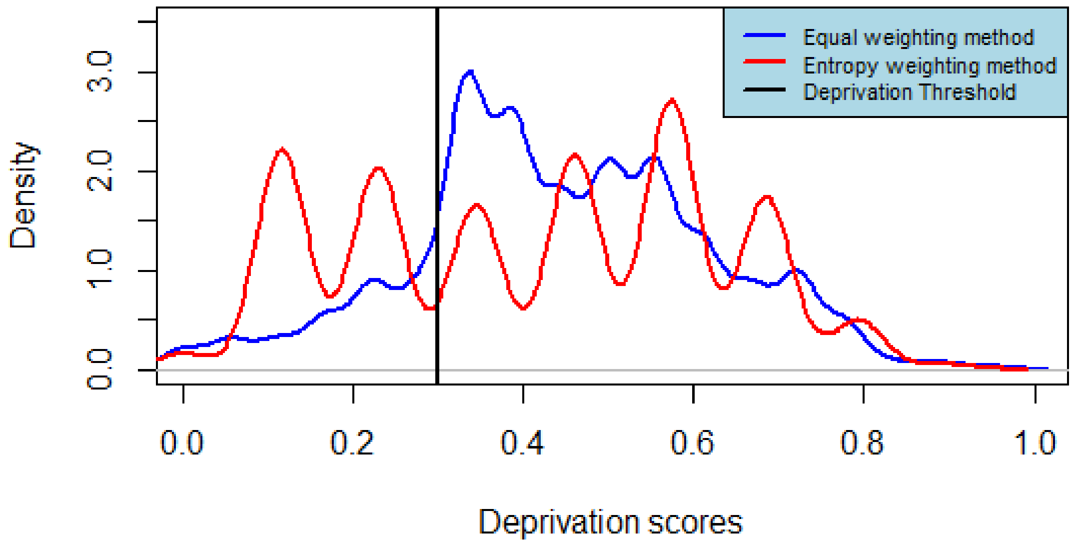


FIGURE 4.2: Density plot for the deprivation scores using entropy and equal weighting with all the nine indicators

When a deprivation threshold of 0.33 (the households with deprivation scores below 0.33 are considered non-deprived while the households with deprivation scores above 0.33 are considered deprived) is considered, we can say that there is a higher proportion of deprived households when we use the equal weighting method as opposed to a low proportion of deprived households when we use the entropy weighting method. This difference may be as a result of entropy being objective and equal weighting being subjective. This shows that there is no consistency between the two weighting methods.

From Figure 4.3, it can be observed that, when using the equal weighting method, the

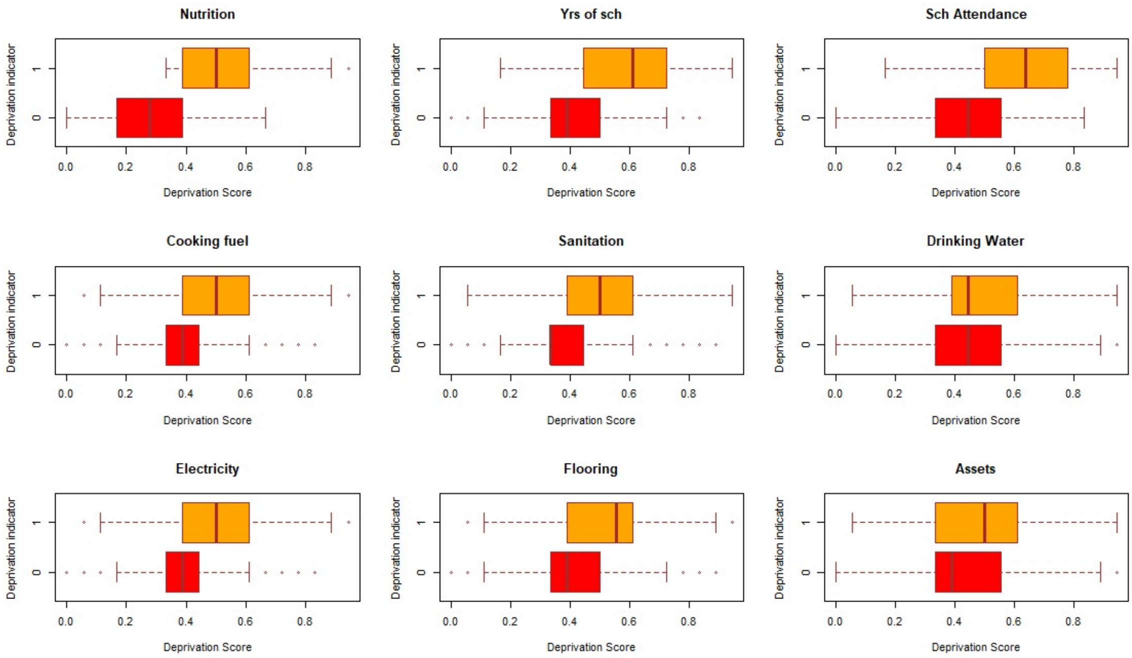


FIGURE 4.3: Boxplots of deprivation score (y_i) for all indicators using the equal weighting method

deprivation scores for the deprived households (orange) on average is higher than that of the non-deprived households (red) for all indicators including the nutrition indicator. This change in the deprivation score of the nutrition indicator which is different from the result in Figure 4.4 may be due to the fact that the nutrition indicator was given a bigger weight when using equal weighting, which was not the case when using the entropy weighting method. The median deprivation score for the deprived households is higher than that of the non-deprived households for all indicators (Figure 4.3). This is not the case for the nutrition indicator when using the entropy method (Figure 4.4).

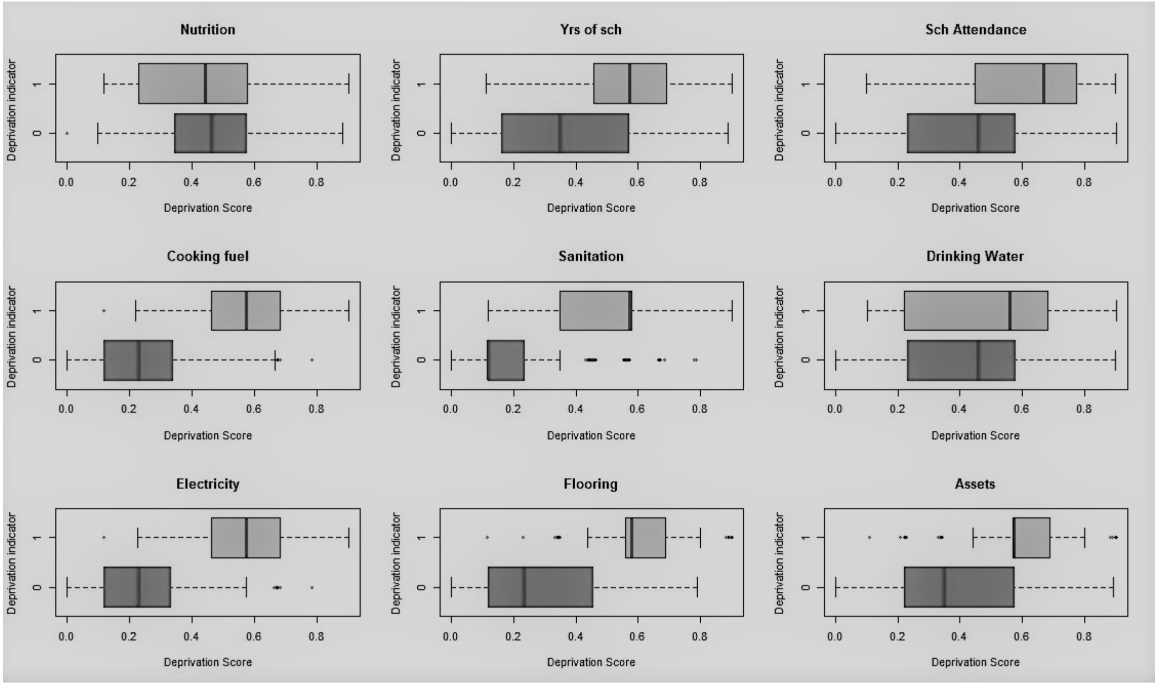


FIGURE 4.4: Boxplots of deprivation score (y_i) for all indicators using the entropy weighting method

From Figure 4.4, it can be noticed that the median deprivation score for the deprived households (grey) is lower than the median deprivation score for the non-deprived households (black) in the nutrition indicator when using the entropy method. On the other hand, when using the same weighting method for all the other indicators, it can be observed that the opposite is true, that is, the median deprivation scores for the deprived households (grey) is higher than the deprivation scores for the non-deprived households (black). We expect similar patterns of distribution for both entropy and equal weighting. However, this is not the case for the nutrition indicator in Figure 4.4 where it can be seen that the non-deprived households had a higher median deprivation score than the deprived households. This raised some concerns about the nutrition indicator which needed to be investigated in order to get an insight on its effect. To gain more insights on the nutrition indicator, it was removed from the model and we recomputed the deprivation scores for both entropy and equal weighting. A comparison of the results was done when the nutrition indicator was included to see it's effect on the deprivation scores for both weighting methods, and the result is seen in Figure 4.5.

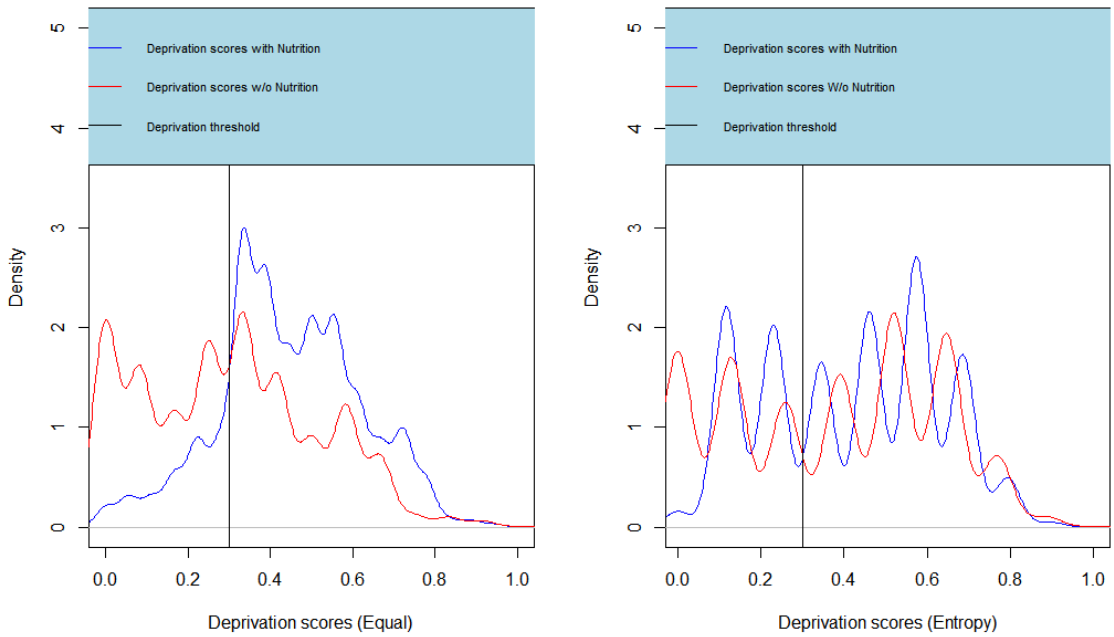


FIGURE 4.5: Density plots for the deprivation scores of entropy and equal weighting with and without the nutrition indicator

Figure 4.5 (left panel), shows that there is a rapid shift away from the deprivation threshold when the nutrition indicator was removed. In particular, the graph moves up and down when we are using the entropy. Moreover, the number of non-deprived individuals increases when the nutrition indicator is not considered compared to when the nutrition indicator is included considering the deprivation threshold. More specifically, it can be noticed that the nutrition indicator had an impact on the deprivation scores when we used the equal weighting method. Also, Figure 4.5 shows how equal weighting is indicator sensitive as it changes with the number of indicators.

In contrast, Figure 4.6 indicates that the entropy method is not indicator sensitive as it is not affected much by the change in the number of indicators. There is an insignificant shift in the distribution of the deprivation scores computed without the nutrition indicator using entropy method. This shows that this method is more stable and consistent than the equal weighting. Equal weighting shifts depending on the number of indicators. Equal weighting is dependent on the number of indicators and hence as the number of indicators changes it also changes rapidly. Also, when data is noisy or there is insufficient data then one might have a problem in the computation of deprivation scores using equal weighting. It is evident that equal weighting can only work well if all the indicators are well measured and are available.

Based on the internationally agreed upon threshold (0.33) for equal weighting proposed by Alkire and Santos (2011), the entropy method yields a proportion of deprived households which is less than the one obtained using the equal weighting method. One would expect to get a more or less similar picture when using entropy to the one when using the equal weighting method. In order for us to get a true picture on the proportion of the deprived households depicted in entropy, we adjust the threshold by setting the proportion of the deprived households in the entropy method to be equal to the proportion of the deprived households for the equal weighting method and deduce the corresponding threshold, which was found to be 0.1172192. This means that a threshold 0.33 under equal weighting is equivalent to a threshold of 0.1172 under entropy weighting.

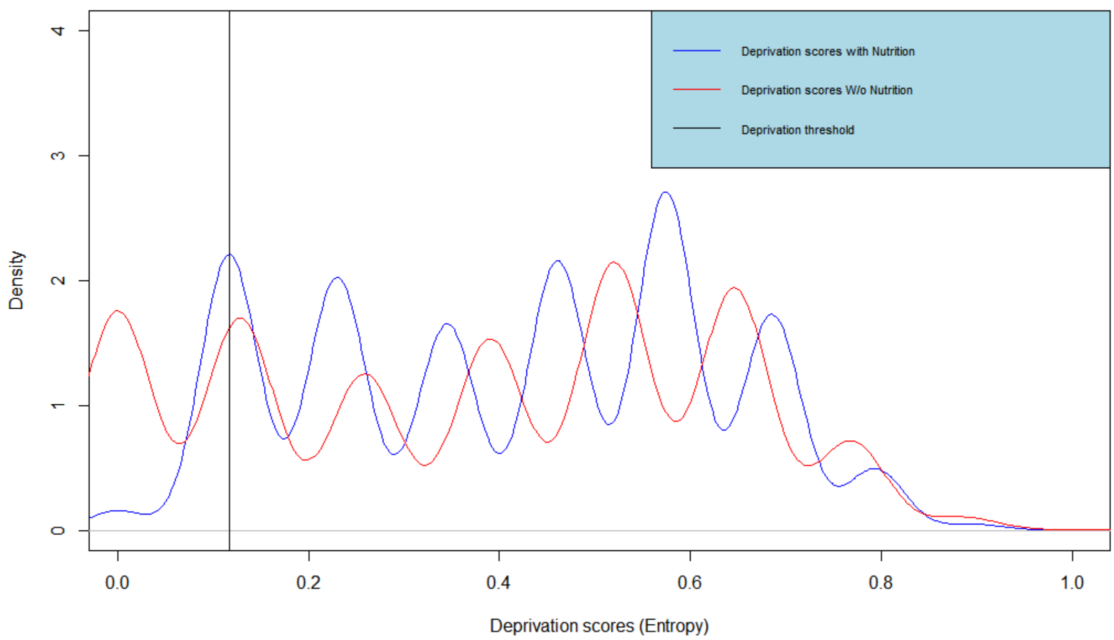


FIGURE 4.6: Density plots for the deprivation scores of entropy weighting with and without the nutrition indicator for the new threshold

Figure 4.7 shows the deprivation scores across all regions in Namibia using entropy and equal weighting methods. The red map shows the deprivation scores across all regions when using the entropy weighting method while the blue one shows the deprivation scores under the equal weighting method. These maps illustrate the variation in deprivation scores across regions depending on the type of weighting method used. It is also important to highlight that when considering a deprivation threshold of 0.33, on average most households have higher deprivation scores when using equal weighting. These different deprivation scores and weighting methods can have an effect on the MPI as they can vary depending on the type of method used.

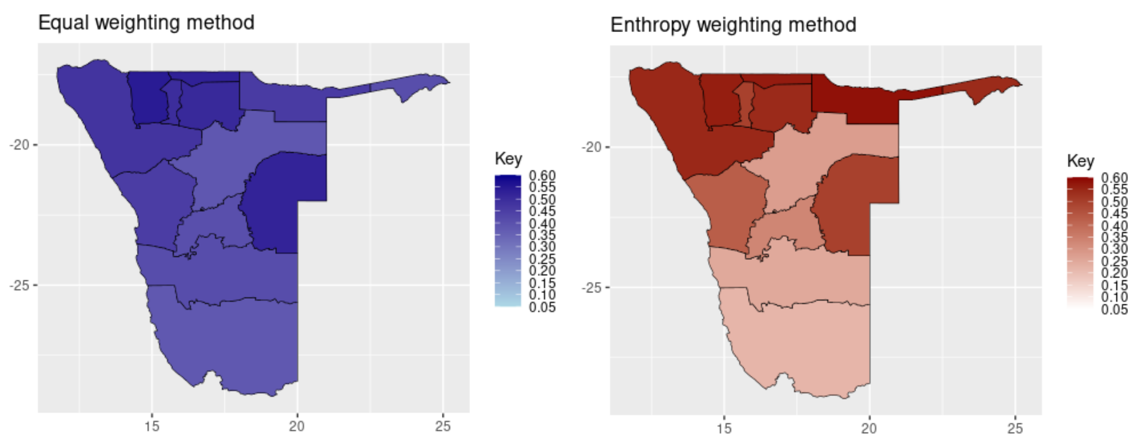


FIGURE 4.7: Namibian map of the average deprivation scores for all the regions using equal and entropy weighting methods

4.3 MPI computation for Namibia using equal- and entropy-weighting methods

The deprivation scores obtained using the two weighting methods (equal weighting and entropy weighting) and two thresholds (0.33 and 0.1172) were used to compute MPI as follows. Firstly, we computed the multidimensional headcount ratio (H) by dividing the total number of people who are multidimensionally poor with the total population. Secondly, the Intensity of Poverty (A) was computed by dividing the sum of the censored deprivation scores with the total number of individuals who are multidimensionally poor. Lastly, the MPI for Namibia was computed by applying Equation 3.9.

Note that, q is the number of individuals who are multidimensionally poor, n is the total population, and $\sum_{i=1}^{n} y_i(p)$ is the censored deprivation for all individuals from the data set in order for us to compute MPI, which is computed as the product of the multidimensional headcount ratio and the intensity of poverty. The results of these computations are shown in Table 4.4.

TABLE 4.4: MPI computation for Namibia using equal and entropy weighting methods

	Equal weighting	**Entropy weighting at threshold of 0.3333**	**Entropy weighting at threshold of 0.1172**
Multidimensional Headcount ratio (H)	0.8265	0.7093	0.9693
Intensity of poverty (A)	0.5092	0.5482	0.4483
MPI	0.4209	0.3888	0.4345

Table 4.4 shows that the proportion of individuals who are multidimensionally poor in Namibia is 83 percent for equal weighting, 70.9% for entropy weighting (with a threshold of 0.33), and 97% for entropy weighting (with a threshold of 0.1172). In other words, more than 70% of Namibians live in multidimensional poverty. In particular, they are deprived at least either in all indicators of a single dimension or a combination across dimensions. In addition, on average the multidimensionally poor individuals are deprived in 51% of the weighted indicators for equal weighting, 55% of the weighted indicators for entropy (with a threshold of 0.33), and 45% of the weighted indicators for entropy (with a threshold of 0.1172). In particular, they are deprived in more than 50% of the weighted indicator when using equal and entropy weighting (with a threshold of 0.33) methods. When entropy is used (with a threshold of 0.1172), they are deprived in less than 50% of the weighted indicators. Furthermore, in the total potential deprivations it could experience overall, Namibia is deprived in 42% for equal weighting, 39% for entropy (with a threshold of 0.33), and 44% for entropy (with a threshold of 0.1172) of those deprivations.

From Table 4.4, it can be observed that the MPI values under equal weighting, entropy (with a threshold of 0.33), and entropy (with a threshold of 0.1172) are close to each other. Precisely, it can be noted that the difference in the MPI values when we adjust the threshold for entropy weighting (from 0.33 to 0.1172) is minimal. Also, we would expect a big change in the MPI value when we adjust the threshold from 0.33 to 0.1172, but from Table 4.4 a minimal change in MPI is observed, meaning that the effect of the threshold adjustment is minimal. This shows that entropy is a robust method because it is not affected much by indicators and change in threshold values.

Considering a threshold of 0.33 (internationally agreed upon threshold) as the benchmark, we calibrated the threshold such that, different thresholds can be used whenever one wants to compute MPI for a particular reason. MPI has a robust functional form which permits comparisons across different environments/situations and having a fixed threshold value would inhibit this type of comparisons and applications across different countries/environments (Alkire & Santos, 2011). In fact, having a fixed value threshold is not recommended when one wants to apply MPI to different environments/situations as different environments had dissimilar realities, needs and availability of data. But, it is necessary to develop a robust model that is suitable for different situations, and one of the ways of coming up with such a model is to have a threshold that is not fixed on particular definitions or theories.

4.4 Namibia regional MPI using equal- and entropy-weighting methods

MPI can also be decomposed by regions to understand how much each region contributes to the national MPI and to understand the regional spatial variation in multidimensional poverty across Namibia using the equal weighting method (at threshold of 0.33) and the entropy weighting method at threshold of 0.33 and threshold of 0.1172, we obtained the incidence (H) of poverty, intensity of poverty (A) and the MPI for each region in Namibia as shown in Tables 4.5 and 4.6.

TABLE 4.5: Namibia regional poverty incidence (H) and intensity (A) using the equal weighting method at threshold of 0.33 (method 1) and the entropy weighting method at thresholds of 0.33 (method 2a) and 0.1172 (method 2b)

	Method 1		Method 2a		Method 2b	
Region	Multidimensional headcount ratio (H)	Intensity of poverty (A)	Multidimensional headcount ratio (H)	Intensity of poverty (A)	Multidimensional headcount ratio (H)	Intensity of poverty (A)
Zambezi	0.6488	0.4884	0.83771	0.5248	0.9945	0.4727
Karas	0.8814	0.4433	0.3091	0.4783	0.9481	0.2763
Erongo	0.8562	0.4158	0.2634	0.4352	0.9109	0.2506
Hardap	0.9086	0.4762	0.4089	0.5212	0.9646	0.3198
Kavango East	0.7077	0.5079	0.8011	0.5642	0.9650	0.4977
Kavango West	0.7666	0.5248	0.9418	0.6010	0.9966	0.5789
Khomas	0.8173	0.4351	0.3828	0.4315	0.9277	0.2860
Kunene	0.7315	0.5389	0.7820	0.5556	0.9662	0.4855
Ohangwena	0.8699	0.5717	0.9306	0.5809	0.9949	0.5565
Omaheke	0.8887	0.5232	0.7575	0.5305	0.9880	0.4487
Omusati	0.9188	0.5755	0.9437	0.5801	0.9980	0.5580
Oshana	0.8353	0.5046	0.7004	0.5074	0.9649	0.4133
Oshikoto	0.8619	0.5455	0.8619	0.5716	0.9877	0.5215
Otjozondjupa	0.8391	0.4798	0.5105	0.5291	0.9368	0.3704

From Table 4.5 and 4.6, it can be noted that the poverty incidence (H), intensity of poverty (A), and the MPI values vary from region to region and also these values depend on the weighting method and threshold value used.

TABLE 4.6: Namibia regional MPI using the equal weighting with a threshold of 0.33 (method 1) and the entropy weighting method with a threshold of 0.33 (method 2a) and 0.1172 (method 2b) respectively

Region	Method 1	Method 2a	Method 2b
Zambezi	0.316849825	0.439596894	0.470064988
Karas	0.390667159	0.147809845	0.261985915
Erongo	0.355956241	0.114606074	0.228269591
Hardap	0.432705437	0.213084536	0.308527336
Kavango East	0.359416695	0.451951771	0.480208139
Kavango West	0.40227369	0.566029252	0.576945548
Khomas	0.355591653	0.16517224	0.265301613
Kunene	0.394178709	0.434535318	0.469043489
Ohangwena	0.497280871	0.540590809	0.5536292
Omaheke	0.464967057	0.401837132	0.443363006
Omusati	0.528810734	0.547416519	0.556909989
Oshana	0.421477804	0.355412552	0.39877871
Oshikoto	0.470154046	0.492688257	0.515082038
Otjozondjupa	0.402602089	0.270145764	0.347012488

Equal-weighting:

Using the equal weighting method with a threshold of 0.33 (Table 4.5), we can observe that the highest incidences of poverty are found in regions such as Omusati (91.9%), Hardap (90.9%), Omaheke (88.%), and !Karas (88.1%). Zambezi region has the lowest poverty incidence (64.9%) followed by Kavango East (70.8%), Kavango West (76.7%), and Kunene (73.2%) just to name a few. Meaning that, the poverty incidence is lower in the two Kavango regions and the Zambezi region and the other northern regions. All this shows that the given headcount ratio (H) percentage of people who are multidimensionally poor in that region. They are deprived in either all the indicators of one dimension or a combination across dimensions.

Table 4.5 reveals that when using equal weighting with a threshold of 0.33, poverty is more intense in northern parts of Namibia such as Omusati (57.6%), Ohangwena (57.2%), Oshikoto (54.6%), and Kunene (53.9%). While the central and southern parts of Namibia such as Erongo (41.6%), Khomas (43.5%), Karas (44.3%) and Hardap (47.6%), and Otjozondjupa (48%) experience a lower intensity of poverty. That is, on average the regions are deprived in the percentage given (Intensity of poverty) of the weighted indicators.

Using the equal weighting method, the highest multidimensionally poor region is Omusati (52.9%), followed by Ohangwena (49.7%), Oshikoto (47%) and Omaheke (46.5%) (Table 4.6). The least multidimensionally poor region is Zambezi (31.7%), followed by Khomas (35.6%), Erongo (35.6%), and Kavango East (35.9%). This result is different from the findings of English (2016), where Zambezi and Kavango East were ranked as some of the poorest regions in Namibia. This difference may be attributed to the instability of the equal weighting method. The deprivation in MPI that each region experiences is given as the percentage of the total potential deprivation it could experience.
Using equal weighting from Table 4.6, in the northern regions Zambezi (31.7%) is the least multidimensionally poor region and Omusati (52.9%) is the worst multidimensionally poor region. Omaheke (46.5%) is the worst multidimensionally poor region and Erongo (35.6%) is the least multidimensionally poor region in the central parts of Namibia. On the other hand, Hardap (43.3%) is the worst multidimensionally poor region and Karas (39.1%) is the least multidimensionally poor region in the southern parts of Namibia.

Entropy-weighting method with a threshold of 0.33:
The findings in Table 4.5 shows that when using the entropy weighting method with a threshold of 0.33, Omusati region has the highest poverty incidence (94.4%), followed by Kavango West (94.2%), Ohangwena (93.1%), and Oshikoto (86.2%). While Erongo has the lowest incidence of poverty (26.3%), followed by Karas (30.9%), Khomas (38.3%), and Hardap (40.9%).

Poverty is more intense in Kavango West (60.1%) and less intense in Khomas (43.2%). This means that, on average people in Kavango west and Khomas are deprived by 60.1% and 43.2% of the weighted poverty indicators respectively (Table 4.5).

From Table 4.6, it can be observed that Kavango West (56.6%) is the worst multidimensionally poor region, and Erongo (11.5%) is the least multidimensionally poor region.

Table 4.6 shows that, Oshana (35.5%) is the least multidimensionally poor region and Kavango west (56.6%t) is the worst multidimensionally poor region in the northern parts of Namibia when using the entropy weighting method at a threshold of 0.33. Omaheke (40.2%) is the worst multidimensionally poor region and Erongo (11.5%) is the least multidimensional poor region in the central parts of Namibia. Whereas, Hardap (21.3%) is the worst multidimensionally poor region and Karas (14.8%) is the least multidimensionally poor region in the southern parts of Namibia when using the entropy weighting method at a threshold of 0.33.

Entropy-weighting method with a threshold of 0.1172:
Table 4.5 portrays the highest incidence level of poverty to be in Omusati (99.8%) and the lowest being in Erongo (91.1%). In Omusati region, 99.8% of people experience multiple deprivation, while 91.1% of people in Erongo region experience multiple deprivations, when using the entropy weighting method with a threshold of 0.1172. Meaning that the people are deprived in either all the indicators of a single dimension or a combination across dimensions.

It can also be observed from Table 4.5 that poverty is more intense in Kavango West (57.9%) and less intense in Erongo (25.1%). On average, people in Kavango West and Erongo are deprived in 57.9% and 25.1% of the weighted poverty indicators, respectively.

MPI is the highest in Kavango West (57.7%) and the lowest in Erongo (22.8%). This shows that Kavango West is the multidimensionally poorest region in Namibia and Erongo is the least multidimensionally poor region (Table 4.6).

When using the entropy weighting method at a threshold of 0.1172, Oshana (39.9%) is the least multidimensionally poor region and Kavango west (57.7%) is the worst multidimensionally poor region in the northern parts of Namibia (Table 4.6). Omaheke (44.3%) is the worst multidimensionally poor region and Erongo (22.8%) is the least multidimensionally poor region in the central parts of Namibia. Whereas, Hardap (30.9%) is the worst multidimensionally poor region and Karas (26.2%) is the least multidimensionally poor region in the southern parts of Namibia when using the entropy weighting method at a threshold of 0.1172.

The entropy weighting method with a threshold of 0.1172 yielded findings that are close to those of English (2016), which point that poverty in Namibia is more intense in the northern areas that include Kavango, Zambezi, Oshikoto, Ohangwena and Kunene. This closeness may be a result of entropy being more stable than equal weighting. From the results in Table 4.6, it can be observed that the regions (Erongo, Khomas, Karas and Hardap) that are less poor are the ones where most economic activities take place, such as mining, commercial agriculture, tourism, and export commodities. The ones with low economic activities like small livestock farming and subsistence farming are the ones that are poorer (Kavango West, Omusati, Ohangwena and Oshikoto). This shows that economic activity may have an impact on MPI.

The entropy weighting method (0.1172) gave us results that are different from the results obtained using the equal weighting method at a threshold of 0.33. The entropy weighting method is more stable and consistent as it is not affected much by the change in thresholds, whereas the equal weighting method is effortlessly altered by adjusting the threshold value.

Proposition: *Let Y be a random variable taking values $y_i \in (0,1)$; $i = 1, 2, ..., n$; denoting deprivation scores for individual i. Also let $p \in (0,1)$ be the deprivation threshold and $Y \sim Beta(\alpha, \beta, p, 1)$; individual i is deprived if $p \leq y_i < 1$ and is given $y_i(p)$, otherwise an individual i is non-deprived if $0 < y_i < p$ and hence equate his/her deprivation score to 0. Then it follows that MPI is the expected value of Y.*

Proof:

From Equations 3.8 it follows that

$$Aq = \sum_{i=1}^{n} y_i(p) \tag{4.1}$$

On the other hand, we know that mean of the deprivation scores $[E(Y)]$ is computed as $\frac{\sum_{i=1}^{n} y_i(p)}{n}$. Thus,

$$\frac{\sum_{i=1}^{n} y_i(p)}{n} = \frac{Aq}{n}, \tag{4.2}$$

Recall that Y is a left truncated beta distributed random variable and hence $E(Y) = \frac{\alpha+p\beta}{\alpha+\beta}$. Therefore,

$$\begin{aligned} E(Y) &= \frac{Aq}{n} \\ \frac{\alpha + p\beta}{\alpha + \beta} &= \frac{Aq}{n} \\ \frac{\alpha + p\beta}{\alpha + \beta} &= A \times \frac{q}{n} \\ A &= \frac{n}{q} \times \frac{\alpha + p\beta}{\alpha + \beta} \end{aligned} \tag{4.3}$$

Substituting Equation (3.7) and Equation (4.3) into Equation (3.9) as follows;

$$\begin{aligned} MPI &= H \times A \\ &= \frac{q}{n} \times \frac{n(\alpha + p\beta)}{q(\alpha + \beta)} \\ &= \frac{q}{n} \times \frac{n}{q} \times \frac{\alpha + p\beta}{\alpha + \beta} \\ &= \frac{\alpha + p\beta}{\alpha + \beta} \end{aligned} \tag{4.4}$$

Which is the expected value of Y.

4.5 MPI as computed using the beta distribution model

The beta distribution model in Equation 4.3 was used to compute Namibia MPI using equal weighting at a threshold of 0.33 and the entropy weighting method at a threshold of 0.3333 and 0.1172. The results are shown in Table 4.7.

TABLE 4.7: Namibia MPI using beta distribution for equal weighting with a threshold of 0.33 (method 1) and entropy weighting method with a threshold of 0.33 (method 2a) and 0.1172 (method 2b) respectively

	Alpha	Beta	MPI
Method 1	3.9731	4.6959	0.6102
Method 2a	1.8721	2.6346	0.6389
Method 2b	1.8721	2.6346	0.4839

Table 4.7 shows that Namibia has higher multidimensional poverty (61%) when we use the equal weighting method at a threshold of 0.3333 and the entropy weighting at a threshold of 0.3333 (63.9%) compared to when we used the entropy weighting at a threshold of 0.1172 (48.4%). The MPI result for Namibia using the beta distribution model for the entropy weighting at a threshold of 0.1172 is more close to the results obtained in Table 4.4 when using the method proposed by Alkire and Santos (2011) and using the same weighting method and threshold. Also, we noted that MPI increases with the increase in the threshold and vice versa as shown by Equation 3.46

TABLE 4.8: 95% confidence interval for Namibia MPI using beta distribution for equal weighting with a threshold of 0.33 (method 1) and entropy weighting method with a threshold of 0.33 (method 2a) and 0.1172 (method 2b) respectively

	95%Confidence Interval	
	Lower bound	Upper bound
Method 1	0.6100	0.6104
Method 2a	0.6387	0.6390
Method 2b	0.4837	0.4841

When using the equal weighting method at 0.33 threshold, the entropy weighting method at 0.33 threshold and the entropy weighting method at 0.1172 threshold, we are 95% confident that the true MPI values in Table 4.7 fall inside the bounds given in Table 4.8.

The beta distribution model was again used to compute the MPI for each region in Namibia and the results are shown in Table 4.9.

TABLE 4.9: Namibia regional MPI using beta distribution for equal weighting with a threshold of 0.33 (method 1) and entropy weighting method with a threshold of 0.33 (method 2a) and 0.1172 (method 2b) respectively

	Method 1			Method 2a			Method 2b		
Region	Alpha	Beta	MPI	Alpha	Beta	MPI	Alpha	Beta	MPI
Zambezi	3.4187	5.0286	0.6012	3.6718	4.1331	0.6452	3.6718	4.1331	0.5325
Karas	5.1506	6.7999	0.6188	1.531	3.7991	0.5225	1.531	3.7991	0.3708
Erongo	6.4608	9.5471	0.6004	1.6753	5.1567	0.4943	1.6753	5.1567	0.3337
Hardap	4.902	5.7119	0.6394	1.4623	2.9606	0.5515	1.4623	2.9606	0.4091
Kavango East	3.2215	4.2401	0.6193	1.9856	2.3221	0.6388	1.9856	2.3221	0.5241
Kavango West	3.0879	3.6215	0.6384	4.3830	3.2455	0.715	4.383	3.2455	0.6244
Khomas	5.1246	7.4814	0.6024	2.1683	5.5035	0.5194	2.1683	5.5035	0.3667
Kunene	3.1647	3.6631	0.6405	2.3927	2.7339	0.6427	2.3927	2.7339	0.5292
Ohangwena	4.0690	3.8678	0.6735	4.1441	3.6357	0.6869	4.1441	3.6357	0.5874
Omaheke	4.8093	4.8793	0.6626	2.6872	3.2601	0.6327	2.6872	3.2601	0.5161
Omusati	5.0800	4.4680	0.6865	4.5555	3.8982	0.691	4.5555	3.8982	0.5929
Oshana	4.0357	4.9445	0.6311	2.0191	3.2487	0.5868	2.0191	3.2487	0.4556
Oshikoto	4.0722	4.0646	0.6653	3.0204	2.9459	0.6692	3.0204	2.9459	0.5641
Otjozondjupa	4.2192	5.0598	0.6346	1.4645	2.7137	0.5648	1.4645	2.7137	0.4266

Northern regions of Namibia

Table 4.9 illustrates that Omusati (68.7%) is the worst multidimensionally poor region and Zambezi (60.1%) is the least multidimensionally poor region in the northern parts of Namibia when using the equal weighting method at a threshold of 0.33. On the other hand, using the entropy weighting method at a threshold of 0.33 and 0.1172, Kavango West is the worst multidimensionally poor region at both thresholds by 71.5% and 62.4%, respectively. Oshana region is the least multidimensionally poor region when using the entropy weighting method at thresholds of 0.33 and 0.1172 with 58.7% and 45.6%, respectively.

Central regions of Namibia

In the central regions of Namibia, Omaheke (66.3%) is the worst multidimensionally poor region and Erongo (60%) is the least multidimensionally poor region when using the equal weighting with a threshold of 0.33. Similarly, when using entropy weighting method with thresholds of 0.33 and 0.1172, Omaheke is the worst multidimensionally poor region by 63.3% (at threshold of 0.33) and 51.6% (at threshold of 0.1172), and Erongo is the least multidimensionally poor region by 49.4% (at threshold of 0.33) and 33.4% (at threshold of 0.1172).

Southern regions of Namibia

Hardap (63.9%) is the worst multidimensionally poor region and Karas (61.9%) is the least multidimensionally poor region when using equal weighting with a threshold of 0.33 (Table 4.9). Similarly from Table 4.9, when using the entropy weighting method with thresholds of 0.33 and 0.1172, Hardap is the worst multidimensionally poor region with 55.2% (at threshold of 0.33) and 40.9% (at threshold of 0.1172), and Karas is the least multidimensionally poor region with 52.3% (at threshold of 0.33) and 37.1% (at threshold of 0.1172).

Overall, it can be noted from Table 4.9 that Omusati (68.7%) is the worst multidimensionally poor region and Zambezi (60.1%) is the least multidimensionally poor region when using equal weighting at a threshold of 0.33. On the other hand, Kavango West (71.5%) is the worst multidimensionally poor region and Erongo (49.4%) is the least multidimensionally poor region when using the entropy weighting method at a threshold of 0.33. Kavango West (62.4%) as the worst multidimensionally poor region and Erongo (33.4%) is the least multidimensionally poor region in Namibia with the entropy weighting method at a threshold of 0.1172. This also shows that the entropy weighting method is more stable as it is not greatly affected by change in threshold. From Table 4.9, it can also be observed that, the MPI values are slightly bigger than the MPI values in Table 4.6. This difference in the MPI values may be due to changes in thresholds and the estimated values of alpha and beta as they are estimates and not the exact values.Table 4.9 shows that the northern regions are the worst multidimensionally poor regions in Namibia and the central regions are the least multidimensionally poor regions.

TABLE 4.10: 95% confidence interval for Namibia regional MPI using beta distribution for equal weighting with a threshold of 0.33 (method 1) and entropy weighting method with a threshold of 0.33 (method 2a) and 0.1172 (method 2b) respectively

	Method 1		Method 2a		Method 2b	
Region	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Zambezi	0.6007	0.6016	0.6447	0.6457	0.5316	0.5334
Karas	0.6184	0.6191	0.5219	0.523	0.3698	0.3718
Erongo	0.6002	0.6006	0.4939	0.4946	0.3331	0.3343
Hardap	0.6391	0.6398	0.5508	0.5522	0.4078	0.4103
Kavango East	0.6187	0.6198	0.6380	0.6397	0.5226	0.5256
Kavango West	0.6378	0.6389	0.7144	0.7155	0.6235	0.6253
Khomas	0.6021	0.6026	0.5191	0.5197	0.3662	0.3672
Kunene	0.6400	0.6411	0.6420	0.6434	0.5280	0.5305
Ohangwena	0.6731	0.6739	0.6865	0.6873	0.5867	0.5882
Omaheke	0.6622	0.6630	0.6321	0.6334	0.5150	0.5172
Omusati	0.6861	0.6868	0.6907	0.6914	0.5923	0.5936
Oshana	0.6307	0.6315	0.5863	0.5874	0.4546	0.4565
Oshikoto	0.6649	0.6657	0.6687	0.6697	0.5632	0.5650
Otjozondjupa	0.6343	0.6350	0.5642	0.5655	0.4255	0.4278

Table 4.10 shows the 95% confidence interval of the regional MPI values when using both the entropy weighting and equal weighting. About 95% is certain that the true MPI values in Table 4.9 are within the lower and upper bounds given in Table 4.10 for each region when we use the equal weighting method at 0.33 threshold, the entropy weighting method at 0.33 threshold and the entropy weighting method at 0.1172 threshold.

4.6 Sensitivity analysis of the parameters

A sensitivity analysis of the parameters α, β and p was performed to assess how changes in these parameters affect the MPI. This analysis was performed using the derivative-based method as follows.
First, we recall from Equation 4.4 that;

$$MPI = \frac{\alpha + p\beta}{\alpha + \beta}$$

Then, taking the partial derivatives of MPI with respect to α, β and p respectively, yields the following;

$$\begin{aligned} \frac{\partial MPI}{\partial \alpha} &= \frac{(\alpha+\beta)-(\alpha+p\beta)}{(\alpha+\beta)^2} \\ &= \frac{\alpha+\beta-\alpha-p\beta}{(\alpha+\beta)^2} \\ &= \frac{\beta-p\beta}{(\alpha+\beta)^2} \\ &= \frac{\beta(1-p)}{(\alpha+\beta)^2}, \end{aligned} \tag{4.5}$$

where $0 < p < 1$ and $\alpha, \beta > 0$.
Since $0 < p < 1$ and $\alpha, \beta > 0$ in Equation 4.5, MPI is an increasing function with respect to α, which means that MPI increases with an increase in the value of α and vice versa.

$$
\begin{aligned}
\frac{\partial MPI}{\partial \beta} &= \frac{(\alpha+\beta)p-(\alpha+p\beta)}{(\alpha+\beta)^2} \\
&= \frac{p\alpha+p\beta-\alpha-p\beta}{(\alpha+\beta)^2} \\
&= \frac{p\alpha-\alpha}{(\alpha+\beta)^2} \\
&= \frac{\alpha(p-1)}{(\alpha+\beta)^2} \\
&= -\Big[\frac{\alpha+p\alpha}{(\alpha+\beta)^2}\Big],
\end{aligned} \tag{4.6}
$$

where $0 < p < 1$ and $\alpha, \beta > 0$.

Equation 4.6 shows that MPI is a decreasing function with respect to β as $0 < p < 1$ and $\alpha, \beta > 0$, which implies that an increase in the value of β leads to a decrease in the MPI and oppositely, a decrease in the value of β increases MPI.

$$
\begin{aligned}
\frac{\partial MPI}{\partial p} &= \frac{1}{\alpha+\beta}\,(\alpha+p\beta) \\
&= \frac{1}{\alpha+\beta}\,\frac{\partial MPI}{\partial p}\,(\alpha+p\beta) \\
&= \frac{1}{\alpha+\beta}\,(\beta) \\
&= \frac{\beta}{\alpha+\beta},
\end{aligned} \tag{4.7}
$$

where $\alpha, \beta > 0$.

Equation 4.7 portrays MPI as an increasing function with respect to p as $\alpha, \beta > 0$. But this shift is dependent on the values of α and β. In this case, increasing the value of α while decreasing the value of β increases MPI and vice versa. This shows that the effect of p on MPI also depends on the values of α and β.

4.7 Numerical simulation of beta distribution

In the numerical simulation of the beta model, we used the entropy deprivation scores data set with a threshold value of 0.1172 to analyse the effect of α and β on the model. In this section we simulated the beta density function for different value combinations of α and β.

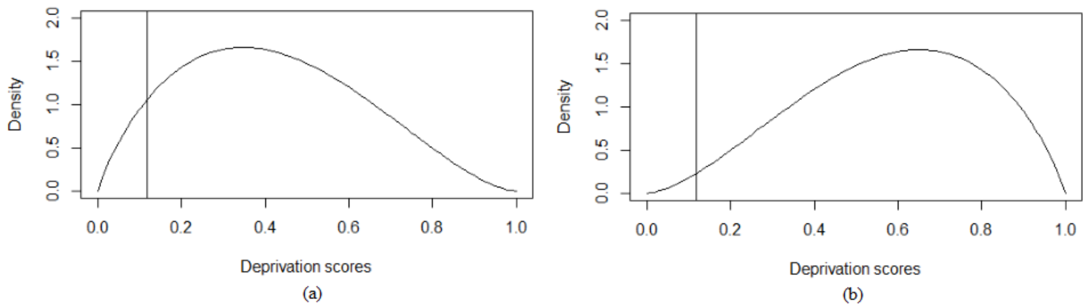


FIGURE 4.8: Beta distribution with a threshold of 0.1172

Figure 4.8 (a) shows the behaviour of beta distribution when the values of α and β are switched. In this case, $\alpha = 1.872071$ and $\beta = 2.634632$ for Figure 4.8 (a), and $\alpha = 2.634632$ and $\beta = 1.872071$ for Figure 4.8 (b)). Figure 4.8 (a) shows a slightly right-skewed distribution whereas Figure 4.8 (b) is slightly left-skewed.Since $\alpha, \beta > 1$, the mode of this distribution occurs at $y = (\alpha - 1)/(\alpha + \beta - 2)$. This means that the shape and skewness of the distribution depend on the values of α and β.

Keeping β constant while increasing α in Figure 4.9 (a) shows that the beta distribution becomes strongly left-skewed as α increases. Since $\alpha, \beta > 1$, the mode of these distribution occurs at $y = (\alpha - 1)/(\alpha + \beta - 2)$. If α increases and β remains constant, MPI will also increase. In contrast, a decrease in α for a fixed β would cause the beta distributions to become right-skewed and thus decreasing MPI. This kind of distribution illustrates how increasing α plays a role in flattening the deprivation in the extreme of the non-deprived. The more α increases, the more the density of the non-deprived (on the left) becomes asymptotic to the y-axis.

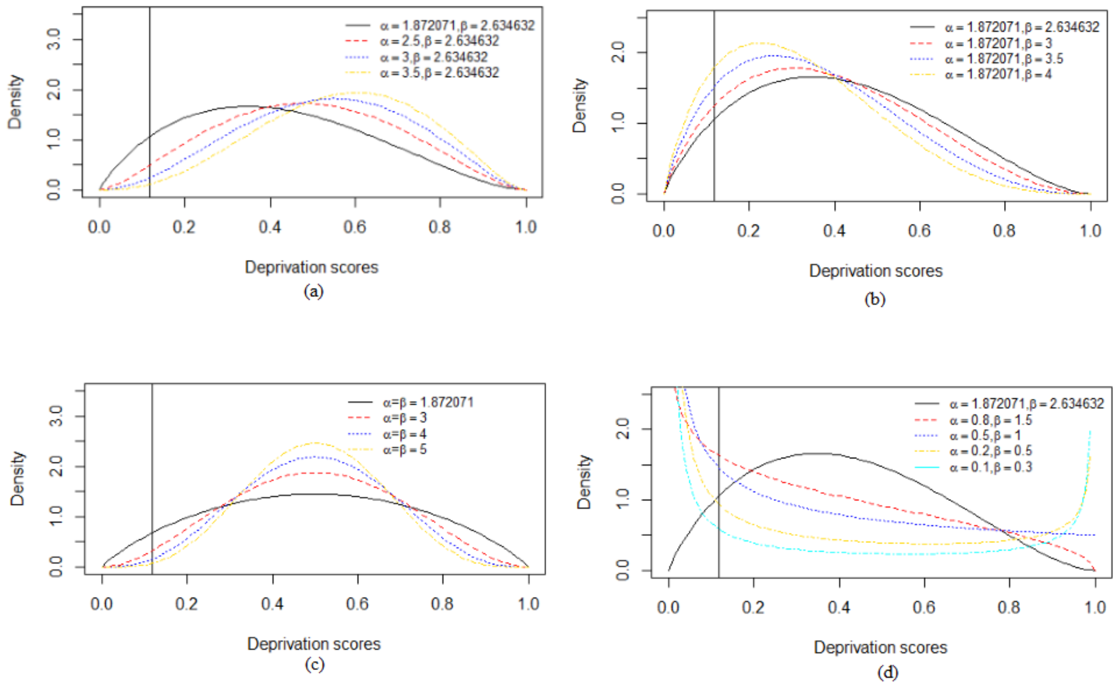


FIGURE 4.9: Beta distributions with different values of α and β

For a constant α and an increasing β, the distribution becomes strongly right-skewed as β increases (Figure 4.9 (b)). Since $\alpha, \beta > 1$, the mode of these distributions occurs at $y = (\alpha - 1)/(\alpha + \beta - 2)$. This shows how increasing β plays a role in flattening the deprivation in the extreme of the deprived.

The beta distributions in Figure 4.9 (c) are unimodal, symmetric (about 0.5) beta distributions, and since $\alpha = \beta$, the distribution is slightly right-skewed, but as α and β increase the distribution becomes peaked and less skewed as the mode also approaches 0.5. In terms of MPI, this type of distribution portrays the situation in most developed countries where the gap between the rich and poor is minimal (at the extreme ends), and there is less inequality between the rich and poor as the distribution is symmetric (about 0.5). Also, it is observed that as α and β decrease, the gap between the rich and poor is reduced as the extreme ends are being narrowed.

Decreasing both α and β in Figure 4.9 (d) yields skewed, U-shaped beta distributions with their antimode occurring at $y = (\alpha - 1)/(\alpha + \beta - 2)$ as α and β becomes <1. As α and β decrease more, the antimode approaches 0.5 and the distribution becomes less skewed. Switching α and β would change the direction of the skew and the shape of the distribution. These beta distributions portray an increase in the deprivation scores at the extreme ends, meaning that there exists a big gap between the rich and poor households.

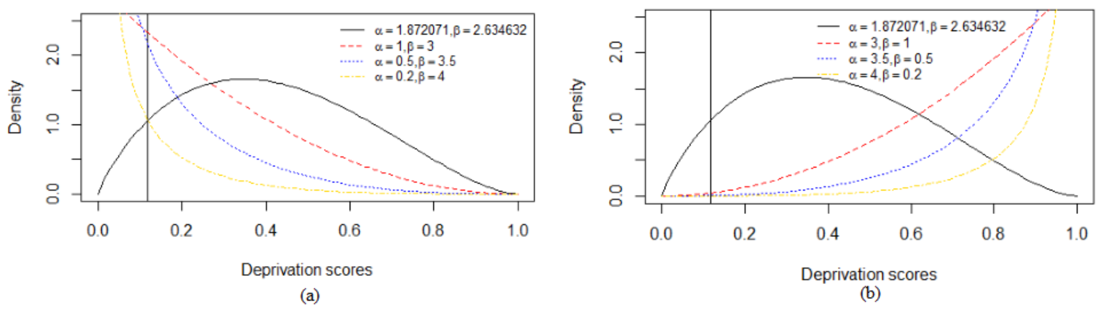


FIGURE 4.10: Beta distributions for decreasing and increasing values of α and β

The distributions in Figure 4.10 (a) become more curved as α decreases and β increases yielding reverse J-shaped beta distribution. Since $\alpha \leq 1$ and $\beta > 1$, the mode/antimode of these distributions occurs at $y = (\alpha - 1)/(\alpha + \beta - 2)$ which is zero (no mode or antimode). Increasing α and decreasing β values yields a J-shaped beta distribution (Figure 4.10 (b)). Increasing β while reducing α reduces the number of the deprived household, causing the graphs to be more asymptotic to the y-axis and increasing the number of the non-deprived households.

Figure 4.11 illustrates the effects that α and β have on MPI. It shows what is happening

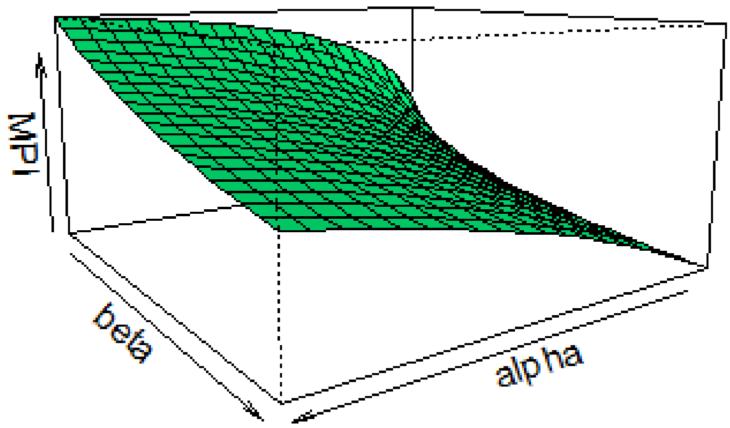


FIGURE 4.11: The effects of alpha and beta on MPI

in Equations 4.5, 4.6, and 4.7. From Figure 4.11 (three dimensional figure), it is confirmed that the MPI is indeed dependent on α and β as discussed above. This effect may be due to the fact that MPI is an increasing function with respect to α, and it is a decreasing function with respect to β which is shown in Equations 4.5, 4.6, and 4.7 respectively where sensitivity analysis was done.

4.8 The beta regression model for poverty measure

Apart from the initial indicators used in this study to compute MPI, we used beta regression to identify additional significant indicators of MPI for Namibia. The dependent variable used is the deprivation scores (y_i) computed using the entropy weighting method. The independent variables used in the beta regression are: housing, access to information or ICT, access to clinic or hospital, food security, and transportation assets. The full description of these independent variables is given in Table 4.11.

TABLE 4.11: MPI indicators used in beta regression

MPI Indicator	Deprivation
Housing	A household is deprived in the event that it has deficient lodging - the rooftop or divider are made of any of the accompanying: sticks and grass, wooden shafts, sticks, clay or cow-dung, thatch, mud, grass, other or none
Access to information or ICT	A household is deprived if it doesn't possess any of these resources: TV, radio, PC, cell phone, tablet or Ipad, or the household doesn't have internet access at home or somewhere else
Access to clinic/hospitals	A household is deprived if a medical clinic or hospital is in excess of 20 km from home or over 30 minutes one way from home
Food security	A household is deprived in the event that it didn't have sufficient food to eat in the 7 days before the survey
Transportation assets	A household is deprived on the off chance that it doesn't claim any of these resources: bakkie/4wheel drive, vehicle/station wagon, bus/minibus, horses/camel, motor cycle/scooter, or bike

In order to fit the beta regression, we transformed the dependent variable (deprivation scores y_i) by using the transformation $(y_i \cdot (n-1) + 0.5)/n$, where n is the sample size (Smithson and Verkuilen (2006)) because it assumed extreme values of zero and one. The results of the fitted beta regression model is given in Table 4.12.

Table 4.12 provides the parameter estimates of the beta regression model. At $\alpha = 0.05$, all the indicators are significantly associated to deprivation scores because their corresponding p-values are small (less than 0.05). However, the diagnostic plots (Figure 4.12) revealed that there were observations corresponding to large cook's distance (Figure 4.12 (b)) and large residuals Figure 4.12 (a), (c), (d)) and hence were most influential in the data. In addition, the Pseudo R-squared of 0.2264, which was very low, indicated that only 22.64% of the variability in the dependent variable is explained by the regression model. In order to improve

TABLE 4.12: Parameter estimates of the beta regression model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.3705	0.02273	-60.28	<2e-16 ***
Housing	0.90383	0.01734	52.12	<2e-16 ***
Access to information or ICT	0.58125	0.03515	16.54	<2e-16 ***
Access to clinic/hospital	0.35664	0.02014	17.7	<2e-16 ***
Food security	0.40466	0.01871	21.63	<2e-16 ***
Transportation assets	0.36898	0.01751	21.08	<2e-16 ***
(ϕ)	5.7615	0.07599	75.82	<2e-16 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Pseudo R-squared:	0.2264			

the model, we had to remove the most influential observations from the data and refit the model (Cribari-Neto & Zeileis, 2010).

TABLE 4.13: Results of the improved Beta regression model.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.73557	0.0161	-107.79	<2e-16 ***
Housing	1.21148	0.01149	105.4	<2e-16 ***
Access to information or ICT	0.3304	0.02934	11.26	9.71e-13 ***
Access to clinic/hospital	0.44691	0.01392	32.11	<2e-16 ***
Food security	0.38014	0.01311	29	<2e-16 ***
Transportation assets	0.58163	0.01192	48.8	<2e-16 ***
(ϕ)	27.4331	0.4896	56.03	<2e-16 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Pseudo R-squared:	0.7594			

The results shown in Table 4.13 indicate that the model has improved as Pseudo R-squared changed from 0.2264 to 0.7594 and the precision parameter ϕ has increased. From the diagnostic plots given in Figure 4.13, it can be observed that the variance is approximately constant (Figure 4.13 (a), (c), (d)) and the distribution is approximately normal (Figure 4.13 (e)). Also, it can be noticed that there are no longer influential outliers in the improved model (Figure 4.13 (b)).

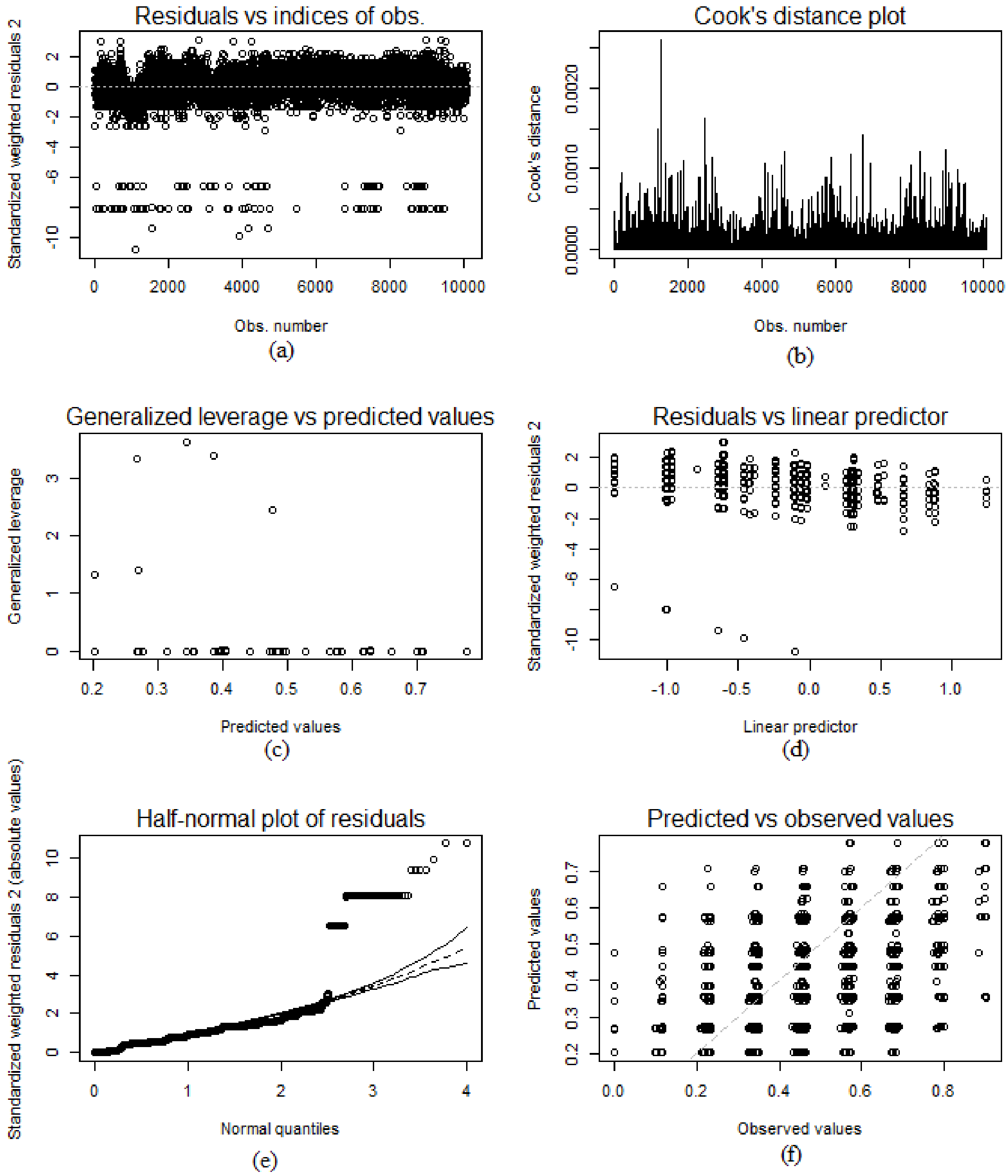


FIGURE 4.12: Diagnostic plots of the beta regression model

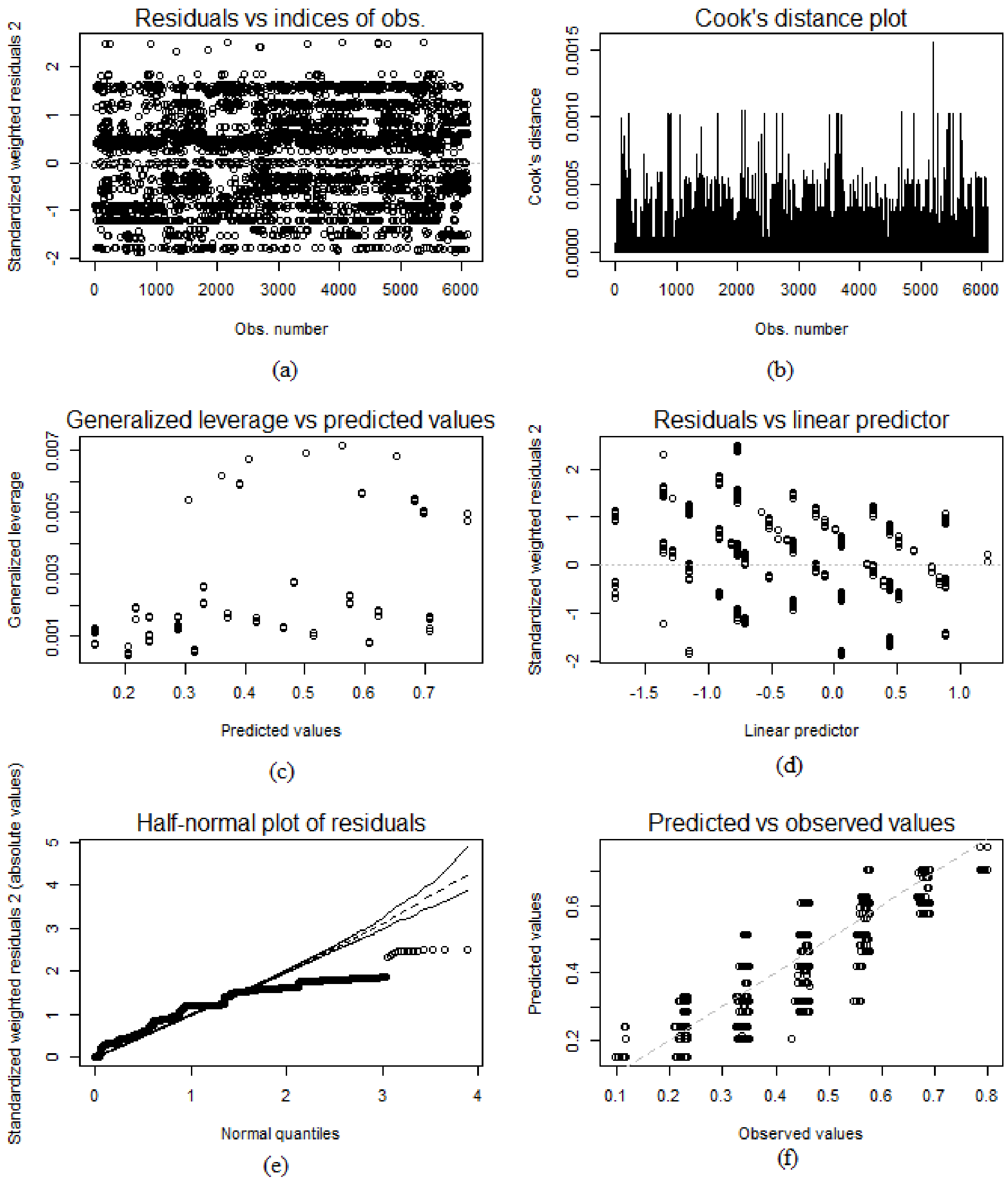


FIGURE 4.13: Diagnostic plots of the improved beta regression model

According to Ferrari and Cribari-Neto Ferrari & Cribari-Neto (2004)), the regression parameter estimate of the beta regression model is interpreted in terms of the odds ratio when the logit link function is used. Thus, each parameter was exponentiated and the results are given in Table 4.14. An individual who is not deprived in none of the indicators (i.e. assuming that all the indicators take zero values) has about 82% reduced chances of being multidimensionally deprived compared to an individual who is deprived in all indicators (OR: 0.1763, 95% CI: 0.1708 to 0.1820). A person who is deprived in housing is 3.3585 times more likely to be multidimensionally deprived compared to an individual who is not deprived in housing (OR: 3.3585, 95% CI: 3.2837 to 3.4350). On the other hand, the odds of being multidimensionally deprived among individuals who have no access to information/ICT is 1.3915 times higher than that of those who have access to information/ICT (OR: 1.3915, 95% CI: 1.3138 to 1.4739). The likelihood of an individual being multidimensionally deprived in regard to access to clinic or hospital is 1.5635 higher than those who are not deprived in regard to access to clinic or hospital (OR: 1.5635, 95% CI: 1.5214 to 1.6067). Individuals who are deprived in food security are 1.4625 times more likely to be multidimensionally deprived relative to those who are not deprived in food security (OR: 1.4625, 95% CI: 1.4254 to 1.5006). The likelihood of an individual who is deprived in transportation assets is 1.7890 more likely to be multidimensionally deprived compared to those who are not deprived in transportation assets (OR: 1.7890, 95% CI: 1.7476 to 1.8312) (Table 4.14).

All the indicators in this model (Table 4.13) are significantly associated with deprivation scores. These results concur with the results of Namibia Statistics Agency (2021b), where all these indicators were used in computing MPI for Namibia. The choice of using the housing indicator for MPI computation was motivated by the first objective of goal 11 of the Sustainable Development Goals (SDGs) which is to warranty access to suitable, safe, and reasonably priced housing and basic services and upgrade informal areas for all (Namibia Statistics Agency, 2021b). The Namibia fifth National Development Plan (NDP5) National Planning Commission (2017) aims for Namibia to have total access to information, inexpensive communication and technology infrastructure and services by 2022. Thus, the indicator "access to information or ICT" would help in identifying who is left behind in terms of access to information or ICT.

TABLE 4.14: Odds ratios of the estimated coefficients of the improved beta regression model and their 95% Confidence Interval

Indicators	Odds Ratio	95% CI: $exp(\hat{\beta}_i \pm 1.96\ S_{\hat{\beta}_i})$
$exp(\beta_0)$	0.1763	(0.1708, 0.1820)
Housing		
Non-deprived (Ref)	1.00	
Deprived	3.3585	(3.2837, 3.4350)
Access to information or ICT		
Non-deprived (Ref)	1.00	
Deprived	1.3915	(1.3138, 1.4739)
Access to clinic/hospital		
Non-deprived (Ref)	1.00	
Deprived	1.5635	(1.5214, 1.6067)
Food security		
Non-deprived (Ref)	1.00	
Deprived	1.4625	(1.4254, 1.5006)
Transportation assets		
Non-deprived (Ref)	1.00	
Deprived	1.789	(1.7476, 1.8312)

The transportation assets indicator is one of the key goals as stressed in the second objective of the SDG 11 with a goal of making transport systems sustainable, affordable, safe and available for all, especially those in susceptible situations such as the elderly, women, children and those with disabilities (Namibia Statistics Agency, 2021b). The food security indicator is connected to SDG 2 whose objective is to attain food security, eliminate hunger, improve nutrition and promote sustainable agriculture. Hence, it is important that this indicator is included in the computation of national MPI if the country is to trace the progress in achieving the second sustainable development goal.

From these results, it can be concluded that it is important to complement the "expert opinion method" of selecting indicators of an MPI with "statistical methods" such as beta regression in order to identify potential indicators that might be left out.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the conclusions resulting from the findings of this study on modelling poverty in Namibia using beta distribution. It also provides recommendations for further research directions.

5.2 Overview of the study

This study used the multidimensional poverty approach which measures poverty in more than one dimension. The study employed a beta distribution approach of modelling poverty where deprivation scores were considered to be the random variable in the interval (0, 1).

The modelling process involved the following steps. Firstly, we identified the variables of interest (indicators) from the data. Secondly, we determined the deprivation threshold for each indicator. A household is considered to be deprived in a specific indicator if their attainment in that specific indicator is below the deprivation threshold, and it is given a score of "1", otherwise, it is given a score of "0" if it not deprived. Thirdly, we determined and allocated weights to each indicator. This was done using two methods, namely, entropy and equal weighting methods. Fourthly, deprivation scores were computed based on the weights obtained. Each household receives a deprivation score between 0 and 1. Finally, a second deprivation threshold was determined following Alkire and Santos's guidelines to identify the multidimensionally poor households.

A household was classified as multidimensionally poor if their deprivation score is equal or above the threshold value (p), and their deprivation scores is replaced with "0". The deprivation scores were computed for both entropy and equal weighting methods. Then, these deprivation scores were used to compute MPI using the Alkire and Santos (2011) and the beta distribution model approaches. Under modelling MPI using beta distribution, we considered the left-truncated beta distribution (with a deprivation threshold value of p, positive shape parameters α and β) and MPI was found to be equivalent to the expected value of the left truncated beta distribution.

Using the entropy and equal weighting methods, we assessed how the change in these weighting methods affected the MPI values under each approach. Lastly, we identified other potential significant indicators that could have been left out in the computation of MPI using beta regression.

5.3 Review and evaluation of the objectives

- **Develop a multidimensional poverty model for Namibia using beta distribution**

This study aimed to develop a basic multidimensional poverty model for Namibia using beta distribution. Using the 2015/2016 Namibia Household Income and Expenditure Survey (NHIES) data, we were able to develop a multidimensional model for Namibia using beta distribution and we showed that the beta model can be used to estimate the MPI at regional and national levels.

- **Perform sensitivity analysis of weighting methods on the multidimensional poverty index for Namibia**

With this objective, we intended to asses how changes in weighting methods may affect MPI resulting from either the Alkire and Santos (2011) method or the beta distribution model. With the Alkire and Santos (2011) method, the results revealed that the overall MPI values were almost similar in both weighting methods with the entropy MPI value being slightly higher than that of equal weighting.

Also the deprivation score values under entropy weighting were low compared to equal weighting. The regional MPI values between the two weighting methods were not consistent with each other, as we observed that the pattern from the least to worst multidimensionally poor regions was different for both weighting methods. This showed that the methods are not consistent with each other in terms of MPI. On the other hand, under the beta distribution model approach, we observed a big difference between the MPI values for both methods with the equal weighting MPI values being bigger than those of the entropy weighting method which shows that changing weighting methods under the beta distribution model has a big effect of the MPI values. The overall MPI value for entropy under the beta distribution model was slightly similar to the overall MPI value for equal weighting under the Alkire and Santos (2011) method. We also observed that the MPI values under the Alkire and Santos (2011) method for both weighting methods were lower compared to the MPI values under the beta distribution model approach for both weighing methods. These differences may be explained by the fact that the beta parameters used to compute MPI are themselves estimates. The results for the worst and least deprived regions showed that the entropy weighting is more stable and consistent as its results concur with those of English (English, 2016).

- **Fit Beta regression model capable of determining other potential significant indicators of multidimensional poverty index for Namibia**

With this objective, we wanted to identify some potential indicators that might have been left out in the initial computation of MPI. In order to identify such indicators, we fitted a beta regression model.

5.4 Lessons learnt

With poverty increasing in Africa, there is a need for statistical methods of estimating poverty that can help decision makers to make more informed decisions and targeted interventions to alleviate poverty. The current methods that are used to estimate poverty such as the Alkire and Santos (2011) weighting method are based on experts' opinions. This method of assigning weights can be subjective and thus affects the MPI. For example, if an expert thinks that the electricity indicator is more important than the food security indicator, he/she will give the electricity indicator a bigger weight than the food security. But this might not necessarily be the case for everyone as others may perceive food security as more important than electricity. This weighting method (equal weighting) is also sensible to the number of indicators, as the change in the number of indicators affects its results. Hence there is a need for more objective methods of computing weights to reduce the bias associated with assigning weights. In this study, we developed a multidimensional poverty model capable of quantifying the uncertainty around the MPI using beta distribution. Entropy and equal weighting methods were assessed on how they affect the MPI. Also, through the beta regression model we found some additional indicators that were omitted in computing MPI.

The findings of this study can be used to compare multidimensional poverty levels among different regions in Namibia. They can also help policy makers to design appropriate targeted poverty interventions across the country through the identification of areas in severe multidimensional poverty conditions. Moreover, the developed model allows to quantify the uncertainty around the computed MPI through the specification of variance (and hence the confidence intervals), which is not the case for the Alkire and Santos (2011) approach.

Developing a multidimensional poverty model for Namibia using beta distribution

A multidimensional poverty model capable of estimating poverty for Namibia using beta distribution was developed. MPI was estimated as an expected value of the left truncated beta distribution. The model was then used to compute regional MPI and overall MPI for Namibia, which involved estimating the parameters of the left truncated beta distribution using maximum likelihood estimation, and specifying the deprivation threshold value as this model is dependent on the parameters α and β, and the threshold value p. We compared the beta distribution model with the Alkire and Santos (2011) approach of computing MPI and the results revealed that the MPI values for the Alkire and Santos (2011) approach were lower than those of the beta distribution model where big MPI values were observed. The difference in the MPI values between the two approaches may be a result of estimating the beta parameters used in computing MPI. The comparison of the two approaches across regions revealed that even though the MPI values for the beta distribution model were higher than those of the Alkire and Santos (2011) approach, the least and worst multidimensionally poor regions were the same in both approaches. This indicates that the two approaches point in the same direction with regards to multiple deprivation.

Performing sensitivity analysis of weighting methods on the multidimensional poverty index for Namibia

In this study, we considered two weighting methods (equal and entropy weighting) to asses their effect on MPI. At a threshold of 0.33, we observed a higher proportion of deprivation when using the equal weighting method as opposed to a low proportion when entropy was used. Hence, the 0.1172 threshold value for entropy was used to give an equivalent proportion to the equal weighting at a threshold value of 0.33. The results revealed that when allocating weights to the indicators under the equal weighting method, the indicators in a dimension with less than the traditional number of indicators (10 indicators) receive higher weights than the other indicators, and this has an effect on the deprivation scores. More specifically, households that are deprived in such an indicator end up having higher deprivation scores compared to households that are not deprived in that specific indicator. Also, the indicator that receives a higher weight becomes the most contributor to the deprivation scores and thus having an effect on MPI.

Removing such an indicator (e.g. nutrition indicator) from the model under both methods showed that the equal weighting method is highly affected by a change in the number of indicators compared to entropy weighting. This result lead to the conclusion that entropy is a better method than equal weighting as it is more stable and consistent irrespective of the data at hand.

Fitting a Beta regression model capable of determining other potential significant indicators of multidimensional poverty index for Namibia

We identified additional indicators that can be used in computing MPI which were not included in the initial computation of MPI by fitting a beta regression model. The results revealed that housing, access to information or ICT, access to clinic/hospital, food security, and transportation assets were statistically significant to be considered when computing MPI. In addition, the results revealed that housing had the greatest effect on the MPI while access to information or ICT had the least effect on MPI.

5.5 Recommendations

In this study, a multidimensional poverty model was developed using beta distribution. However, this model can only estimate the poverty levels but does not quantify the incidence and intensity of poverty. Future research studies could focus on developing a multidimensional poverty model using beta distribution that is capable of estimating the incidence and intensity of poverty. Also, since this study only looked at decomposition of MPI by region to estimate how each region contributes to the overall MPI, future researches could focus on the contribution of each indicator to the overall MPI for Namibia.

This study only used two weighting methods (entropy and equal weighting) in computing the Namibian MPI. The entropy weighting method was found to be better than equal weighting and hence future studies on MPI could use the entropy weighting method in computing MPI. Also, more weighting methods such as the ratio weighting method, point allocation weighting method, criteria importance through inter-criteria (CRITIC) weighting method, simple multi-attribute rating technique (SMART) weighting method, integrated weighting method, and mean weight (MW) weighting method could be investigated.

In order to identify statistically significant indicators that could have been left out in the initial computation of MPI, we fitted a beta regression model. This was performed by assuming our data was beta distributed. Our model utilized the logit link function to map a combination of covariate (indicators) values from a set of real numbers into the bounded interval [0, 1]. The logit link function was used in our study due to its simple interpretation as the logarithm of the odds ratio. Our study only used this link function and did not investigate the other link function to see how they affect the results of our model. It is in this regard that future studies could investigate the effect of other link functions such as the logit specification, the probit function, the complementary log-log link and the log-log link on the results of beta regression.

The indicators that were identified using the beta regression can be used as alternative indicators along each domain. These indicators can also be added to the indicators that were initially used to compute MPI. Having this additional indicators would bring another dimension to MPI which is mostly restrained to 10 indicators. Since this study only focused on national and regional MPI, future research can investigate the further decomposition of MPI to town, local authority or constituency levels.

Diagnostic measures were performed to check the goodness-of-fit of our model. Other aspects such as the correlation between the variables was not done. On the other hand, when we fitted the beta regression model we note that our link function is strictly monotonic and twice differentiable. We also noticed that changing the number of indicators in a domain has an effect on MPI. Some measures such as level of agreement of variables, unidimensionality and independence can further be studied for modelling MPI using beta distribution.

APPENDIX A

R codes used in Thesis

```
#Importing data with deprivation scores from R
library(readxl)
Final_Data_with_the_Deprivation_scores <- read_excel("C:/Users/Admin/Dropbox/Masters/
MASTERS THESIS/Data Sets/Final Data with the Deprivation scores.xlsx")
deprivation_scores_data <- read_excel("C:/Users/Admin/Dropbox/Masters/
MASTERS THESIS/Data Sets/Final Data with the Deprivation scores.xlsx")

#Density plots for the deprivation scores using entropy and equal weighting
plot(density(deprivation_scores_data$yi_Equal),ylim=c(0.0,3.5),lwd=2,pch=16,
xlab="Deprivation scores",ylab="Density",col="blue",main="",xlim=c(0.01,1))
lines(density(deprivation_scores_data$yi_Entropy),pch=16,lwd=2,col="red")
abline(v=0.3,pch=19,col="black",lty="solid",lwd=2)

#Adding a legend to the plot
legend("topright",
legend=c("Equal weighting method", "Entropy weighting method","Deprivation Threshold"),
col=c("blue","red","black"),
lty=1,lwd=2, cex=0.7, bg='lightblue')
#Importing new data set without the nutrition indicator from excel
Dep_scores_without_Nutrition<- read_excel("C:/Users/Admin/Dropbox/Masters/
MASTERS THESIS/Data Sets/Deprivation scores without Nutrition.xlsx")
#Density plots for the deprivation scores using
#entropy and equal weighting with and without the nutrition indicator
par(mfrow=c(1,2))
```

```
plot(density(deprivation_scores_data$yi_Equal),ylim=c(0.0,5),lwd=1.5,pch=16,
xlab="Deprivation scores (Equal)",ylab="Density",col="blue",main="",xlim=c(0.0,1.0))
lines(density(Dep_scores_without_Nutrition$yi_Equal_8),pch=16,lwd=1.5,col="red")
abline(v=0.3,pch=19,col="black",lty="solid",lwd=1.5)

#Adding a legend to the plot
legend("top",
legend=c("Deprivation scores with Nutrition",
 "Deprivation scores w/o Nutrition","Deprivation threshold"),
col=c("blue","red","black"),
lty=1,lwd=1.5, cex=0.7, bg='lightblue')

plot(density(deprivation_scores_data$yi_Entropy),ylim=c(0.0,5),lwd=1.5,pch=16,
xlab="Deprivation scores (Entropy)",ylab="Density",col="blue",main="",xlim=c(0.01,1))
lines(density(Dep_scores_without_Nutrition$yi_Entropy_8),pch=16,lwd=1.5,col="red")
abline(v=0.3,pch=19,col="black",lty="solid",lwd=1.5)

#Adding a legend to the plot
legend("top",
legend=c("Deprivation scores with Nutrition",
"Deprivation scores W/o Nutrition","Deprivation threshold"),
col=c("blue","red","black"),
lty=1,lwd=1.5, cex=0.7, bg='lightblue')

#Ploting boxplots for each indicator using entropy and equal weighting
par(mfrow=c(3,3))
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind1_Nutrition,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Nutrition",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind3_Yrs_sch,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Yrs of sch",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
```

```
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind4_Sch_att,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Sch Attendance",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind5_CookFuel,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Cooking fuel",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind6_sanitation,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Sanitation",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind7_drinking_H20,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Drinking Water",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind8_Electricity,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Electricity",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind9_Floortype,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Flooring",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Entropy~deprivation_scores_data$Ind10_assets,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Assets",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
par(mfrow=c(3,3))
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind1_Nutrition,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Nutrition",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind3_Yrs_sch,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Yrs of sch",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind4_Sch_att,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Sch Attendance",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind5_CookFuel,
```

```
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Cooking fuel",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind6_sanitation,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Sanitation",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind7_drinking_H20,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Drinking Water",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind8_Electricity,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Electricity",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind9_Floortype,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Flooring",col=c("red","orange"),notch=FALSE,horizontal=TRUE)
boxplot(deprivation_scores_data$yi_Equal~deprivation_scores_data$Ind10_assets,
xlab="Deprivation Score", ylab="Deprivation indicator ",border="brown",
main="Assets",col=c("red","orange"),notch=FALSE,horizontal=TRUE)

#########################################################################################
#Getting Quantiles of the deprivation scores
quantile(deprivation_scores_data$yi_Equal, probs = c(.17001,1-0.17001))
quantile(deprivation_scores_data$yi_Entropy, probs = c(.17001,1-0.17001))
par(mfrow=c(1,1))
#Density plot for the deprivation scores using entropy weighting with the new
#recalibrated threshold for with and without the nutrition indicator
plot(density(deprivation_scores_data$yi_Entropy),ylim=c(0.0,4.0),lwd=1.5,pch=19,
xlab="Deprivation scores (Entropy)",ylab="Density",col="blue",main="",xlim=c(0.01,1))
lines(density(Dep_scores_without_Nutrition$yi_Entropy_8),pch=19,
xlim=c(0,1),lwd=1.5,col="red")
abline(v=0.1172192,pch=19,col="black",lty="solid",lwd=1.5)

#Adding a legend to the plot
```

```
legend("topright",
legend=c("Deprivation scores with Nutrition",
"Deprivation scores W/o Nutrition","Deprivation threshold"),
col=c("blue","red","black"),
lty=1,lwd=1.5, cex=0.7, bg='lightblue')

#Checking the number of values above the thresholds
length( which( deprivation_scores_data$yi_Entropy > 0.1172 ) )
length( which( deprivation_scores_data$yi_Equal> 0.3 ) )

##################################################################################
#Ploting the Namibian maps for average deprivations using entropy
#and equal weighting methods
par(mfrow=c(1,1))

#loading packages to be used in the maps
library(ggplot2)
library(ggmap)
library(maps)
library(mapdata)
library(sp)

#Downloading the Namibian map file
link <- "http://biogeo.ucdavis.edu/data/gadm2.8/rds/NAM_adm1.rds"
download.file(url=link, destfile="file.rda", mode="wb")
gadmNab <- readRDS("file.rda")

#Ploting the map for Namibia
library(rmapshaper)
gadmNabS <- ms_simplify(gadmNab, keep = 0.02)
plot(gadmNabS)
```

```
library(broom)
library(gpclib)
library(ggplot2)
namMapsimple <- tidy(gadmNabS, region="NAME_1")
ggplot(data = namMapsimple,
aes(x = long, y = lat, group = group)) +
geom_polygon(aes(fill = "red"))

# check region names
unique(namMapsimple$id)
# the exclamation mark before Karas will cause a problem.
namMapsimple$id <- gsub("!", "", namMapsimple$id)
# check it worked
unique(namMapsimple$id)
# now no exclamation mark
unique(deprivation_scores_data$Region)

#combining "Kavango East" "Kavango West" to Kavango
deprivation_scores_data$Region[deprivation_scores_data$Region == "Kavango West"] <-
"Kavango"
deprivation_scores_data$Region[deprivation_scores_data$Region == "Kavango East"] <-
"Kavango"

#Subset only the data that is needed for the maps
data<-deprivation_scores_data[,c(3,4,15,16)]
data_means<-data.frame(aggregate(data[, 3:4], list(data$Region), mean))
data_means_All<-data.frame(aggregate(data[, 3:4],
list(data$Region, data$Urban_Rural), mean))
data_means_All
data_means
## step 4: let's get our data and bind it...
# download the data about urbanisation
```

```
# change colnames so that we can merge by id...

data_col_names1 <- colnames(data)
data_col_names <- colnames(data)
data_col_names[1] <- c("id")
data_col_names1[1] <- c("id")
data_col_names

colnames(data_means_All) <- data_col_names1

# bind the scores data into map with merge
mapNabValsAll <- merge(namMapsimple,
data_means_All,
by.y = 'id')

p_All <- ggplot() +     # sometimes don't add any aes...
geom_polygon(data = mapNabValsAll,
aes(x = long, y = lat,
group = group,
fill = yi_Equal ),color="black", size = 0.2)+
scale_fill_continuous(name="Deprevation Score", low = "lightblue",
high = "darkblue",limits = c(0.2,0.6),
breaks=c(0.33,0.36,0.39,0.42,0.45,0.48,0.51,0.54,0.57)) +
labs(title="Equal Weighting")

#########################################
#ENTHROPY WEIGHIT
p_All1 <- ggplot() +     # sometimes don't add any aes...
geom_polygon(data = mapNabValsAll,
aes(x = long, y = lat,
group = group,
```

```
fill = yi_Entropy ),color="black", size = 0.2)+
scale_fill_continuous(name="Deprevation Score", low = "white",
high = "darkred",limits = c(0,0.4), breaks=c(0.05,0.1,0.15,0.2,0.25,0.3,0.35)) +
labs(title="Enthropy Weighting")

library(gridExtra)
grid.arrange( p_All, p_All1,ncol=2, nrow=1)
#########################################################################################
#Installing packages to be used in fitting the beta distribution
install.packages("betafunctions")
#loading packages to be used
library(fBasics)
library(fitdistrplus)
library(stats4)
library(EnvStats)
library(betafunctions)

#Estimating the parameters alpha and beta using MLE from entropy data set
testdata <-deprivation_scores_data$yi_Entropy
bmle<-ebeta(testdata, method = "mle")
bmle
bmle$parameters
testdata1 <-deprivation_scores_data$yi_Equal
bmle1<-ebeta(testdata1, method = "mmue")
bmle1
bmle1$parameters
testdata2 <-deprivation_scores_data$‘Censored_yi_Entropy(0.1172)‘
bmle2<-ebeta(testdata2, method = "mle")
bmle2
bmle2$parameters

beta.mle(deprivation_scores_data$yi_Entropy, tol=1e-09)
```

```
#Importing REgional deprivation scores data
Regional_MPI_Final_Deprivation_scores <- read_excel("C:/Users/Admin/Dropbox/Masters/
MASTERS THESIS/Data Sets/Regional MPI Final Deprivation scores.xlsx")
Regional_dep_scores<-Regional_MPI_Final_Deprivation_scores

#Estimating the parameters alpha and beta using MLE for each region
Zambezi_Region_MPI_Final_Deprivation_scores <- read_excel("C:/Users/Admin/Dropbox/Masters/
MASTERS THESIS/Data Sets/Zambezi Region MPI Final Deprivation scores.xlsx")
Zambezi_Region_dep_scores<-Zambezi_Region_MPI_Final_Deprivation_scores
Zambezi<-Zambezi_Region_dep_scores$yi_Equal
Zambezibmle<-ebeta(Zambezi, method = "mle")
Zambezibmle
Zambezi1<-Zambezi_Region_dep_scores$yi_Entropy
Zambezi1bmle<-ebeta(Zambezi1, method = "mle")
Zambezi1bmle
Zambezibmle$parameters
Zambezi1bmle$parameters

Karas_Region_MPI_Final_Deprivation_scores <- read_excel("C:/Users/Admin/Dropbox/Masters/
MASTERS THESIS/Data Sets/Karas Region MPI Final Deprivation scores.xlsx")
Karas_Region_dep_scores<-Karas_Region_MPI_Final_Deprivation_scores
Karas<-Karas_Region_dep_scores$yi_Equal
Karasbmle<-ebeta(Karas, method = "mle")
Karasbmle
Karas1<-Karas_Region_dep_scores$yi_Entropy
Karas1bmle<-ebeta(Karas1, method = "mle")
Karas1bmle
Karasbmle$parameters
Karas1bmle$parameters

Erongo_Region_dep_scores<-Erongo_Region_MPI_Final_Deprivation_scores
```

```
Erongo<-Erongo_Region_dep_scores$yi_Equal
Erongobmle<-ebeta(Erongo, method = "mle")
Erongobmle
Erongo1<-Erongo_Region_dep_scores$yi_Entropy
Erongo1bmle<-ebeta(Erongo1, method = "mle")
Erongo1bmle
Erongobmle$parameters
Erongo1bmle$parameters

Hardap_Region_dep_scores<-Hardap_Region_MPI_Final_Deprivation_scores
Hardap<-Hardap_Region_dep_scores$yi_Equal
Hardapbmle<-ebeta(Hardap, method = "mle")
Hardapbmle
Hardap1<-Hardap_Region_dep_scores$yi_Entropy
Hardap1bmle<-ebeta(Hardap1, method = "mle")
Hardap1bmle
Hardapbmle$parameters
Hardap1bmle$parameters

Kavango_East_Region_dep_scores<-Kavango_East_Region_MPI_Final_Deprivation_scores
Kavango_East<-Kavango_East_Region_dep_scores$yi_Equal
Kavango_Eastbmle<-ebeta(Kavango_East, method = "mle")
Kavango_Eastbmle
Kavango_East1<-Kavango_East_Region_dep_scores$yi_Entropy
Kavango_East1bmle<-ebeta(Kavango_East1, method = "mle")
Kavango_East1bmle
Kavango_Eastbmle$parameters
Kavango_East1bmle$parameters

Kavango_West_Region_dep_scores<-Kavango_West_Region_MPI_Final_Deprivation_scores
Kavango_West<-Kavango_West_Region_dep_scores$yi_Equal
Kavango_Westbmle<-ebeta(Kavango_West, method = "mle")
```

```
Kavango_Westbmle
Kavango_West1<-Kavango_West_Region_dep_scores$yi_Entropy
Kavango_West1bmle<-ebeta(Kavango_West1, method = "mle")
Kavango_West1bmle
Kavango_Westbmle$parameters
Kavango_West1bmle$parameters


Khomas_Region_dep_scores<-Khomas_Region_MPI_Final_Deprivation_scores
Khomas<-Khomas_Region_dep_scores$yi_Equal
Khomasbmle<-ebeta(Khomas, method = "mle")
Khomasbmle
Khomas1<-Khomas_Region_dep_scores$yi_Entropy
Khomas1bmle<-ebeta(Khomas1, method = "mle")
Khomas1bmle
Khomasbmle$parameters
Khomas1bmle$parameters

Kunene_Region_dep_scores<-Kunene_Region_MPI_Final_Deprivation_scores
Kunene<-Kunene_Region_dep_scores$yi_Equal
Kunenebmle<-ebeta(Kunene, method = "mle")
Kunenebmle
Kunene1<-Kunene_Region_dep_scores$yi_Entropy
Kunene1bmle<-ebeta(Kunene1, method = "mle")
Kunene1bmle
Kunenebmle$parameters
Kunene1bmle$parameters

Ohangwena_Region_dep_scores<-Ohangwena_Region_MPI_Final_Deprivation_scores
Ohangwena<-Ohangwena_Region_dep_scores$yi_Equal
Ohangwenabmle<-ebeta(Ohangwena, method = "mle")
Ohangwenabmle
```

```
Ohangwena1<-Ohangwena_Region_dep_scores$yi_Entropy

Ohangwena1bmle<-ebeta(Ohangwena1, method = "mle")

Ohangwena1bmle

Ohangwenabmle$parameters

Ohangwena1bmle$parameters


Omaheke_Region_dep_scores<-Omaheke_Region_MPI_Final_Deprivation_scores

Omaheke<-Omaheke_Region_dep_scores$yi_Equal

Omahekebmle<-ebeta(Omaheke, method = "mle")

Omahekebmle

Omaheke1<-Omaheke_Region_dep_scores$yi_Entropy

Omaheke1bmle<-ebeta(Omaheke1, method = "mle")

Omaheke1bmle

Omahekebmle$parameters

Omaheke1bmle$parameters


Omusati_Region_dep_scores<-Omusati_Region_MPI_Final_Deprivation_scores

Omusati<-Omusati_Region_dep_scores$yi_Equal

Omusatibmle<-ebeta(Omusati, method = "mle")

Omusatibmle

Omusati1<-Omusati_Region_dep_scores$yi_Entropy

Omusati1bmle<-ebeta(Omusati1, method = "mle")

Omusati1bmle

Omusatibmle$parameters

Omusati1bmle$parameters


Oshana_Region_dep_scores<-Oshana_Region_MPI_Final_Deprivation_scores

Oshana<-Oshana_Region_dep_scores$yi_Equal

Oshanabmle<-ebeta(Oshana, method = "mle")

Oshanabmle

Oshana1<-Oshana_Region_dep_scores$yi_Entropy

Oshana1bmle<-ebeta(Oshana1, method = "mle")
```

```
Oshana1bmle
Oshanabmle$parameters
Oshana1bmle$parameters

Oshikoto_Region_dep_scores<-Oshikoto_Region_MPI_Final_Deprivation_scores
Oshikoto<-Oshikoto_Region_dep_scores$yi_Equal
Oshikotobmle<-ebeta(Oshikoto, method = "mle")
Oshikotobmle
Oshikoto1<-Oshikoto_Region_dep_scores$yi_Entropy
Oshikoto1bmle<-ebeta(Oshikoto1, method = "mle")
Oshikoto1bmle
Oshikotobmle$parameters
Oshikoto1bmle$parameters

Otjozondjupa_Region_dep_scores<-Otjozondjupa_Region_MPI_Final_Deprivation_scores
Otjozondjupa<-Otjozondjupa_Region_dep_scores$yi_Equal
Otjozondjupabmle<-ebeta(Otjozondjupa, method = "mle")
Otjozondjupabmle
Otjozondjupa1<-Otjozondjupa_Region_dep_scores$yi_Entropy
Otjozondjupa1bmle<-ebeta(Otjozondjupa1, method = "mle")
Otjozondjupa1bmle
Otjozondjupabmle$parameters
Otjozondjupa1bmle$parameters

#Plotting the fitted basic beta distribution using
#the estimated parameters alpha and beta
curve(dbeta(x, 1.872071, 2.634632), pch=15, ylab="Density", xlab="Deprivation scores",
col=c("black"), ylim=c(0.0,2.0), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)

#switching the values of alpha and beta
curve(dbeta(x, 2.634632, 1.872071), pch=15, ylab="Density", xlab="Deprivation scores",
```

```
col=c("black"), ylim=c(0.0,2.0), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)

#Running simmulations of the beta distributions
#with different values of alpha and beta
curve(dbeta(x, 1.872071, 2.634632), col=c("black"), pch=19, ylab="Density",
xlab="Deprivation scores", ylim=c(0.0,2.5), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)
curve(dbeta(x, 1.872071, 3.0), col=c("red"), pch=19, lty="dashed",lwd=1.5, add = TRUE)
curve(dbeta(x, 1.872071, 3.5), col=c("blue"), pch=19, lty="dotted",lwd=1.5, add = TRUE)
curve(dbeta(x, 1.872071, 4.0), col=c("gold"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
#Adding a legend to the plot
legend(0.55,2.59, bty="n",
legend=c(expression(paste(alpha ==1.872071, "," ,beta ==2.634632)),
expression(paste(alpha ==1.872071, "," ,beta ==3.0)),
expression(paste(alpha ==1.872071, "," ,beta ==3.5)),
expression(paste(alpha ==1.872071, "," ,beta ==4.0))),
lty=c("solid","dashed","dotted","dotdash"),col=c("black","red", "blue", "gold"),
lwd=1.5, cex=0.8, bg='lightblue')

curve(dbeta(x, 1.872071, 2.634632), col=c("black"), pch=19, ylab="Density",
xlab="Deprivation scores", ylim=c(0.0,3.5), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)
curve(dbeta(x, 2.5, 2.634632), col=c("red"), pch=19, lty="dashed",lwd=1.5, add = TRUE)
curve(dbeta(x, 3.0, 2.634632), col=c("blue"), pch=19, lty="dotted",lwd=1.5, add = TRUE)
curve(dbeta(x, 3.5, 2.634632), col=c("gold"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
legend(0.55,3.59, bty="n",
legend=c(expression(paste(alpha ==1.872071, "," ,beta ==2.634632)),
expression(paste(alpha ==2.5, "," ,beta ==2.634632)),
expression(paste(alpha ==3.0, "," ,beta ==2.634632)),
expression(paste(alpha ==3.5, "," ,beta ==2.634632))),
lty=c("solid","dashed","dotted","dotdash"),col=c("black","red", "blue", "gold"),
```

```
lwd=1.5, cex=0.8, bg='lightblue')

curve(dbeta(x, 1.872071, 1.872071), col=c("black"), pch=19, ylab="Density",
xlab="Deprivation scores", ylim=c(0.0,3.5), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)
curve(dbeta(x, 3, 3), col=c("red"), pch=19, lty="dashed",lwd=1.5, add = TRUE)
curve(dbeta(x, 4, 4), pch=19, col=c("blue"), lty="dotted",lwd=1.5, add = TRUE)
curve(dbeta(x, 5, 5), col=c("gold"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
legend(0.55,3.59, bty="n",
legend=c(expression(paste(alpha, "=" ,beta ==1.872071)),
expression(paste(alpha, "=" ,beta ==3)),
expression(paste(alpha, "=" ,beta ==4)),
expression(paste(alpha, "=" ,beta ==5))),
lty=c("solid","dashed","dotted","dotdash"),col=c("black","red", "blue", "gold"),
lwd=1.5, cex=0.8, bg='lightblue')

curve(dbeta(x, 1.872071, 2.634632), col=c("black"), pch=19, ylab="Density",
xlab="Deprivation scores", ylim=c(0.0,3.5), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)
curve(dbeta(x, 2.5, 3.0), col=c("red"), pch=19, lty="dashed",lwd=1.5, add = TRUE)
curve(dbeta(x, 3.5, 4.0), col=c("blue"), pch=19, lty="dotted",lwd=1.5, add = TRUE)
curve(dbeta(x, 4.5, 5.0), col=c("gold"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
legend(0.55,3.59, bty="n",
legend=c(expression(paste(alpha ==1.872071, "," ,beta ==2.634632)),
expression(paste(alpha ==2.5, "," ,beta ==3.0)),
expression(paste(alpha ==3.5, "," ,beta ==4.0)),
expression(paste(alpha ==4.5, "," ,beta ==5.0))),
lty=c("solid","dashed","dotted","dotdash"),col=c("black","red", "blue", "gold"),
lwd=1.5, cex=0.8, bg='lightblue')

curve(dbeta(x, 1.872071, 2.634632), col=c("black"),pch=19, ylab="Density",
xlab="Deprivation scores", ylim=c(0.0,2.5), lty="solid",lwd=1.5)
```

```
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)
curve(dbeta(x, 0.8, 1.5), col=c("red"), pch=19, lty="dashed",lwd=1.5, add = TRUE)
curve(dbeta(x, 0.5, 1.0), col=c("blue"), pch=19, lty="dotted",lwd=1.5, add = TRUE)
curve(dbeta(x, 0.2, 0.5), col=c("gold"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
curve(dbeta(x, 0.1, 0.3), col=c("cyan"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
legend(0.55,2.59, bty="n",
legend=c(expression(paste(alpha ==1.872071, "," ,beta ==2.634632)),
expression(paste(alpha ==0.8, "," ,beta ==1.5)),
expression(paste(alpha ==0.5, "," ,beta ==1.0)),
expression(paste(alpha ==0.2, "," ,beta ==0.5)),
expression(paste(alpha ==0.1, "," ,beta ==0.3))),
lty=c("solid","dashed","dotted","dotdash"),col=c("black","red", "blue", "gold","cyan"),
lwd=1.5, cex=0.8, bg='lightblue')

curve(dbeta(x, 1.872071, 2.634632), col=c("black"),pch=19, ylab="Density",
xlab="Deprivation scores", ylim=c(0.0,2.5), lty="solid",lwd=1.5)
abline(v=0.1172,pch=19,col="black",lty="solid",lwd=1.5)
curve(dbeta(x, 1.0, 3.0), col=c("red"), pch=19, lty="dashed",lwd=1.5, add = TRUE)
curve(dbeta(x, 0.5, 3.5), col=c("blue"), pch=19, lty="dotted",lwd=1.5, add = TRUE)
curve(dbeta(x, 0.2, 4.0), col=c("gold"), pch=19, lty="dotdash",lwd=1.5, add = TRUE)
legend(0.55,2.59, bty="n",
legend=c(expression(paste(alpha ==1.872071, "," ,beta ==2.634632)),
expression(paste(alpha ==1.0, "," ,beta ==3.0)),
expression(paste(alpha ==0.5, "," ,beta ==3.5)),
expression(paste(alpha ==0.2, "," ,beta ==4.0))),
lty=c("solid","dashed","dotted","dotdash"),col=c("black","red", "blue", "gold"),
lwd=1.5, cex=0.8, bg='lightblue')

#plotting the graph/plot showing the overall effect of alpha and beta on MPI
install.packages('rgl')
library(rgl)
knit_hooks$set(rgl = hook_rgl, webgl = hook_webgl)
```

```
p<-c(0.1172)
alpha<- seq(1, 10, length= 20)
beta <- seq(1, 10, length= 20)
cone <- function(alpha, beta, p){
(alpha+(p*beta))/(alpha+beta)
}
MPI <- outer(alpha,beta, cone,p)
persp(alpha,
beta,
MPI,
col = "springgreen", shade = 0.3, ylab="beta", zlab="MPI",
theta = 90, phi = -1, expand = 0.5)

##########################################################################################
####  BETA REGRESSION

#Installing packages used in beta regression
install.packages('betareg')

if(!require(psych)){install.packages("psych")}
if(!require(betareg)){install.packages("betareg")}
if(!require(lmtest)){install.packages("lmtest")}
if(!require(rcompanion)){install.packages("rcompanion")}
if(!require(multcompView)){install.packages("multcompView")}
if(!require(emmeans)){install.packages("emmeans")}
if(!require(ggplot2)){install.packages("ggplot2")}
library(betareg)
library(psych)
library(lmtest)
library(rcompanion)
library(multcompView)
library(emmeans)
```

```
library(ggplot2)

#Importing data for beta regression
library(readxl)
FINAL_Indicators_Data_set_used_in_BETA_REGRESSION <- read_excel("C:/Users/Admin/Dropbox/
Masters/MASTERS THESIS/Data Sets/FINAL Indicators Data set used in BETA REGRESSION.xlsx")
View(FINAL_Indicators_Data_set_used_in_BETA_REGRESSION)
Beta<-FINAL_Indicators_Data_set_used_in_BETA_REGRESSION

#Fitting the beta regression model1
Betareg<-betareg(data=Beta,Transformed_yi_entropy~Ind11_Housing +
Ind12_Access_to_information_or_ICT + Ind13_Access_to_clinic_or_hospital +
Ind14_Food_security +Ind17_Transportation_assets)
summary(Betareg)

#Diagnositic plots for beta regression model1
par(mfrow = c(3, 2))
plot(Betareg, which = 1:6)

#Removing the most influential observations in regression model1
par(mfrow = c(1, 1))
cooksd <- cooks.distance(Betareg)
sample_size <- nrow(Beta)
plot(cooksd, pch="*", cex=2, main="Influential Obs by Cooks distance")
abline(h = 4/sample_size, col="red")
text(x=1:length(cooksd)+1, y=cooksd,
labels=ifelse(cooksd>4/sample_size, names(cooksd),""), col="red")
influential <- as.numeric(names(cooksd)[(cooksd > (4/sample_size))])

#Reffiting the model without the initial influential observations
Beta21<- Beta[-influential, ]
Betareg20<-betareg(data=Beta21,Transformed_yi_entropy~Ind11_Housing +
```

```
Ind12_Access_to_information_or_ICT + Ind13_Access_to_clinic_or_hospital +
Ind14_Food_security +Ind17_Transportation_assets)
summary(Betareg20)

#Re-runing the diagnostics
par(mfrow = c(3, 2))
plot(Betareg, which = 1:6)

#Removing more influential observations
par(mfrow = c(1, 1))
cooksd21 <- cooks.distance(Betareg20)
sample_size <- nrow(Beta22)
plot(cooksd21, pch="*", cex=2, main="Influential Obs by Cooks distance")
abline(h = 0.00007, col="red")
text(x=1:length(cooksd21)+1, y=cooksd21,
labels=ifelse(cooksd21>0.00007, names(cooksd),""), col="red")
influential21 <- as.numeric(names(cooksd21)[(cooksd21 > (0.00007))])

#Reffiting the model without the most influential observations and running diagnostics
Beta22<- Beta21[-influential21, ]
Betareg22<-betareg(data=Beta22,Transformed_yi_entropy~Ind11_Housing +
Ind12_Access_to_information_or_ICT + Ind13_Access_to_clinic_or_hospital +
Ind14_Food_security +Ind17_Transportation_assets)
summary(Betareg22)
par(mfrow = c(3, 2))
plot(Betareg22, which = 1:6)

#Exponentiating the coefficients of the final beta model to get the odds ratio
X0<-exp(-1.73557)     #X0 is the intercept
X0

X1<-exp(1.21148)     #X1 is the the housing indicator
```

```
X1

X2<-exp(0.33040)      #X2 is the access to information or ICT indicator
X2

X3<-exp(0.44691)     #X3 is the access to clinic or hospital indicator
X3

X4<-exp(0.38014)   #X4 is the food security indicator
X4

X5<-exp(0.58163)  #X5 is the transportation assets indicator
X5

v<-confint(Betareg22)
v
X6<-exp(0.6049863)
X6
```

BIBLIOGRAPHY

Akram, S. & Ann, Q. (2015). Newton raphson method. *International Journal of Scientific & Engineering Research*, 6(7), 1748–1752.

Alkire, S. & Jahan, S. (2018). The new global mpi 2018: aligning with the sustainable development goals. Oxford Poverty and Human Development Initiative (OPHI).

Alkire, S., Roche, J. M., Ballon, P., Foster, J., Santos, M. E., & Seth, S. (2015). *Multidimensional poverty measurement and analysis*. Oxford University Press, USA.

Alkire, S., Roche, J. M., Santos, M. E., & Seth, S. (2011). Multidimensional poverty index 2011: brief methodological note.

Alkire, S. & Santos, M. E. (2011). Training material for producing national human development reports: The multidimensional poverty index (mpi).

Bonat, W. H., Petterle, R. R., Hinde, J., & Demétrio, C. G. (2019). Flexible quasi-beta regression models for continuous bounded data. *Statistical Modelling*, 19(6), 617–633.

Boundless (2014). Measuring poverty - boundless open textbook. https://web.archive.org/web/20150402090643/https://www.boundless.com/sociology/textbooks/bo sociology-textbook/stratification-inequality-and-social-class-in-the-u-s-9/poverty-78/measuring-poverty-463-3312/.

Cameron, D. (2011). David cameron: Free trade in africa shows a way out of poverty. https://www.globalpolicy.org/social-and-economic-policy/poverty-and-development/poverty-and-development-in-africa/50467-david-cameron-free-trade-in-africa-shows-a-way-out-of-poverty-.html?itemid=id969.

Cavapozzi, D., Han, W., & Miniaci, R. (2015). Alternative weighting structures for multidimensional poverty assessment. *The Journal of Economic Inequality*, 13(3), 425–447.

Chen, L. & Singh, V. P. (2017). Generalized beta distribution of the second kind for flood frequency analysis. *Entropy*, 19(6), 254.

Chetty, P. (2016). Importance of ethical considerations in a research. https://www.projectguru.in/importance-ethical-considerations-research/.

Chotikapanich, D., Griffiths, W. E., Rao, D., & Karunarathne, W. (2014). Income distributions, inequality, and poverty in asia, 1992–2010.

Cribari-Neto, F. & Zeileis, A. (2010). Beta regression in r. *Journal of statistical software*, 34(1), 1–24.

English, J. (2016). Inequality and poverty in namibia: A gaping wealth gap. https://borgenproject.org/inequality-and-poverty-in-namibia/.

Ferrari, S. & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of applied statistics*, 31(7), 799–815.

Gao, S. & Sun, K. (2020). Poverty measure based on hesitant fuzzy decision algorithm under social network media. *Symmetry*, 12(3), 384.

Glen, S. (2015). Method of moments definition and example. https://www.statisticshowto.com/method-moments/.

Hamel, K., Tong, B., & Hofer, M. (2019). Poverty in africa is now falling–but not fast enough. *Brookings Institution,[www. brookings. edu]*.

Johnson, N., Kotz, S., & Balakrishnan, N. (1994). Beta distributions. *Continuous univariate distributions. 2nd ed. New York, NY: John Wiley and Sons*, (pp. 221–235).

Johnson, R. W. (2013). Applications of the beta distribution part 1: Transformation group approach. *arXiv preprint arXiv:1307.6437*.

Jose, T. (2015). What is purchasing power parity? https://www.indianeconomy.net/splclassroom/what-is-purchasing-power-parity/.

Kadapa, V. (2020). Entropy method for weight in multi-criteria decision making objective weight estimation in mcdm. https://www.youtube.com/watch?v=8OeXP9tAadM.

Kenton, W. (2021). What is the international poverty line? https://www.investopedia.com/terms/i/international-poverty-line.asp.

Klasen, S. (2014). Measuring poverty and inequality in sub-saharan africa: Knowledge gaps and ways to address them. *Africa Can End Poverty*.

Korn, R. (2017). Wealth and poverty in namibia. https://www.borgenmagazine.com/wealth-and-poverty-in-namibia/.

Löffler, G. & Posch, P. N. (2011). *Credit Risk Modeling Using Excel and VBA with DVD*. John Wiley Sons, Ltd.

Massuanganhe, I. J. (2005). Modelling prsp ii & poverty reduction in mozambique local development: Econometric analysis of factors determining millennium. *Governance*.

Mazucheli, J. & Menezes, A. F. B. (2019). L-moments and maximum likelihood estimation for the complementary beta distribution with applications on temperature extremes. *Journal of Data Science*, 17(2).

Namibia Statistics Agency (2012). Poverty dynamics in namibia: A comparative study using the 1993/94, 2003/04 and the 2009/10 nhies surveys. Namibia Statistics Agency.

Namibia Statistics Agency (2015). Namibia household income and expenditure survey 2015/2016 interviewer manual. Namibia Statistics Agency.

Namibia Statistics Agency (2016). Namibia household income and expenditure survey (nhies) 2015/2016 key poverty indicators (preliminary figures). Namibia Statistics Agency.

Namibia Statistics Agency (2021a). Errata on 2015/2016 nhies report. Namibia Statistics Agency.

Namibia Statistics Agency (2021b). Namibia multidimensional poverty index (mpi) report 2021. Namibia Statistics Agency.

National Planning Comission (2008). A review of poverty and inequality in namibia. Central Bureau OF Statistics.

National Planning Commission (2015). Namibia index of multiple deprivation. Republic of Namibia National Planning Commission of Namibia.

National Planning Commission (2016). Namibia poverty mapping. National Planning Commission.

National Planning Commission (2017). Namibia's 5th national development plan (ndp5). Windhoek: National Planning Commission.

Ng, D., Koh, S., Sim, S., & Lee, M. (2018). The study of properties on generalized beta distribution. In *Journal of Physics: Conference Series*, volume 1132 (pp. 012080).: IOP Publishing.

Ospina, R. & Ferrari, S. L. (2010). Inflated beta distributions. *Statistical Papers*, 51(1), 111–115.

Owen, C. B. (2008). *Parameter estimation for the beta distribution*. Brigham Young University.

Oxford Poverty and Human Development Initiative (2017). Ophi country briefing 2017: Namibia global multidimensional poverty index (mpi) at a glance. http://www.ophi.org.uk/multidimensional-poverty-index/mpi-country-briefings/.

Oxford Poverty and Human Development Initiative (2019a). Global multidimensional poverty index 2019: illuminating inequalities. Oxford Poverty and Human Development Initiative (OPHI).

Oxford Poverty and Human Development Initiative (2019b). Sierra leone multidimensional poverty index 2019. Oxford Poverty and Human Development Initiative.

Packtor, J. (2017). 10 shocking facts about poverty in africa. https://borgenproject.org/10-quick-facts-about-poverty-in-africa/.

Patel, N. (2018). Figure of the week: Understanding poverty in africa. https://www.brookings.edu/blog/africa-in-focus/2018/11/21/figure-of-the-week-understanding-poverty-in-africa/.

Peer, A. (2018). Global poverty: facts, faqs, and how to help. *World Vision. https://www. worldvision. org/sponsorship-news-stories/global-poverty-facts# end*.

Pogge, T. (2015). Measuring poverty. https://web.archive.org/web/20160129022900/http://thomaspogge pogge/measuring-poverty/.

Quinn, K. (2001). The newton raphson algorithm for function optimization. *Seattle: The Center for Statistics and the Social Sciences, University of Washington*, 6.

Ravallion, M. (2011). On multidimensional indices of poverty. *The Journal of Economic Inequality*, 9(2), 235–248.

Satapathy, S. S. & Jaiswal, K. K. (2018). A study on poverty estimation and current state of poverty in india. *International Journal of Advanced Scientific Research and Management*, 3(6), 86–92.

Schmidt, M. (2009). Measuring poverty in namibia. https://www.namibian.com.na/index.php?id=53477page=archive-read.

Schröder, C. & Rahmann, S. (2017). A hybrid parameter estimation algorithm for beta mixtures and applications to methylation state classification. *Algorithms for Molecular Biology*, 12(1), 21–24.

Silver, M. & Gharib, M. (2017). What's the meaning of the world bank's new poverty lines? https://www.npr.org/sections/goatsandsoda/2017/10/25/558068646/whats-the-meaning-of-the-world-banks-new-poverty-lines.

Singh, B., Pudir, P., & Maheshwari, S. (2014). Parameter estimation of beta-geometric model with application to human fecundability data. *arXiv preprint arXiv:1405.6392*.

Smithson, M. & Verkuilen, J. (2006). A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*, 11(1), 54.

Sustainable Development Solutions Network (2020). Multidimensional poverty index. https://indicators.report/indicators/i-3/.

The World Bank (2022). Lifting 800 million people out of poverty – new report looks at lessons from china's experience. https://www.worldbank.org/en/news/press-release/2022/04/01/lifting-800-million-people-out-of-poverty-new-report-looks-at-lessons-from-china-s-experience.

United Nations Economic Commission for Europe (2017). Guide on poverty measurement. United Nations New York.

United States Agency for International Development (2015). Gender and extreme poverty getting to zero: A usaid discussion series. United States Agency for International Development.

Watson, I. & Derrill, D. (2014). Poverty and basic needs. *Encyclopedia of Food and Agricultural Ethics. Springer Netherlands*, (pp. 1529–1535).

Wong, S. Y. (2012). Understanding poverty: Comparing basic needs approach and capability approach. *Available at SSRN 2066179*.

World Bank Group (2020). Namibia overview. https://www.worldbank.org/en/country/namibia/overvie

Zaninetti, L. (2013). The initial mass function modeled by a left truncated beta distribution. *The Astrophysical Journal*, 765(2), 128.