COMPARISON OF LINEAR AND QUANTILE REGRESSION METHODS FOR DETERMINATION OF FACTORS ASSOCIATED WITH NUTRITIONAL STATUS OF WOMEN OF REPRODUCTIVE AGE IN NIGERIA

BY

OYEDAPO, HEPHZIBAH ABIMBOLA

B.Sc. (Hons) Statistics (Ilorin)

MATRIC NUMBER: 182955

A DISSERTATION SUBMITTED TO

THE DEPARTMENT OF EPIDEMIOLOGY AND MEDICAL STATISTICS,

FACULTY OF PUBLIC HEALTH, COLLEGE OF MEDICINE,

UNIVERSITY OF IBADAN

IN PARTIAL FULFILLMENT OF

THE REQUIREMENT FOR THE AWARD OF

MASTER OF SCIENCE DEGREE IN BIOSTATISTICS

JANUARY, 2016

ABSTRACT

Mean regression approach explores the average effect effectively of the change in mean BMI, but may not be able to identify how extreme values affect Body Mass Index (BMI). Therefore, mean regression based methods may not be able to answer how the factors may affect large extreme BMI values. The use of quantile regression allowed the impact of the explanatory variable to vary along the whole range of BMI intake. Reproduction is associated with nutritional status because the various roles played by women give rise to serious health problems. While warnings about health penalty of excess weight (less attention seems to be given to the consequences of being underweight. The study aims to evaluate the performance of parametric Ordinary Least Square (OLS) and non parametric regression (quantile function) method to determine factors affecting nutritional status of women of reproductive age in Nigeria.

A nationally representative sample of women within reproductive age (15-49 years) within households in communities was obtained from the Nigeria Demographic and Health Survey (NDHS 2013). The Body Mass Index of women defined as weight/height^2 was the outcome variable while the explanatory variables include age, family size, total children ever born, marital status, highest level of education, place of residence region, wealth index. OLS regression was used to determine the average effect of factors on BMI while Quantile Regression was used to determine how a particular quantile of the BMI distribution was associated with covariates Analysis was done on STATA version 12 while graphs were done using R version 3.2.0.

A total of 31, 828 women were included in the study. The mean age of women was 29 years (SD=7.0), 49.2% had no formal education and 23.5% belonged to the poorest wealth quintile. It was shown that only 16% of variation has been predicted by linear regression. Results of ordinary least square regression analysis show that women's age, number of children, place of residence, level of education, some regions (North East, North West and South West), wealth index ($p<0.001$) were found to have effects on women's Body Mass Index. While in the 10th quantile, the effect of children ever born, place of residence were not significant. Family size did not contribute significantly to the BMI effect produced.

Quantile regression was able to detect the amount of under estimation and over estimation produced by the Ordinary Least Square regression of BMI values. The magnitude of the changes differed depending on the location of the woman in the BMI distribution

Key words: Body Mass Index, Linear regression, Quantile regression, Underweight, Overweight

Word Count: 410

ACKNOWLEDGEMENT

I am grateful to God for His sufficient grace and help.

My sincere appreciation goes to my supervisor, Prof. Olusola Ayeni for his immense contribution, encouragement, valuable advice and fatherly love. I thank him for his time.

My heartfelt thanks go to all my lecturers; Dr. (Mrs) O.B. Yusuf, Dr. J.O. Akinyemi, Dr. O.M. Akpa, Dr. S.A. Adebowale, Dr. S.A. Fagbamigbe, Dr. R.F. Afolabi and Dr. B.M. Gbadebo.

I am grateful to Mr. N.B. Afolabi for his resolute support.

To my course mates, I thank them for their acquaintance. Recognition goes to Kayode Fowobaje, Samson Olorunju, Mariam Jaiyeola, Mrs Akinpelu, Alhaja Amina, Mr. Inubile, Michael Ekholutale, Opeyemi Ojo, Mrs Odunlami, Anthony Wegbom, Olawumi Adewumi, and Ayoola Agbaje.

I also appreciate Tunde Bakre and David Ajayi for their efforts.

DEDICATION

To my family for their unwavering support

CERTIFICATION

I certify that this work titled "Comparison of Linear and Quantile Regression Methods for Determination of Factors Associated with Nutritional Status of Women of Reproductive Age in Nigeria" was carried out by Oyedapo, Hephzibah Abimbola in the Department of Epidemiology and Medical Statistics, Faculty of Public Health, University of Ibadan under my supervision.

....

Supervisor

Prof. Olusola Ayeni

B.Sc. (Ibadan), M.Sc. (London), Ph.D. (London)

Department of Epidemiology and Medical Statistics,

Faculty of Public Health,

University of Ibadan, Nigeria.

CERTIFICATION

I certify that this work titled "Comparison of Linear and Quantile Regression Methods for Determination of Factors Associated with Nutritional Status of Women of Reproductive Age in Nigeria" was carried out by Oyedapo, Hephzibah Abimbola in the Department of Epidemiology and Medical Statistics, Faculty of Public Health, University of Ibadan under my supervision.

....

Supervisor

Prof. Olusola Ayeni

B.Sc. (Ibadan), M.Sc. (London), Ph.D. (London)

Department of Epidemiology and Medical Statistics,

Faculty of Public Health,

University of Ibadan, Nigeria.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

OLS	Ordinary Least Square
QR	Quantile Regression
BMI	Body Mass Index
NDHS	Nigeria Demographic and Health Survey
WHO	World Health Organization
CED	Chronic Energy Deficiency
SE	Standard Error
LRM	Linear Regression Model
QRM	Quantile Regression Model
MAC	Mid Arm Circumference
MAMC	Mid Arm Muscle Circumference
BIA	Bioelectrical Impedance Analysis
DEXA	Dual-Energy X-ray Absorptiometry
GLM	Generalized Linear Models
GAMLSS	Generalized Additive Models for Location, Scale and Shape
LMS	Lambda-Mu-Sigma
EA	Enumeration Area

CHAPTER ONE

1.0 INTRODUCTION

1.1 BACKGROUND

Ordinary Least Square (OLS) Regression model focuses on modeling the conditional mean of a response variable without accounting for the full conditional distributional properties of the response variable. Hence, an OLS regression analysis does not give a complete picture of the relationship between variables, as it centers on changes at the conditional mean. Linear regression is the one of the popular analytic tools used to detect how a continuous dependent variable is associated with a set of predictors in health and population studies (Kleinbaum et al., 2007). Linear regression centers on the expectation of a variable *Y* conditional on the values of a set of variables X, (that is, E(Y|X)), which can be more or less complex, but it limits solely on a specific location of the Y (dependent variable) conditional distribution.

The linear regression approach basically fails to characterize the relationship between a dependent variable's distribution and predictors, and also, does not answer the question of how changes in predictors have an effect on the shape of a dependent variable's distributions (Koenker, 2005). Linear regression requires the residuals to be normally distributed and homoscedastic (Ho et al., 2006; Adams, 2008) but the normality assumption may perhaps not hold especially when a dependent variable has a heavy-tailed distribution (Hao and Naiman, 2007). Failing to meet these modeling assumptions may result in biased estimates and misleading conclusions. Quantile regression was an approach recently developed to deal with these issues.

Quantile regression (QR) was introduced in 1978 by Koenker and Bassett (Hao and Naiman, 2007), which models conditional quantiles as functions of predictors. The quantile regression model is an extension of the linear-regression model. Fifteen years after Koenker and Bassett first introduced quantile regression, empirical applications of quantile regression started to develop. QR methods have been popularized among economists and ecologists over the past decade.

Quantile regression is one of the statistical analyses used in medicine and public health (Austin et al., 2005) which is capable of identifying more effects than the conventional OLS methods. It does not restrict attention to the conditional mean, permits estimation of the whole conditional distribution of a response variable, and approximates quantile exact effects which explain the impact of the covariates not only on the middle but also on the extremes of the distribution of the outcome variables (Mohamad et al., 2009). QR also broadens this approach by studying the conditional distribution of dependent variable(Y) on independent variable(X) at different locations and thus offering a global view on the interrelations between Y and X and also describes the relationship at different positions in the conditional distribution of dependent variable. For instance, in the analysis of the correlates of nutritional status, quantile regression evaluates and provides a complete view of change in nutritional status with factors when the distribution is not normal (Shankar, 2010).

In quantile regression, a real valued random variable Y of the distribution function $F(y) = Pr(Y \leq y)$ can be defined as $Q(\tau) = \inf\{y: F(y) \geq \tau\}$. The median is a special quantile, one that describes the central location of a distribution. Conditional median regression is a special case of quantile regression in which the conditional 0.5th quantile is modeled as a function of covariates. Non central positions of the dependent variable can be described by other quantiles. The "quantile" concept takes a broad view of specific terms like quartile, quintile, decile, and percentile (Hao and Naiman, 2007). QR partitions the whole distribution into quantiles and estimates the conditional quantiles as functions of explanatory variables Equal-sized partitioning of a distribution are called quantiles. For example, quintile partitions the data into five equally-sized groups.

The Quantile Regression Model (QRM) and Linear Regression Model (LRM) are similar in certain aspects, as both models deal with a continuous response variable that is linear in unknown parameters, but they model different quantities and rely on different assumptions about error terms. However, limiting attention to the mean and standard deviation alone leads us to ignore other important properties that offer more insight into the distribution. In addition, one may be interested in other position parameter instead of the mean (Cizek, 2011). Thus, quantile regression is more preferable in such situations.

The advantages of quantile regression over traditional approaches (OLS) are: flexibility in dealing with non-normally distributed errors, robustness against outliers, and ability to detect heterogeneity (Koenker 2005). Quantile regression has this capability to analyse the whole

Gannoun et al, 2003; Hao and Naiman, 2007). OLS alone could not help since the effect of skewness can be assessed only through the evaluation of different quantiles. By focusing on the mean as a measure of central location, information about the tails of the distribution is lost (Montenegro, 2001).

Nutrition is the intake of food, considered in relation to the body's dietary needs. Good quality nutrition is a basis of good health which is an adequate, well balanced diet. Poor nutrition can lead to reduced immunity, increased susceptibility to disease, impaired physical and mental development, and less productivity (WHO, 2013).

Nutritional status of women is essential for the health of women and their ability to perform work as well as the wellbeing of their children. However, poor nutritional status in a woman signifies a low Body Mass Index (BMI), short stature, anaemia, a greater risk of obstructed labour and having a low birth weight baby (NDHS, 2013). About 30% of all women of reproductive age are presumed to have anaemia whereby the highest proportion is from the African women-48% to 57% (De Benoist et al., 2009). Anaemia increases the risk of haemorrhage and sepsis during childbirth. Women are most vulnerable to anaemia due to deficient iron in their diets, menstrual blood loss and periods of speedy growth (WHO, 2013). One in every three Nigerian woman suffers from iron deficiency, 49% of women of reproductive age have anaemia, 24.3% have low iron levels, while 12.7% are iron deficient (Nutrition Society of Nigeria, 2015).

Reproduction is associated with nutritional status, whereby inhibitors in the hind brain suppress ovulation with weight loss (ESHRE Capri Workshop Group, 2006). Reproduction has been identified as a possible cause of under-nutrition among women in developing societies known as "maternal depletion" (Winkvist et al., 1998).

Nutritional status of women is measured by medical and social history, diet history and intake, clinical examination, anthropometrics and biochemical data. Anthropometric methods are assessed by Body Mass Index(evaluation of body weight independent of height), Frame Size(determined using wrist circumference and elbow breadth), Skin Fold Thickness(estimates subcutaneous fat stores to estimate total body fat), Body Circumferences and Areas(estimates skeletal muscle mass-mid arm or upper arm circumference(MAC), Mid Arm Muscle or Arm Muscle circumference(MAMC), Total Upper Arm Area, Upper Arm Muscle Mass and Upper

Arm Fat Area, Bioelectrical Impedance Analysis(BIA measures electrical conductivity through water in difference body compartments) and Dual-Energy X-ray Absorptiometry(DEXA measures whole body scan with x-rays of different intensity).

BMI is an attractive index for measuring women's undernutrition and overnutrition because it is an indicator of body composition. A woman, by her structure has reduced mass and more fat mass. At her child bearing age, she tends to build up more fat due to child-birth. BMI is related to body fat mass, fat-free mass and fat proportion which makes it a valid indicator for women because women have a larger proportion of fat than men (WHO, 1998). Low BMI values indicate reduced fat and fat free mass while the upper end of the BMI distribution signifies that the body fat mass is strong which makes BMI a suitable index for measuring obesity.

The nutritional and health status of women is of great concern in the modern world, because the various roles played by women give rise to serious health problems. The situation is even worse in countries where societal norms and sex discrimination have forcefully subjected women to satisfy the health and nutritional needs of their families at their own expense.

Hunger and malnutrition are devastating problems, mainly for the poor and unprivileged. Women in the reproductive age group and children are most vulnerable to malnutrition due to low dietary intakes, inequitable sharing of food within the household, improper food storage and preparation, dietary taboos, infectious diseases, and care. Mostly for women, the high nutritional costs of pregnancy and lactation also contribute significantly to their poor nutritional status.

Some evidence in developing countries show that malnourished individuals, that is, women with a body mass index (BMI) below 18.5, show an advancement in mortality rates as well as increased risk of illness (Rotimi, 1999). For social and biological reasons, women of the reproductive age are among the most vulnerable to malnutrition (Ene-Obong et al., 2001). Increased perinatal and neonatal mortality, a higher risk of low birth weight babies, stillbirths, and miscarriage are some of the consequences of malnutrition in women. Women below 1.45 meters in height are considered too short or stunted (NDHS, 2013). Unfortunately, problems of under as well as over-nutrition prevail in female populations worldwide (Winkvist, 2003).

Obesity is a global health problem and the rate of increase is high (Kumanyika et al., 2002). Overweight is one of the top of ten health risks in the world. A report showed that more than one billion people are overweight worldwide, whereby 250 million are estimated to be clinically obese (WHO, 1998). A study carried out in Rivers state, Nigeria reported a prevalence of obesity as 16.3% (Siminnialayi et al., 2008).

At the same time, in some parts of the world, increasing prevalence of obesity (BMI >=30) have been documented especially among women. For instance, obesity is at slightly three times more prevalent among women compared to men in several African nations (WHO, 2000).

Various factors such as eating habit, educational background, socioeconomic status (Dressler and Smith, 2013; Ball and Crawford, 2005), working condition and cultural features of individuals are observed to have an effect on etiology of obesity. Under-nutrition is a menace for low productivity, poor health and mortality (Benson and Shekar, 2006). Moreover, under-nutrition among women leads to poor reproductive health outcomes (Mora and Nestel, 2000).

Finally, health professionals have warned about the adverse health outcomes of overweight and obesity (Tikimoto et al., 2004). While warnings about health penalty of excess weight (Berg, 1995; Must et al., 1999), less attention seems to be given to the consequences of being underweight (Che, 2002; Pawlinska-Chmara, et al., 2007).

1.2 PROBLEM STATEMENT

Both underweight and obesity are forms of under-nutrition and over-nutrition; they both have greater health consequences. Few studies carried out have focused on under-nutrition (Mokhtar et al., 2001). However, with respect to human development, both under-nutrition and over-nutrition demand as high degree of main concern as they signify the lack of food security and also obesity which indicates over consumption of food (Letamo and Navaneetham, 2014).

Under-nutrition and poor health from preventable causes greatly have an effect on the well-being of millions of people in the developing countries (Kiday et al., 2013). One of the goals of the National Plan of Action on Food and Nutrition in Nigeria to address the food and nutrition problems in the country is to reduce under-nutrition especially among women and children by 30% by 2010 (NPAN, 2005) which has not been accomplished. In recent years, the prevalence of overweight and obesity has increased obviously in women (Mokhtar et al., 2001). Fear of being fat may bring unnecessary attempts to reduce body weight producing thinness that in some cases is associated with nutritional deficiencies, irregular menstruation, and eating disorders.

Previous studies suggest that underweight in women of childbearing age is an aetiology for adverse pregnancy outcomes and intrauterine growth retardation or low-birth weight infants (Nandi, 2000). Low BMI can be a sign of Chronic Energy Deficiency (CED) and lack of adequate weight gain during pregnancy can lead to low birth weight babies leading to adverse health implications (Singh, 2011). Women with CED have increased morbidity while other studies have linked low BMI to decreased work capacity (Dharmalingam et al., 2005). Also among individuals who are HIV-positive, those who have lower BMI may progress to AIDS more quickly (Nube et al., 2003).

Obesity is a critical public health problem for women of reproductive age. Obesity was formerly considered as a problem linked to affluence but now the trend is increasing fast in many developing countries and even in the poor neighborhoods of the developed countries in the world. Women who are obese and overweight are at high risks of adverse health outcomes like cardiovascular disease, diabetes, kidney disease and cancers related to obesity (Flegal et al., 2007).

The techniques of parameter estimation have led to several problems in the realm of regression analysis such as underestimation and overestimation of parameters and some are imprecise (Green et al, 1994). These differences in the distribution make the use of quantile regression more compelling than OLS, which has the shortcoming of assuming the same effects across the entire distribution. Previous studies on nutritional status have either used an Ordinary Least Squares (OLS) framework or have focused on specific groups (Kamal et al., 2005).

Linear regression requires the residuals to be normally distributed and homoscedastic (Ho et al., 2006; Adams, 2008) but the normality assumption may not hold particularly when a dependent variable has a heavy-tailed distribution (Hao and Naiman, 2007). Failure to meet these modeling assumptions may result in biased estimates and confusing conclusions. To address these issues, quantile regression method has been developed.

Most researches done on the nutritional status of women have emphasized on mean regression which is limited in its ability to capture cross-distribution variations in effects.

1.3 JUSTIFICATION

Women of child-bearing age have certain nutritional requirements above those of adult males. The loss of blood during menstruation results in a regular deficiency of iron and other nutrients and makes women more vulnerable than men to anaemia. However, in many developing countries, women work much harder than men. In rural areas, they are often heavily involved in agriculture, and in urban areas they may work long hours in factories and elsewhere, yet when they return home from the field or the factory they still have much work to do in the household, including food preparation and child care. Often, the heavy burden of collecting water and fuel falls on women. All of this labour increases women's needs for nutritional energy and other nutrients.

The nutritional status of women before, during and after pregnancy contributes a good deal to their own general well-being, but also to that of their children and other members of the family. The field of maternal nutrition focuses attention on females as mothers. It has often concentrated on their nutritional status mainly as it is related to the well-being of the infants that they produce and their ability to breastfeed and nurture. The health and well-being of the mother herself has been neglected. In the same way, the field of maternal and child health has placed major importance on providing services and help to women mainly so that they can have successful pregnancies and lactations; this is also in the interests of the infant. The dual role of women as mothers and productive workers is compromised by poor diets and ill health; not only their own well-being but that of the whole family is affected. A heavy work load may push a woman with marginal food intake over the brink and into a state of malnutrition.

Researchers have encouraged healthy weight status among women of reproductive age (Flegal et al., 2010; Ashton et al., 2009) but little is known about the determinants of the progression from normal weight status to underweight and normal weight status to both overweight and obesity in Nigeria.

A researcher would be unable to assess the differences that occur in the relationship between the dependent and explanatory variables without performing quantile regression at different quantiles. Linear regression is widely used to assess the associated factors of the nutritional status of women of reproductive age in Nigeria, but in most cases of interest, the biometrical measurements like BMI are not normally distributed(Ouyang, 2015; Shankar, 2010), so parametric model such as linear regression that needs normality assumption are not informative enough (Mohamad et al, 2009). Therefore a non- parametric approach (QR) was also used.

Some analyses on nutritional status of women have used mean regression yet modeling using quantile regression is more appropriate than using mean regression in that the former provides flexibility to analyze the determinants of nutritional status corresponding to quantiles of interest whereas the latter allows only analyzing the determinants of mean nutritional status.

But I go beyond the previous studies and indeed much of the previous regression based literature on the determinants of nutritional status of women by specifying and estimating a richer type of regression model that overcomes the limited scope of linear regression typically used. This quantile regression (QR) approach allows the effects of the covariates on BMI to vary across the conditional distribution of BMI ranging from the relatively underweight to the relatively obese.

1.4 OBJECTIVES

1.4.1 Main Objective: To evaluate the performance of parametric (OLS) and non parametric regression (quantile function) method to determine factors affecting nutritional status of women of reproductive age in Nigeria.

1.4.2 Specific Objectives

The specific objectives of this study are to:

i. Compare the suitability of OLS and quantile regression models in assessing factors influencing the nutritional status of women of reproductive age.
ii. Assess the nutritional status of women of reproductive age in Nigeria.
iii. Determine factors influencing the nutritional status of women of reproductive age in Nigeria.
iv. Examine different directions of the effects of nutritional status on different quintiles of the distribution.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Anthropometric Measurements

There are mainly two ways to measure the incidence of malnutrition among vulnerable groups of the society. (i) Caloric/nutrition intake (ii) Anthropometric (Svedberg, 2001). For many years it has been agreed that anthropometric approach is a better measurement than calorie approach.

This study focuses on anthropometric approach which is considered as more consistent measurement over calorie intake approach. Nutritionists argue that the energy intake is a poor measure of evaluating nutritional status, which depends not only on the nutrient intake but also on non-nutrient food attributes, privately, and publicly provided inputs and health status (Svedberg, 2001). It is recommended that the estimation of malnutrition should be based on outcome measures rather than input measures.

The suggested outcome measure is the anthropometric measure. Outcome indicators are more closely related to health and functional capacity. One essential indicator of measuring under nutrition is the BMI (Jiten et al., 2012). It was reported by Siminnialayi et al., 2008 that Body Mass Index (BMI) is the most commonly marker used to assess body-weight and it is also highly correlated with body fat. The medical hazard of being obese is related with body fat distribution, therefore, abdominal fat is considered as important a medical risk as the overall fat in the body (Finer, 2003), thus, the Body Mass Index (BMI) serves as a platform for assessing obese, overweight and also underweight women. Igiri et al., 2009, in a study conducted in Cross River State, Nigeria, used BMI to ascertain the health status of young adults in the Calabar metropolis and concluded that they had a normal health status.

BMI as an indicator offers a direct measure of underweight and overweight, it also acts as a proxy for mortality risk and fat mass and has an improved overall performance than any other weight stature index (Flegal et al., 2010).

Among the outcome measures, anthropometric measures are considered to have an advantage over other indicators since body measurements are sensitive to even minor levels of malnutrition. Information about nutrition is often collected with the clear aim of selecting people for targeted intervention. Then it needs to know who is undernourished and who is not.

and standardized dietary intake norms cannot be used to detect under-nutrition in individuals. For this purpose, anthropometric and related methods have to be applied. Nevertheless, several development economists and contemporary nutritionists (Svedberg, 2000) have shown that anthropometric measures provide more reliable and useful indications of nutritional status than do dietary intake measurements.

The fundamental advantages of the anthropometric approach are simplicity and accuracy. Based on the above arguments, the anthropometric approach is preferred to calorie intake approach because it reveals the past nutritional status in terms of Body Mass Index. In general, it is agreed that anthropometric technique offers more consistent estimates.

2.2 Morbidity and Mortality Due to Malnutrition

Malnutrition is an important public health problem and a threat to proper living which can be in two forms-under nutrition and over nutrition. Underweight, obesity and even overweight are considered as unhealthy and abnormal status which can be traced to several mortalities and morbidities (Flegal et al., 2007). A local Nigerian newspaper described obesity as "a new killer in town" which also joins the ranks of HIV/AIDS and malaria; this was reported by Ogundipe and Obinna, 2010.

Ezzati et al., 2002 evaluated diverse causes of mortality and morbidity which revealed that malnutrition is the single leading cause of ill health worldwide. The poor are exposed to greater risk of health problem which leads to a heightened risk of disease (WHO, 2013). Agetta, 2010 reported that malnutrition increases the risk of death. Poor nutritional status of women of reproductive age is linked to women's susceptibility to ill health. Nutritional status of women is targeted by the millennium development goal 5 to reduce maternal mortality by three-quarters by 2015. Akpa and Mato, 2008 in their study revealed that those underweight do not have enough nutritional reserve to carry them through during illness which makes them have a higher risk of mortality. Therefore, females in their reproductive ages are more vulnerable to communicable diseases. In developing countries, maternal mortality rate is about a hundred fold higher.

2.3 Factors Affecting the Nutritional Status of Women

Understanding the factors that affect women's nutritional status is important. The factors are:

2.3.1 Wealth Index

According to Islam et al., 2004, malnutrition is a common trait among low income rural women. Malnutrition rates are higher among rural households who rely more on agriculture than on other sectors for their livelihoods. The results show that generally, across rural and urban areas, household economic status is positively associated with women nutrition, though the effect is limited. Its impact is found in rural North East and the Southern region (Ajieroh, 2010), and in urban areas of the North East and North West. Although relatively weak in effect, the consistency and wide reach of the effect of household economic status indicates its relative importance to maternal nutrition.

Letamo and Navaneetham, 2014 in their study revealed that women who were in the poorest household were more likely to be undernourished compared with those who were in the richest class. Adedoyin et al., 2005 reported that low socio economic status is responsible for overweight and obesity in that women have poor food habits. They estimated that of the obese women, 15.3% are of low socio economic status while 3.75% are of high socio economic status.

2.3.1.1 Women's Employment and Control over Income

Employment status is also a factor that affects body weights of women. The more earnings women control, the better their wellbeing and nutrition (Ene-Obong et al., 2001). In a study conducted, it was found that employed women had a significantly lower body mass index and healthier eating habits compared to those employed (Arslan and Ceviz, 2007). It is also known that obesity possibility is higher for women with sedentary occupation (Mummery et al., 2005).

2.3.2 Educational Level

Education is related to illiteracy and lack of access to information. It was shown in some studies that educational backgrounds of women affect their habits positively (Sharma et al., 2008; Carlsson et al., 2013). Also, it is plainly seen in this study that BMI of individuals considerably decreases with increase in women's level of education. It can be said that education influences the health and nutritional status of rural women in Nigeria (Ene-Obong et al., 2001).

Low BMI was more prevalent among adults with no education compared to those with higher levels of educational attainment. Women who had no education were more likely to be underweight compared with those who had post secondary education (Letamo and Navaneetham, 2014). It was also reported by Adedoyin et al., 2005 that women with low level of education have poor eating habits, believing that quantity is better than quality.

2.3.3 Place of Residence

Prior studies in developing countries suggest that women living in urban areas are more likely to be overweight and obese than those in rural areas (Popkin, 2001). Uthman, 2009 revealed in his research work conducted in Nigeria that women who live in urban areas are at risk of being obese. Adult population living in rural areas portrayed a high prevalence of low BMI than those residing in urban areas (Letamo and Navaneetham, 2014).

2.3.4 Age of Women

Chinedu and Emiloju, 2014 demonstrated that persons aged 20-35 are generally at the peak of their strength. The biological functions of the body are not subject to weakening. A research carried out by Teller and Yimar, 2000 showed that women aged 15-19 years and women aged 45-49 years are mostly affected by under-nutrition. Letamo and Navaneetham, 2014 also demonstrated that younger women age 20-24 years were more likely to be underweight compared to older women. He concluded that young age was associated with low BMI.

In India, women in the age group 15-19 years and 20-24 years are 3.1 and 1.7 times respectively more prone to suffer from under-nutrition compared to women in the age group 35-49 (Jiten, 2012). Another study conducted showed that women's nutritional status has significant association on age of women (Kiday et al., 2013).

2.3.5 Marital Status of Women

Ball and Crawford, 2005 demonstrated that marriage, divorce or loss of spouse which change the societal roles have an effect on women's body weights. It is found in their study that single individuals have a much lower BMI compared to married or widow individuals ($p<0.001$). The proportion of women who affirmed that their bodyweight increased after marriage was 31.5%. Also, according to the study carried out by Sobal et al., 2003; in the ten years of follow-up, weight gain was more usually observed in the women getting married. Low BMI was more prevalent among single women (Letamo and Navaneetham, 2014).

2.3.6 Region

Studies have shown a high rate of malnutrition in the rural part of Nigeria; 56% reported in a rural community of South West and also in the Northern part of the country, 86.3% was reported in three rural areas (Okwu et al., 2008; Oluwatayo et al., 2008). At the same time, Body Mass Index (BMI) and obesity have increased among women in most regions (Finucane et al., 2011; Stevens et al., 2012).

2.3.7 Family Size

Agetta, 2010 in her study revealed that households that had fewer members were more likely to have good nutritional status. Nutritional status of study participants according to BMI classification had no significant association on family size (Kiday et al., 2013)

2.3.8 Total Children Ever Born

Low parity was associated with low BMI. Women of high parity were less likely to experience low BMI (Letamo and Navaneetham, 2014). Kiday et al., 2013 also demonstrated that nutritional status of study participants according to BMI classification had no significant association on number of parity (Kiday et al., 2013). In a study conducted in Gombe, North East Nigeria, Yahaya et al., 2007 found out that a woman with a higher parity has a more difficult chance to obtain optimum nutritional requirements.

2.4 Issues on Nutritional Status of Women in Sub- Saharan Africa

Ahmed et al., 2014 found out that the relationships between nutritional status of mothers and factors were not significant. The findings contradict the outcome of a study in Ethiopia, Burkina Faso, and Republic of Congo which showed that, wealth index is one of the most important determinants of nutritional status in women (Girma and Genebo, 2002; Adebowale et al., 2015).

At the same time, Body Mass Index (BMI) and obesity have increased among women in most regions (Finucane et al., 2011; Stevens et al., 2012). Nutritional status of study participants according to BMI classification had no significant association on maternal educational status, household asset (radio) ownership, family size, maternal age, number of meals per day, number of parity, farm animal ownership and residence of the lactating mothers (Kiday et al., 2013).

This was in contrary to the study conducted on women's nutritional status which had significant association on marital status, household assets, age of women and maternal education (Kiday et al., 2013).

2.5 Robustness in Linear and Quantile Regression Analysis

Robustness refers to insensitivity to outliers and the violation of model assumptions concerning the dependent variable. The Linear Regression Model (LRM) estimates can be sensitive to outliers while Quantile Regression Model (QRM) estimates are not sensitive to outliers (Lingxin, 2007). Robustness is studying phenomena of highly skewed distributions e.g. health outcomes like obesity. Robustness in linear and quantile regression models for both categorical and continuous dependent variables has been studied by several authors (Zaman et al., 2001). The coefficients for the median and other quantiles remain the same even when an extreme value is added to the data (Onyedikachi, 2015)

2.6 Empirical Applications of Quantile Regression

Quantile regression has been applied to a broad range of studies. Quantile regression also spread to medicine and public health (Austin et al., 2005). Alexander et al., 2011 demonstrated that studying the coefficients and their uncertainty for different percentiles generates new insights, especially for non-normally distributed data. A small number of recent applications to explaining variations in BMI (Beyerlein et al., 2008; Terry et al., 2007; Sturm et al., 2005) have demonstrated the value added by QR methods in this setting. Abreveya, 2002 in his study "Demographics impact on infant birth weight" showed that the quantile regression estimates signify that several factors have higher impact at lower quantiles while lower impact at higher quantiles.

The vast majority of applications of quantile regression in the field of health have focused on geographically defined contexts such as countries (Shankar, 2010), states, and neighborhood (Hill et al., 2014; Hoss and Fischbeck, 2014, Ramokolo et al., 2015).

Empirical researchers took advantage of quantile regression's ability to examine the impact of predictor variables on the response distribution. Two of the earliest empirical papers by economists (Hao and Naiman, 2007) provided practical examples of how to apply quantile regression to the study of wages. Quantile regression allowed them to examine the entire conditional distribution of wages and determine if the returns to schooling, and experience and the effects of union membership differed across wage quantiles. The use of quantile regression to analyze wages increased and expanded to address additional topics such as changes in wage distribution (Machado and Mata, 2005; Melly, 2005), wage distributions within specific industries (Budd and McCall, 2001), and educational attainment and wage inequality (Lemieux, 2006).

The use of quantile regression also expanded to address the quality of schooling (Bedi and Edwards, 2002; Eide et al. 2002) which showed a strong positive effect on the labor earnings distribution with the strongest impact occurring at the lower quartile. Quantile regression also spread to other fields, notably sociology, ecology and environmental sciences.

2.7 Determinants of Nutritional Status of Women of Reproductive Age

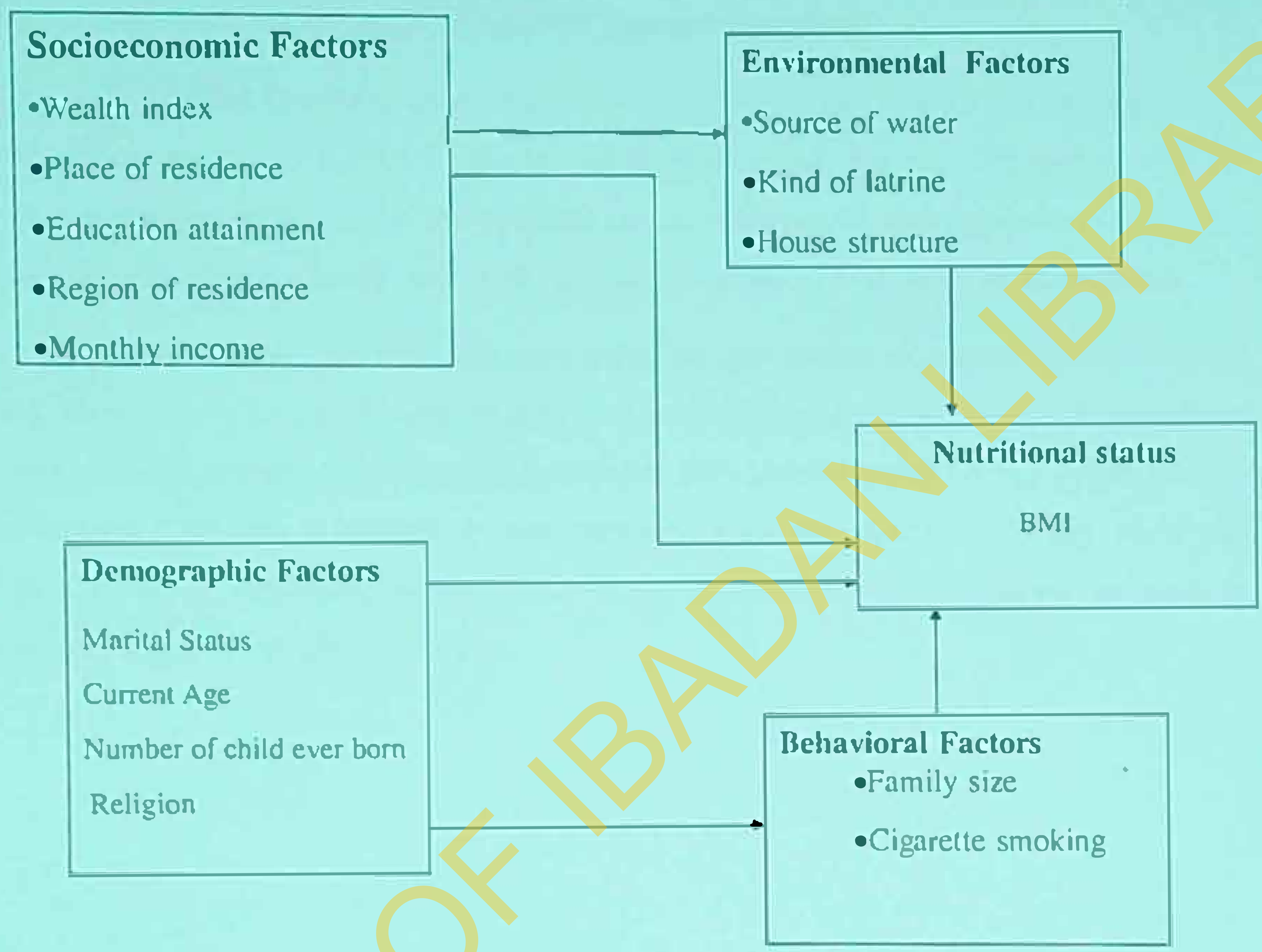


Figure 2.1: Conceptual Framework for Determinants of Nutritional Status of Women of Reproductive Age

Some researchers have examined the determinants of nutritional status of women of reproductive age (Nemati and Mogadam, 2008; Agetta, 2010). The customized conceptual framework builds on existing knowledge to analyze the socio-economic and demographic factors associated to nutritional status among reproductive aged women in Nigeria. United Nations (UNICEF, 2005) and Adebowale et al., 2014 reported that in countries where nutrition improvements have lagged behind, economic growth was slow and social discrimination against women was common. The environmental factors, religion and cigarette smoking were not used in this study.

2.8 Comparison of Linear and Quantile Regression Methods of Estimation.

Ouyang et al, 2015 used the lambda-mu-sigma (LMS) method to describe the changes in the BMI distribution of adults. He also investigated the changes in the BMI distribution over time using a separate sex stratified longitudinal quantile regression analysis.

Asirvatham, 2009 studied the differences between eating habits and BMI. He compared the quantile regression results to the OLS results. He demonstrated that to fully understand the behavior of the relationships across the conditional distribution of the dependent variables: energy intake, Healthy Eating Index, and BMI, quantile regression will be more appropriate.

Beyerlein et al., 2008 in their research work used different approaches to predict childhood BMI which were Generalized Linear Models (GLM), Quantile Regression (QR), and Generalized Additive Models for Location, Scale and Shape (GAMLSS). He compared GLM, GAMLSS and QR models among BMI data to identify the best method for the risk factors of obesity. He found out that QR allowed for additional interpretation of pre specified distribution quantiles, such as quantiles referring to overweight or obesity.

CHAPTER THREE

3.0 METHODOLOGY

3.1 Study Area

Nigeria is one of the Sub-Saharan African countries located in the West African region which came into existence as a nation in 1914 through the amalgamation of the Northern and Southern protectorates. The British established a crown settlement type of government after the amalgamation. The dealings of the colonial administration were carried out by the British until 1942, when a few Nigerians became involved. In the early 1950s, Nigeria achieved partial self government with a parliament in which the majority of the members were elected into an executive council of which most were Nigerians. Nigerians became fully independent in October 1960 as an alliance of three regions (Northern, Western, and Eastern) under a constitution that provided for a parliamentary system of governance.

Nigeria is in the West African sub region, lying between latitudes 4°16´ and 13°53´ north and longitudes 2°40´ and 14°41´ east. It is surrounded by Niger in the North, Chad in the North east, Cameroon in the East, and Benin in the west. To the South, Nigeria is bordered by approximately 850 kilometres of the Atlantic Ocean, stretching from Badagry in the West to the Rio del Rey in the East. It covers a total area of 923,768 kilometer square and it is the 14th largest country in Africa and the world's 32nd largest country. Nigeria comprises of 36 states and a Federal Capital Territory(FCT), these states are grouped into six geopolitical zones; North Central, North East, North West, South East, South-South And South West. There are 774 local government areas in the country. According to the 2006 National Census estimate, the population of the country was 140,431,790 with an annual growth rate of 3.2 percent. Nigeria is the most populous country in Africa; it has an estimate of 374 ethnic groups. The major ethnic groups are Hausa/Fulani, Yoruba and Igbo which account for 68 percent, Edo, Ijaw, Ibibio, Kanuri, Ebira, Nupe And Tiv make up 27%, other minority ethnic groups make up 5%.

3.2 Study Design

Data was obtained from the Nigeria Demographic and Health Survey (NDHS 2013) and secondary data analysis was conducted to answer the study objectives. The survey made use of a cross-sectional population based study design. This study explores the factors affecting nutritional status of women of reproductive age.

3.3 Study Population

The population for the 2013 Nigeria Demographic and Health Survey was drawn from females aged 15-49 years and males aged 15-59 years in Nigeria. Samples were derived from the target population by random selection of households in the country, the selected individuals were interviewed. For the purpose of the study, women within the reproductive age (15-49years) in Nigeria were the target population. The study population was obtained as subsamples from the samples of women interviewed in the survey of the year 2013.

3.4 Inclusion Criteria

Reproductive aged women between 15-49 years.

3.5 Exclusion Criteria

1. Women who are pregnant.
2. Women who had given birth in the two months preceding the survey will be excluded from the analysis.

3.6 Sampling Frame and Technique

NDHS 2013 is a nationally representative sample. The 36 states in Nigeria and the Federal Capital Territory were divided into Local Government Areas (LGA) and each LGA was further divided into smaller localities. The 36 states were regrouped by geopolitical locality into six zones and using the 2006 population census implementation, each locality was subdivided into Enumeration Areas (EAs). A complete list of the EAs served as the sample frame for the survey. The sampling technique for the 2013 NDHS was a stratified sample, selected at random in three stages from the sampling frame. The first stage; each state was stratified into urban and rural areas: this resulted in a list of localities. Second stage; one enumeration area was selected at random from a selected locality with equal probability of selection, the resulting list of households served as sampling frame for the selection of households in the third stage. The third stage: fixed number of 45 households were selected in every urban and rural cluster through probability systematic sampling using the household listing.

3.7 Sample Size

A representative sample of 40,320 households was selected for the NDHS 2013 survey. Of the households occupied, 38,522 were successfully interviewed of which 39,902 women aged 15-49 years and 18922 males aged 15-49 years were eligible for interview however 38,948 females and 17,359 males were successfilly interviewed. For this study, a sample size of 31,482 women within the reproductive age of 15-49 years was used but after the sample weighting, it increased to 31,828.

3.8 Study Variables

The following variables were used in the study;

3.8.1 Dependent Variable

The Body Mass Index (BMI) of women which is a continuous variable was specified as the dependent variable. The nutritional status of women was assessed by anthropometric method-weight and height, using the Body Mass Index (BMI), which is defined as a woman's weight in kilograms divided by the square of her height in meters (kg/m^2). Underweight, overweight and obesity were defined using the WHO BMI classification as follows: underweight = < 18.5 kg/m^2, normal = >18.5-24.9kg/m^2, overweight = >25-29.9 kg/m^2 and obese = > 30kg/m^2 (WHO, 2000).

3.8.2 Independent Variables

The choice of the explanatory variables was based on literature of factors influencing women nutritional status of reproductive age (Van et al, 2008, Agee, 2010; Owen et al, 2014). It consists of Current Age, Family Size, Total Children Ever Born, Marital Status, Level of Education, Place of Residence, Region and Wealth Index. The explanatory variables that are categorical were dummy coded to examine differences across variables given in Table 3.1 below.

Table 3.1: Variables Definition

Variables	Subcategories	Definitions/Dummy code
Body Mass Index		BMI in kg/m^2
Age	15-24 years	=1 if 15-24 years, else=0
	25-34 years	=1 if 25-34 years, else=0
	35-49 years	=1 if 35-49 years, else=0
Family Size	<5 members	=1 if <5 members, else=0
	5-8 members	=1 if 5-8 members, else=0
	>8 members	=1 if >8 members, else=0
Total Children Ever born	1-3 children	=1 if 1-3 children, else=0
	4-6 children	=1 if 4-6 children, else=0
	>=7 children	=1 if >=7 children, else=0
Marital Status	Married	=1 if Married, Unmarried=0
	Unmarried	=1 if Unmarried, Married=0
Highest Level Of Education	No education	=1 if No Education, else=0
	Primary	=1 if Primary, else=0
	Post Primary	=1 if Post Primary, else=0
Place of Residence	Urban	=1 if Urban, Rural=0
	Rural	=1 if Rural, Urban=0
Region	North Central	=1 if North Central, else=0
	North East	=1 if North East, else=0
	North West	=1 if North West, else=0
	South East	=1 if South East, else=0
	South South	=1 if South South, else=0
	South West	=1 if South West, else=0
Wealth Index	Poorest	=1 if poorest, else=0
	Poorer	=1 if poorer, else=0
	Middle	=1 if middle, else=0
	Richer	=1 if richer, else=0
	Richest	=1 if richest, else=0

3.9 Data Management and Analysis

SPSS version 20 was used for extraction of relevant data from the NDHS 2013 dataset and was also used for data cleaning. Descriptive statistics were analysed where the variables in classes were presented as frequencies (percentages) and continuous data were presented as mean (standard deviation) for normally distributed data while median and range were presented for skewed data.

STATA version 12 was used for fitting the two regression models (quantile regression and multiple linear regression). For quantile regression, the relationship between Body Mass Index and the explanatory variables were analyzed in specified conditional quintile (first, second, third, fourth and fifth) of the outcome variable. As is in literature, the selected quintiles are: 0.10 (the lowest conditional quantile function estimated), 0.25, 0.50 (median), 0.75 and 0.90 (the highest conditional quantile function estimated (Shankar, 2010; Asirvatham, 2009). Quantile regression method was used to examine the effects of covariates at these different points in the distribution. Also the graphs were plotted using R Console. Adjusted R-square was employed to assess the goodness of fit in linear regression while pseudo R^2 was used to assess quantile regression. Two hundred bootstrap replications were performed in the quantile equations to compute the standard errors (SE) of the estimates and also to obtain unbiased estimates (Asirvatham, 2009). The 95% confidence intervals were derived from standard errors generated from 200 bootstrap replications.

3.10 Statistical Models, Assumptions and Specifications

The statistical models used in the study include the following:

3.10.1 Linear Regression

A linear regression is a statistical method used for analyzing dataset with one or more independent variable(s) determining an outcome. The outcome variable is continuous. In this study, the linear regression model was fitted; the dependent variable BMI measured in weight/height^2 and independent variables such as age of women, region, residence, e.t.c as mentioned earlier. OLS is the typical method for regular linear regression.

The linear regression function which is:

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\ldots\ldots + \beta_pX_p+ \varepsilon_i \qquad \varepsilon_i \sim N(0,\sigma^2)$$

where **Y** is the value of the response variable,

β_0 is the intercept of the model that is, the effect of Body Mass Index when all the explanatory variables are zero,

β_p is the slope of the model that is the effect of each explanatory variable on the Body Mass Index.

X_s are the explanatory variables in which the variables in classes once were dummy coded as "1" to be presence of the class of variable and "0" as absence of that class.

ε_i is a random error term that is assumed to normally independent and identically distributed with mean $(E(\varepsilon_i))=0$ and variance σ^2.

The least squares method is a common method in linear regression and it is used to find a function that best fits a given set of data (Barnes, 2001). The strength of the least square method is that it minimizes the sum of the n squared errors (SSE) of the predicted values on the fitted line ($\hat{y}_i$) and the observed value(y).

$$\sum_i^n (y_i - \hat{y}_i)^2$$

3.10.2 Assumptions:

a. Relationship between independent and dependent variable is linear.

b. $E(\varepsilon_i)=0$.

c. Errors are normally distributed i.e. $\varepsilon_i \sim N(0,\sigma^2)$.

d. Error terms:

- $Var(\varepsilon)= \sigma^2$, or homoskedastic errors.
- $E(r_{\varepsilon_i,\varepsilon_j}) = 0$, or no autocorrelation.

3.10.3 Quantile Regression

Quantile regression is a statistical technique that provides a more detailed analysis of the relationship between the dependent variable and its independent variables because it provides conditional regression coefficients for each quantile. QR allows the impact of the explanatory variable to vary along the whole range of BMI intake ("Quantile" is a general terminology for what may be referred to as percentile, quintile, decile, quartile, etc., in specific cases). QR is also necessary because it identifies the lower and upper extremes of the BMI distribution which is related to the underweight and overweight part of the population of women of reproductive age. QR methodology is also necessary because it helps in understanding the relationship between variables outside of the mean of the data.

For a random response variable Y with probability distribution function:

$$F(y) = \text{Prob}\ (Y \leq y),$$

The τ^{th} quantile of Y is defined as the inverse function:

$$Q(\tau) = \inf\ \{y : F\ (y) \geq \tau\}\ \text{where}\ 0 < \tau < 1.$$

Let $X = (x_1, \ldots, x_n)$ denote the matrix consisting of n observed vectors of the random vector X, and $Y = (y_1, \ldots, y_n)$ denote the n observed responses.

The model for linear quantile regression in this study is given by:

$$Y_i = \alpha_\tau + \beta_{i\tau} X_i + \varepsilon_i$$

where $\beta_\tau = (\beta_{1\tau}, \ldots, \beta_{p\tau})$ is the unknown p-dimensional vector of parameters and

$\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ is the n dimensional vector of unknown errors

(Assumption: the τ^{th} quantile of ε_i is zero).

The β_τ is a solution of:

$$\min_{\beta\tau \in R^P} \left[\sum_{i \in \{i: yi \leq \alpha\tau + xi\beta\tau\}} \tau |yi - \alpha\tau - xi\beta\tau| + \sum_{i \in \{i: yi \leq \alpha\tau + xi\beta\tau\}} (1-\tau)\ |yi - \alpha\tau - xi\beta\tau| \right]$$

3.11 Assessing the Goodness of Fit for Linear and Quantile Regression

R^2 also known as the coefficient of variation is a quantitative measure of how well the independent variables account for the dependent variable(s). The R^2 indicates how much a dependent variable can be accounted for by having knowledge of the independent variables instead of assuming that each one had the mean value on the dependent variable.

After fitting the linear model to the data set, an assessment was made for the adequacy of fit. The assessment of goodness of fit for the QR model exploited the general idea leading to the typical R^2 goodness of fit index in classical regression analysis (Koenker, 2005). In linear regression models, the goodness of fit was measured by the adjusted R-squared (the coefficient of determination) method:

$$R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2}$$

A quantity related to R^2, known as the adjusted R-squared, R^2_a was also used for judging the goodness of fit.

Pseudo R^2 measures the relative success of the corresponding quantile regression models at a specific quantile in terms of an appropriately weighted sum of absolute residuals. The obtained pseudo R^2 was considered as an index comparing the residual absolute sum of weighted differences using the selected model with the residual absolute sum of weighted differences using a model with only the intercept. The obtained pseudo R^2 ranges between 0 and 1.

It is worth noting that the index cannot be considered a measure of the goodness of fit of the whole model because it is related to a given quantile. In practice, for each considered quantile, the corresponding pseudo R^2 was evaluated at a local level, thereby indicating whether the presence of the covariates influences the considered quantile. The pseudo R^2 was be used to assess the model with the best goodness of fit between nested models.

3.12 Bootstrap Method for Regression Models

Bootstrapping is a nonparametric approach to statistical inference that substitutes computation for more traditional assumptions and asymptotic results. Bootstrap is the most suitable resampling method in QR analysis (Davino et al., 2014; (He and Hu, 2002; Kocherginsky, 2003; Kocherginsky et al., 2005).

Bootstrapping offers precise inferences when the data are not well behaved. Also, it is also applied to data with sampling distributions difficult to derive. Furthermore, it is useful for complex stratified and clustered samples.

The bootstrap approach can be used to estimate standard errors without requiring any assumption of the error distribution. Finally, the capabilities of bootstrap methods are also explored to estimate standard errors in QR.

CHAPTER FOUR

4.0 RESULTS

This study involved analysis of the 2013 NDHS dataset for women aged 15-49 years. The total number of women was 31 482, however sample weights were applied and the sample size was 31 828.

4.1 SOCIO DEMOGRAPHIC CHARACTERISTICS OF RESPONDENTS

4.1.1 Continuous variables

Characteristics of the sample were depicted in Table 4.1. The mean BMI for the study sample was 23.12kg/m^2 (SD=4.36), whereas the 25th percentile was 19.55 kg/m^2, the median was 22.36 kg/m^2 (IQR=4.63), and the 75th percentile was 23.55 kg/m^2.

The mean age was 29.36 years (SD=6.97). The family size is the number of members in the household had a mean of 7(SD=4) while the median number of family size was 6(IQR=4). The mean children ever born i.e the mean number of parity was 4 (SD=3) and the median of 4(IQR=4).

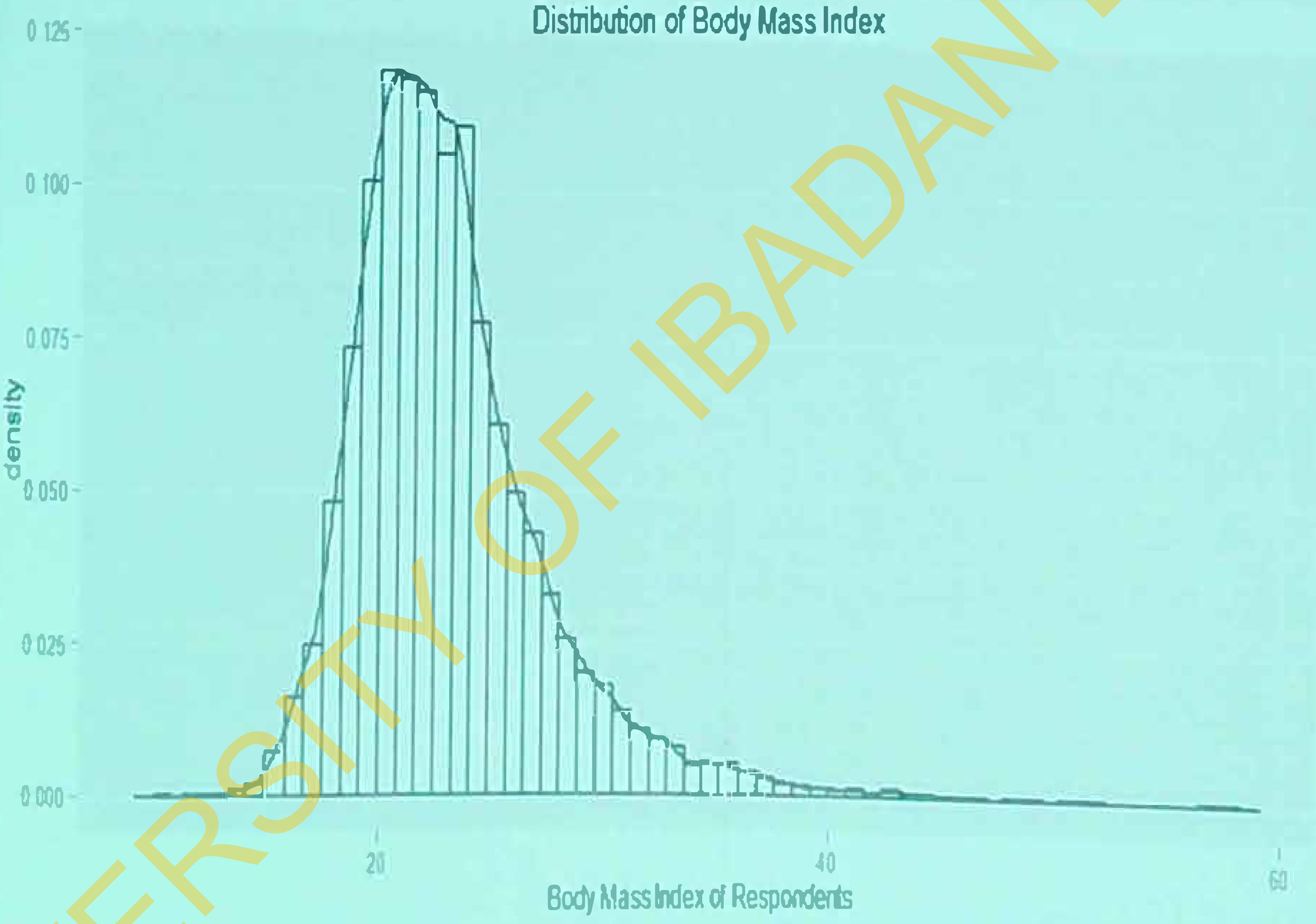
Table 4.1: Descriptive statistics of continuous varible

Variables	N	Mean	Standard deviation	Median	Interquartile range
BMI	31828	23.12	4.36	22.36	4.63
Age	31828	29.36	6.97	29.00	10.00
Family Size	31828	7.01	3.56	6.00	4.00
Total Children Ever Born	31828	4.32	2.58	4.00	4.00

Body Mass Index: Skewness =1.784, Kurtosis=6.546

4.1.2 Distribution of Body Mass Index

Fig. 4.1 shows the distribution of BMI, from which there were some women who belong to the extremely lower parts and extremely higher parts of the distribution. The lowest BMI was 11.72 kg/m^2. There was a shift to the right where the distribution became wider and a large proportion of the samples had a higher BMI. There was a little shift in the lower end of the distribution. There was proportionately much more shifting of the distribution curve at the upper end than the lower end of the distribution.

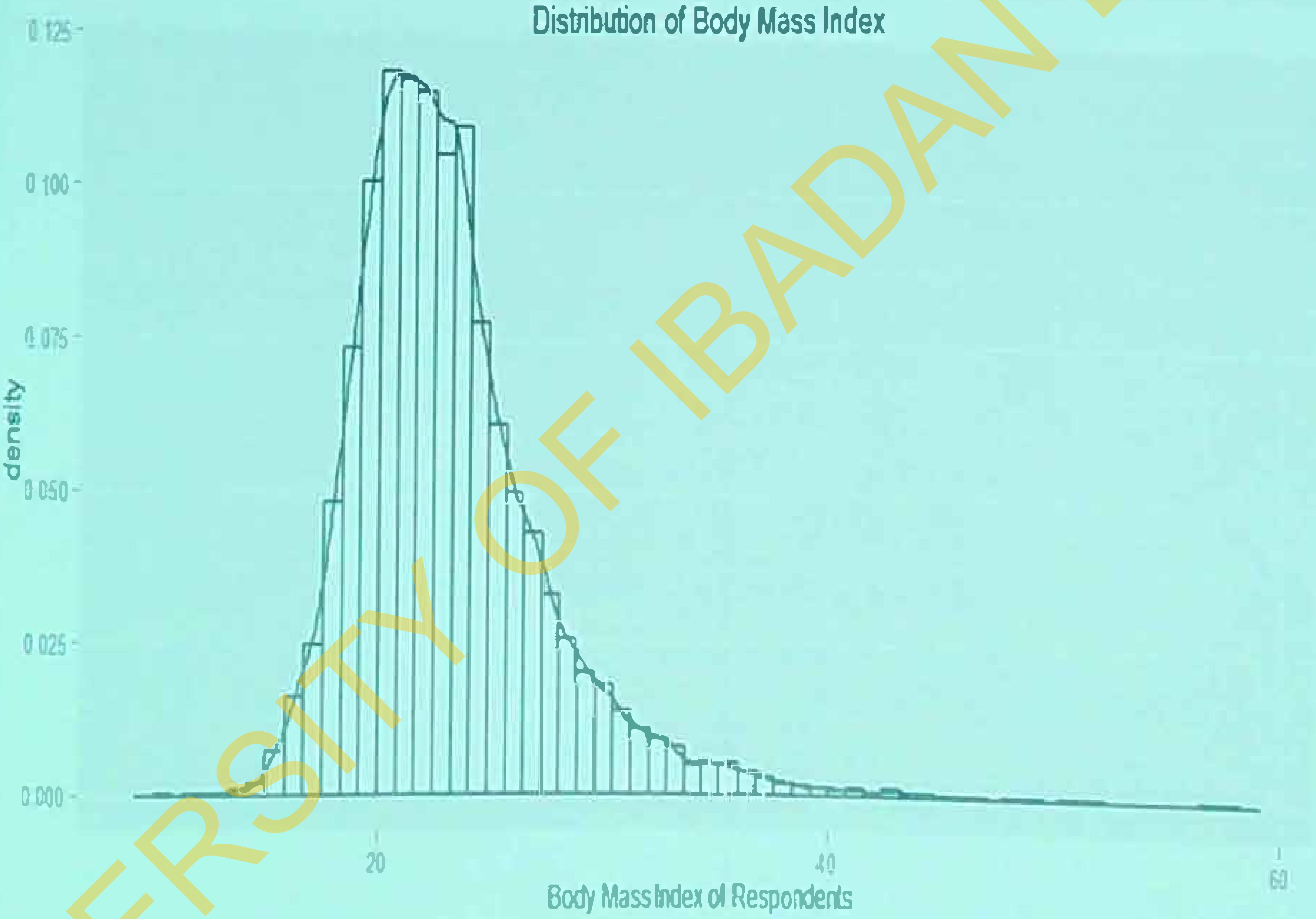


Skewness=1.784, Kurtosis=6.546

Figure 4.1: Body Mass Index Distribution of Women

4.1.2 Distribution of Body Mass Index

Fig. 4.1 shows the distribution of BMI, from which there were some women who belong to the extremely lower parts and extremely higher parts of the distribution. The lowest BMI was 11.72 kg/m^2. There was a shift to the right where the distribution became wider and a large proportion of the samples had a higher BMI. There was a little shift in the lower end of the distribution. There was proportionately much more shifting of the distribution curve at the upper end than the lower end of the distribution.



Skewness=1.784, Kurtosis=6.546

Figure 4.1: Body Mass Index Distribution of Women

4.1.3 Categorical Variables

The distributions for variables in classes are presented in Table 4.2. Women in the age group 25-34 years constitute the highest proportion (49.9%). Women with less than five household members had a proportion of 40.4%. Women with 1-3 children had a proportion of 45.4%. Women who were married had a higher proportion of 95.8% compared to unmarried women. Women with no formal education had the highest proportion of 49.2%. Almost two-third, 65% of the women lived in rural women. The North West geopolitical zone had the highest proportion of women 37.0%.

The distribution of women in terms of wealth quintile revealed that the lowest wealth quintile had the highest percentage of women i.e poorest (23.5%) while the highest wealth quintile had the lowest proportion i.e richest 17.8%.

Table 4.2: Descriptive statistics of categorical variables

Variables	Frequency	Percentage
Age(years)		
15-24	7834	24.6
25-34	15867	49.9
35-49	8128	25.5
Family Size		
<5	12869	40.4
5-8	10704	33.6
>8	8256	25.9
Total Children Ever Born		
1-3	14461	45.4
4-6	11076	34.8
>=7	6291	19.8
Marital Status		
Married	30491	95.8
Unmarried	1337	4.2
Educational Level		
No Education	15657	49.2
Primary	6127	19.3
Post Primary	10044	31.6
Residence		
Urban	11126	35.0
Rural	20702	65.0
Region		
North Central	4340	13.6
North East	5578	17.5
North West	11775	37.0
South East	2840	8.9
South South	2935	9.2
South West	4360	13.7

Table 4.2 (contd): Descriptive statistics of categorical variables

Variables	Frequency	Percentage
Wealth Index		
Poorest	7496	23.5
Poorer	7355	23.1
Middle	6001	18.9
Richer	5656	17.8
Richest	5320	16.7

4.1.4 Distribution of BMI of Women by Socio Demographic Variables

The summary distribution of Body Mass Index of women for some socio-demographic variables is presented in Table 4.3. Average BMI is lowest for women in age group 15-24 years (21.77±3.42). Women with more than 8 household members have a mean BMI of (22.81±4.28). The distribution of BMI of women in terms of children ever born revealed that those with (4-6 children) had the highest BMI (23.40±4.07). Women who are married have an average BMI of (23.11±4.34). Women with the post primary education have a BMI of (24.72±4.85). Women that are urban dwellers have higher mean BMI of (24.47±4.89) compared to rural dwellers. South South regional women have the highest BMI of (24.70±4.73). The poorest women have the lowest BMI out of the wealth quintile (21.66±3.64).

Table 4.3: Summary Statistics of Body Mass Index of women by Socio Demographic Variables

Variables	Mean	SD
Age(years)		
15-24	21.77	3.42
25-34	23.27	4.31
35-49	24.11	4.92
Family Size		
<5	23.12	4.27
5-8	23.34	4.51
>8	22.81	4.28
Total Children Ever Born		
1-3	22.83	4.07
4-6	23.40	4.59
>=7	23.27	4.55
Marital Status		
Married	23.11	4.34
Unmarried	23.31	4.74
Educational Level		
No Education	22.02	3.75
Primary	23.27	4.13
Post Primary	24.72	4.85
Residence		
Urban	24.47	4.89
Rural	22.39	3.86
Region		
North Central	23.40	4.06
North East	22.40	3.97
North West	22.14	3.88
South East	24.65	5.06
South South	24.70	4.73
South West	24.31	4.69

Table 4.3 (contd): Summary Statistics of Body Mass Index of women by Socio Demographic Variables

Variables	Mean	SD
Wealth Index		
Poorest	21.66	3.64
Poorer	22.07	3.52
Middle	22.94	3.84
Richer	23.94	4.49
Richest	25.93	5.17

4.2 GOODNESS OF FIT / MODEL FIT

Table 4.4 shows the estimates of the regression coefficients and their standard errors for the linear (OLS) regression and for the selected quantiles of the quantile (QR) regression. It also shows the adjusted R^2 for the OLS and the Pseudo R^2 for each selected quantile.

R^2 also known as the coefficient of variation is a quantitative measure of how well the independent variables account for the dependent variable(s). The R^2 indicates how much a dependent variable can be accounted for by having knowledge of the independent variables instead of assuming that each one had the mean value on the dependent variable. Pseudo R^2 constitutes a local measure of goodness of fit for a particular quantile.

An R^2 of 0.16 for the OLS regression means that 0.16 or 16% of the variation in the values of BMI can be explained on the basis of the variation in the independent variables that is age, family size, total children ever born, marital status, place of residence, e.t.c, explain 16% of these differences in the BMI. This shows that few of the independent variables have been accounted for.

It can also be seen that the Pseudo R^2 increases with increasing quantile. A Pseudo R^2 of 0.05 for the lowest quantile implies that just 5% of the variation in the values of BMI can be explained on the basis of the regression line of the 10^{th} quantile. The pseudo R^2 of 0.06 means 6% of the variation in the values of BMI can be explained by the regression line of the 25^{th} quantile. The pseudo R^2 of 0.08 for the median shows that 8% of the variation in the values of BMI can be explained, i.e the independent variables explain 8% of the differences in BMI. The pseudo R^2 of 0.12 for the 75^{th} percentile shows that 12% of the variation in the values of BMI can be explained, i.e the independent variables explain 12% of the differences in BMI. The pseudo R^2 of 0.15 for the uppermost quantile shows that 15% of the variation in the values of BMI can be explained, i.e the independent variables explain 15% of the differences in BMI. This shows that the 90^{th} quantile model better predicts the BMI distribution among the quantiles.

Table 4.4: Regression Coefficients and Standard Errors of the Linear Regression and for Selected Quantiles of the Quantile Regression Models

Variables	OLS Regression β (SE)	Quantile Regression 10th Quantile β (SE)	25th Quantile β (SE)	50th Quantile β(SE)	75th Quantile β (SE)	90th Quantile β(SE)
Age Group						
15-24[C]						
25-34	0.87(0.06)	0.39(0.08)	0.59(0.06)	0.83(0.07)	1.04(0.09)	1.30(0.13)
35-49	1.72(0.09)	0.64(0.10)	1.09(0.10)	1.61(0.09)	2.18(0.14)	2.80(0.21)
Family Size						
<5[C]						
5-8	0.01(0.06)	0.16(0.07)	0.11(0.06)	0.04(0.06)	-0.02(0.07)	-0.21(0.13)
>8	-0.11(0.07)	0.03(0.08)	-0.03(0.07)	-0.10(0.07)	-0.14(0.09)	-0.19(0.16)
Children Ever Born						
1-3[C]						
4-6	0.48(0.06)	0.09(0.09)	0.19(0.07)	0.32(0.06)	0.38(0.08)	0.96(0.13)
>=7	0.53(0.09)	0.17(0.11)	0.21(0.09)	0.26(0.10)	0.39(0.12)	1.16(0.17)
Marital Status						
Unmarried[C]						
Married	0.16(0.11)	0.53(0.14)	0.46(0.14)	0.08(0.16)	-0.07(0.17)	-0.05(0.27)
Place of Residence						
Urban[C]						
Rural	-0.22(0.06)	0.06(0.08)	-0.01(0.07)	-0.18(0.07)	-0.32(0.10)	-0.50(0.16)
Level of Education						
No Education[C]						
Primary	0.42(0.07)	0.36(0.07)	0.33(0.05)	0.34(0.07)	0.34(0.10)	0.49(0.15)
Postprimary	0.96(0.08)	0.69(0.09)	0.72(0.07)	0.80(0.08)	0.78(0.11)	1.27(0.21)

SE = Bootstrapped Standard Error with 200 replications

C = Reference Category

Table 4.4 (contd): Regression Coefficients and Standard Errors of the Linear Regression and for Selected Quantiles of the Quantile Regression Models

Variables	OLS Regression	Quantile Regression				
	β (SE)	10th Quantile β (SE)	25th Quantile β (SE)	50th Quantile β(SE)	75th Quantile β (SE)	90th Quantile β(SE)
Level of Education						
No Education[C]						
Primary	0.42(0.07)	0.36(0.07)	0.33(0.05)	0.34(0.07)	0.34(0.10)	0.49(0.15)
Postprimary	0.96(0.08)	0.69(0.09)	0.72(0.07)	0.80(0.08)	0.78(0.11)	1.27(0.21)
Region						
North Central[C]						
North East	-0.49(0.08)	-0.92(0.09)	-0.71(0.07)	-0.47(0.08)	-0.24(0.12)	-0.16(0.18)
North West	-0.71(0.08)	-0.79(0.09)	-0.56(0.06)	-0.73(0.07)	-0.86(0.09)	-1.05(0.18)
South East	-0.06(0.10)	-0.21(0.11)	-0.16(0.11)	-0.18(0.12)	0.04(0.15)	-0.28(0.27)
South South	0.13(0.09)	-0.05(0.10)	-0.03(0.09)	0.04(0.10)	0.31(0.15)	0.07(0.21)
South West	-0.78(0.09)	-0.81(0.11)	-0.84(0.10)	-0.86(0.09)	-0.64(0.13)	-0.83(0.28)
Wealth Index						
Poorest[C]						
Poorer	0.33(0.07)	0.15(0.08)	0.32(0.06)	0.30(0.06)	0.39(0.08)	0.65(0.13)
Middle	0.82(0.08)	0.29(0.08)	0.57(0.07)	0.62(0.07)	1.01(0.10)	1.65(0.20)
Richer	1.51(0.09)	0.48(0.11)	0.90(0.09)	1.34(0.10)	1.99(0.12)	2.86(0.21)
Richest	3.20(0.11)	1.43(0.12)	2.29(0.12)	2.90(0.13)	4.11(0.18)	5.71(0.32)
Constant	20.98(0.15)	17.77(0.19)	18.83(0.17)	20.85(0.19)	22.83(0.22)	24.48(0.39)
Adj. R^2	0.16					
Pseudo R^2		0.05				
Pseudo R^2			0.06			
Pseudo R^2				0.08		
Pseudo R^2					0.12	
Pseudo R^2						0.15

SE = Bootstrapped Standard Error with 200 replications. C= Reference Category

4.3 Comparison of the results of the OLS with the QR analysis of factors affecting the Nutritional Status of Women

The regression coefficients are presented again in Table 4.5 but now together with their p-values to facilitate the identification of significant effects of factors. It should be remembered that relatively low BMI in the lower end of the distribution indicates problems associated with being underweight, while higher BMI at the upper levels of the distribution indicates more problems with being overweight.

Figures 4.2 to 4.6 illustrate graphically the effects of each of the demographic factors on BMI. In each figure and in each panel, the solid line represents estimates of the coefficient from the quantile regression while the shaded area represents the 95% confidence interval. Superimposed on the plot is a dashed line representing the ordinary least squares (OLS) estimate of the regression coefficient.

The coefficient estimates of the indicator or dummy variables reported in the tables are relative to the reference category.

In Table 4.5, according to the linear regression model, for each unit rise in the mean BMI of mothers aged 15-24 years, that of mothers aged 25-34 years increased by 0.87kg/m^2 while that of mothers aged 35-49 years increased by 1.72kg/m^2. But in the quantile regression results, the magnitude of the changes differed depending on the location of the woman in the BMI distribution. The corresponding rises for these age groups at the 10th quantile are 0.39 and 0.64 while they are 1.04 and 2.18 at the 75th quantile. Thus the OLS model overestimated the effect for thinner women (those at the lower quantiles) but underestimated it for obese women at the upper quantiles.

Panels 1 and 2 show that on the average, women aged 25-34 years and 35-49 years have larger BMI than those aged 15-24 years by about 0.87 kg/m^2 and 1.72kg/m^2 respectively according to the OLS estimate but the quantile regression estimate was smaller in the lower quantiles of the distribution and considerably higher (1.30 kg/m^2 and 2.80kg/m^2) in the upper tail of the distribution.

For each unit rise in the mean BMI of mothers who have a family size of less than 5 household members, that of mothers with 5-8 members increased by 0.01, while that of mothers with more than 8 members in the household decreased by 0.11. In the QR results, the magnitude of the changes differs. The negative family size effect was smaller among obese women.

The QR estimates in panels 3 and 4 show the BMI of women with 5-8 household members increased at the lower ends of the BMI distribution and also that of women with more than 8 household members was larger at the lower ends of the BMI distribution, and there is a dramatic widening for those with less than 5 children at the upper end of the BMI distribution. However, family size did not contribute significantly to the BMI effect.

For each unit rise in the mean BMI of mothers who have ever given birth to 1-3 children, that of mothers who have given birth to 4-6 children increased by 0.48 while that of mothers with seven or more children increased by 0.53 according to the OLS result. The QR results showed that the magnitude of the changes differed as the location of the women in the BMI distribution changed. OLS overestimated the effect in all women except those at the highest quantiles. The effect was smaller in underweight women but much bigger in obese women. This was well illustrated in panels 5 and 6 where the BMI of women with 4-6 children increased at the upper tail of the distribution and that of women with more than 6 children rises at the upper tail of the distribution

For each unit rise in the mean BMI of mothers who were unmarried, that of married women increased by 0.16 which was not statistically significant. In the QR results, the effect was bigger and statistically significant at the 10^{th} and 25^{th} quantiles but decreased and not statistically significant at the 75^{th} and 90^{th} quantiles. Panel 7 shows that on average, married women have larger BMI than unmarried women while the QR estimates indicate that the BMI of married women was larger at the lower tails of the distribution while it was lower at the upper tails of the distribution and also there was a dramatic widening for unmarried women at the upper end of the BMI distribution.

For each unit rise in the mean BMI of mothers who resided in the urban area, that of mothers who resided in rural area decreased by 0.22. In the QR results, the magnitude of the changes differed as the quantile increases. There were bigger decreases among women at the 75^{th} and 90^{th} quantiles. This was well illustrated in panel 8.

For each unit rise in the mean BMI of mothers who had no education, that of mothers with primary education increased by 0.42, while that of mothers with post primary increased by 0.96. In the QR results, the magnitude of the changes differed depending on the location of women in the BMI distribution. The OLS model consistently overestimated this effect in all but those at the highest quantile where it was underestimated. The QR results showed these effects to be smaller except at the 90th quantile. Panels 9 and 10 show that the effect was more pronounced at the 90th quantile

As regards regional variation, compared with the North Central region, the OLS results show BMI to be lower by 0.49 in the North East, by 0.71 in the North West and by 0.78 in the South West. The QR analysis shows that these OLS results underestimated these effects in the North East at the 10th and 25th quantiles where the decreases are respectively 0.92 and 0.91, but overestimated them at the 75th and 90th quantiles where the decreases are 0.24 and 0.16 respectively. In the North West, the OLS model underestimated the effects at the 75th and 90th quantiles. In the South West, there is fair agreement between the OLS and the QR results. This was well demonstrated in panels 11-15.

With regards to wealth index, compared to the poorest class, the OLS analysis show that BMI increased progressively to 0.33 kg/m^2 for the poorer class right up to 3.20 kg/m^2 for the richest class. These results overestimated the effects shown by the QR model at all quantiles below the 50th, but underestimated the effects seen at the 75th and the 90th quantiles.

Panels 16-19 show that the effects of the other wealth quintiles compared to poorest was more pronounced at the 90th quantile.

Table 4.5: Regression coefficients and p-values of Linear Regression and Quantile Regression Estimates for Selected Quantiles

Variables	OLS Regression	Quantile Regression				
	β (p-value)	10th Quantile β (p-value)	25th Quantile β (p-value)	50th Quantile β(p-value)	75th Quantile β (p-value)	90th Quantile β(p-value)
Age Group						
15-24[C]						
25-34	0.87(<0.001)	0.39(<0.001)	0.59(<0.001)	0.83(<0.001)	1.04(<0.001)	1.30(<0.001)
35-49	1.72(<0.001)	0.64(<0.001)	1.09(<0.001)	1.61(<0.001)	2.18(<0.001)	2.80(<0.001)
Family Size						
<5[C]						
5-8	0.01(0.841)	0.16(0.015)	0.11(0.056)	0.04(0.466)	-0.02(0.818)	-0.21(0.102)
>8	-0.11(0.085)	0.03(0.716)	-0.03(0.666)	-0.10(0.138)	-0.14(0.110)	-0.19(0.229)
Children Ever Born						
1-3[C]						
4-6	0.48(<0.001)	0.09(0.283)	0.19(0.007)	0.32(<0.001)	0.38(<0.001)	0.96(<0.001)
>=7	0.53(<0.001)	0.17(0.109)	0.21(0.021)	0.26(0.007)	0.39(0.001)	1.16(<0.001)
Marital Status						
Unmarried[C]						
Married	0.16(0.140)	0.53(<0.001)	0.46(0.001)	0.08(0.636)	-0.07(0.666)	-0.05(0.857)
Place of Residence						
Urban[C]						
Rural	-0.22(<0.001)	0.06(0.467)	-0.01(0.942)	-0.18(0.011)	-0.32(0.001)	-0.50(0.002)
Level of Education						
None[C]						
Primary	0.42(<0.001)	0.36(<0.001)	0.33(<0.001)	0.34(<0.001)	0.34(0.001)	0.49(0.001)
PostPrimary	0.96(<0.001)	0.69(<0.001)	0.72(<0.001)	0.80(<0.001)	0.78(<0.001)	1.27(<0.001)

C = Reference Category

Table 4.5: Regression coefficients and p-values of Linear Regression and Quantile Regression Estimates for Selected Quantiles

Variables	OLS Regression	Quantile Regression				
	β (p-value)	10th Quantile β (p-value)	25th Quantile β (p-value)	50th Quantile β(p-value)	75th Quantile β (p-value)	90th Quantile β(p-value)
Age Group						
15-24[C]						
25-34	0.87(<0.001)	0.39(<0.001)	0.59(<0.001)	0.83(<0.001)	1.04(<0.001)	1.30(<0.001)
35-49	1.72(<0.001)	0.64(<0.001)	1.09(<0.001)	1.61(<0.001)	2.18(<0.001)	2.80(<0.001)
Family Size						
<5[C]						
5-8	0.01(0.841)	0.16(0.015)	0.11(0.056)	0.04(0.466)	-0.02(0.818)	-0.21(0.102)
>8	-0.11(0085)	0.03(0.716)	-0.03(0.666)	-0.10(0.138)	-0.14(0.110)	-0.19(0.229)
Children Ever Born						
1-3[C]						
4-6	0.48(<0.001)	0.09(0.283)	0.19(0.007)	0.32(<0.001)	0.38(<0.001)	0.96(<0.001)
>=7	0.53(<0.001)	0.17(0.109)	0.21(0.021)	0.26(0.007)	0.39(0.001)	1.16(<0.001)
Marital Status						
Unmarried[C]						
Married	0.16(0.140)	0.53(<0.001)	0.46(0.001)	0.08(0.636)	-0.07(0.666)	-0.05(0.857)
Place of Residence						
Urban[C]						
Rural	-0.22(<0.001)	0.06(0.467)	-0.01(0.942)	-0.18(0.011)	-0.32(0.001)	-0.50(0.002)
Level of Education						
None[C]						
Primary	0.42(<0.001)	0.36(<0.001)	0.33(<0.001)	0.34(<0.001)	0.34(0.001)	0.49(0.001)
PostPrimary	0.96(<0.001)	0.69(<0.001)	0.72(<0.001)	0.80(<0.001)	0.78(<0.001)	1.27(<0.001)

C = Reference Category

Table 4.5 (contd): Regression coefficients and p-values of Linear Regression and Quantile Regression Estimates for Selected Quantiles

Variables	OLS Regression	Quantile Regression				
	β (p-value)	10th quantile β (p-value)	25th quantile β (p-value)	50th quantile β(p-value)	75th quantile β (p-value)	90th quantile β(p-value)
Region						
North Central[C]						
North East	-0.49(<0.001)	-0.92(<0.001)	-0.71(<0.001)	-0.47(<0.001)	-0.24(0.046)	-0.16(0.389)
North West	-0.71(<0.001)	-0.79(<0.001)	-0.56(<0.001)	-0.73(<0.001)	-0.86(<0.001)	-1.05(<0.001)
South East	-0.06(0.581)	-0.21(0.055)	-0.16(0.162)	-0.18(0.111)	0.04(0.813)	-0.28(0.316)
South South	0.13(0.135)	-0.05(0.644)	-0.03(0.740)	0.04(0.706)	0.31(0.043)	0.07(0.752)
South West	-0.78(<0.001)	-0.81(<0.001)	-0.84(<0.001)	-0.86(<0.001)	-0.64(<0.001)	-0.83(0.003)
Wealth Index						
Poorest[C]						
Poorer	0.33(<0.001)	0.15(0.049)	0.32(<0.001)	0.30(<0.001)	0.39(<0.001)	0.65(<0.001)
Middle	0.82(<0.001)	0.29(<0.001)	0.57(<0.001)	0.62(<0.001)	1.01(<0.001)	1.65(<0.001)
Richer	1.51(<0.001)	0.48(<0.001)	0.90(<0.001)	1.34(<0.001)	1.99(<0.001)	2.86(<0.001)
Richest	3.20(<0.001)	1.43(<0.001)	2.29(<0.001)	2.90(<0.001)	4.11(<0.001)	5.71(<0.001)
Constant	20.98(<0.001)	17.77(<0.001)	18.83(<0.001)	20.85(<0.001)	22.83(<0.001)	24.48(<0.001)
Adj. R2	0.16					
Pseudo R2		0.05				
Pseudo R2			0.06			
Pseudo R2				0.08		
Pseudo R2					0.12	
Pseudo R2						0.15

C=Reference Category

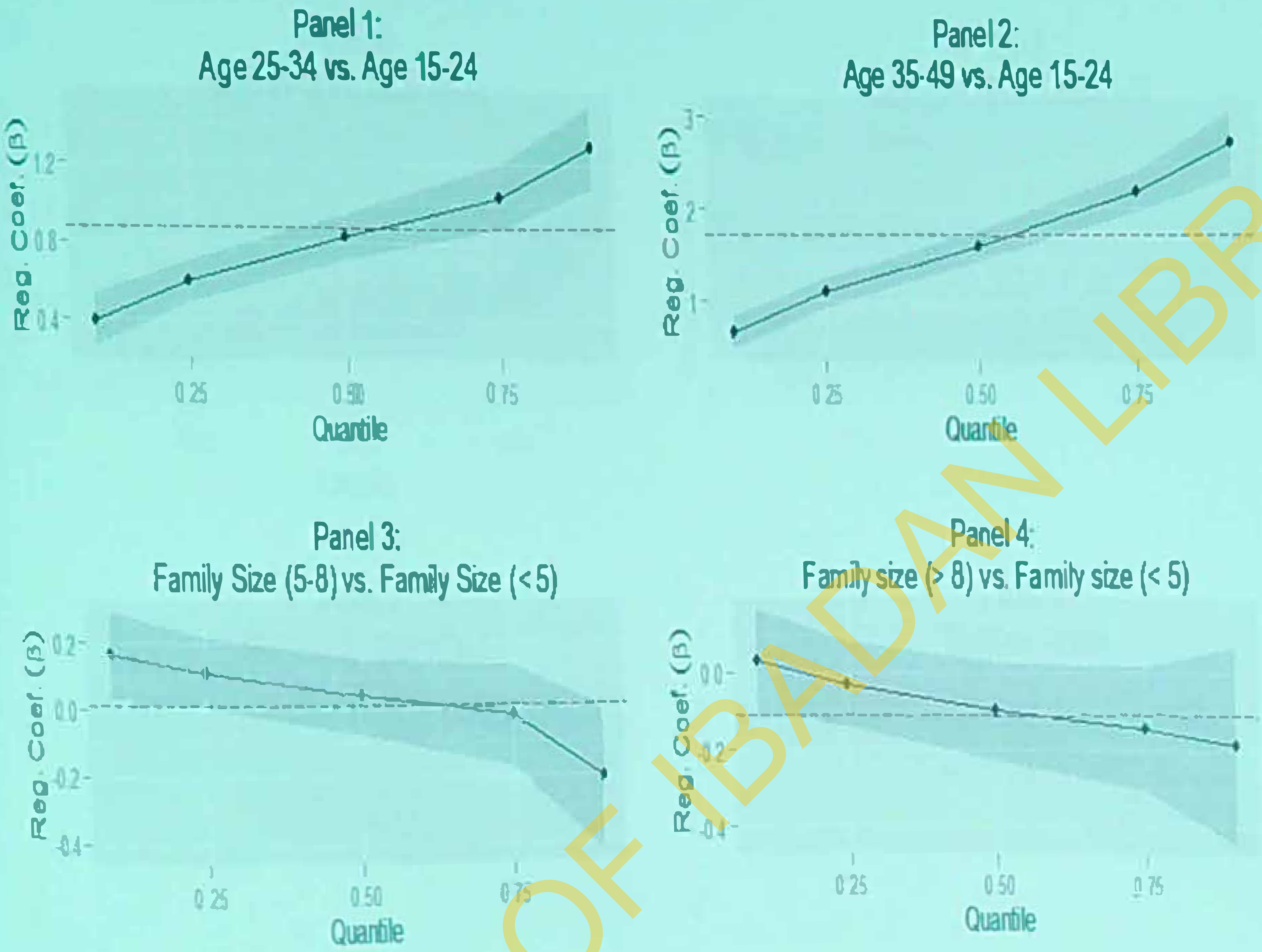


Figure 4.2: Quantile regression models demonstrating effects of woman's age (compared with 15-24 years) and family size (compared with <5) on BMI.

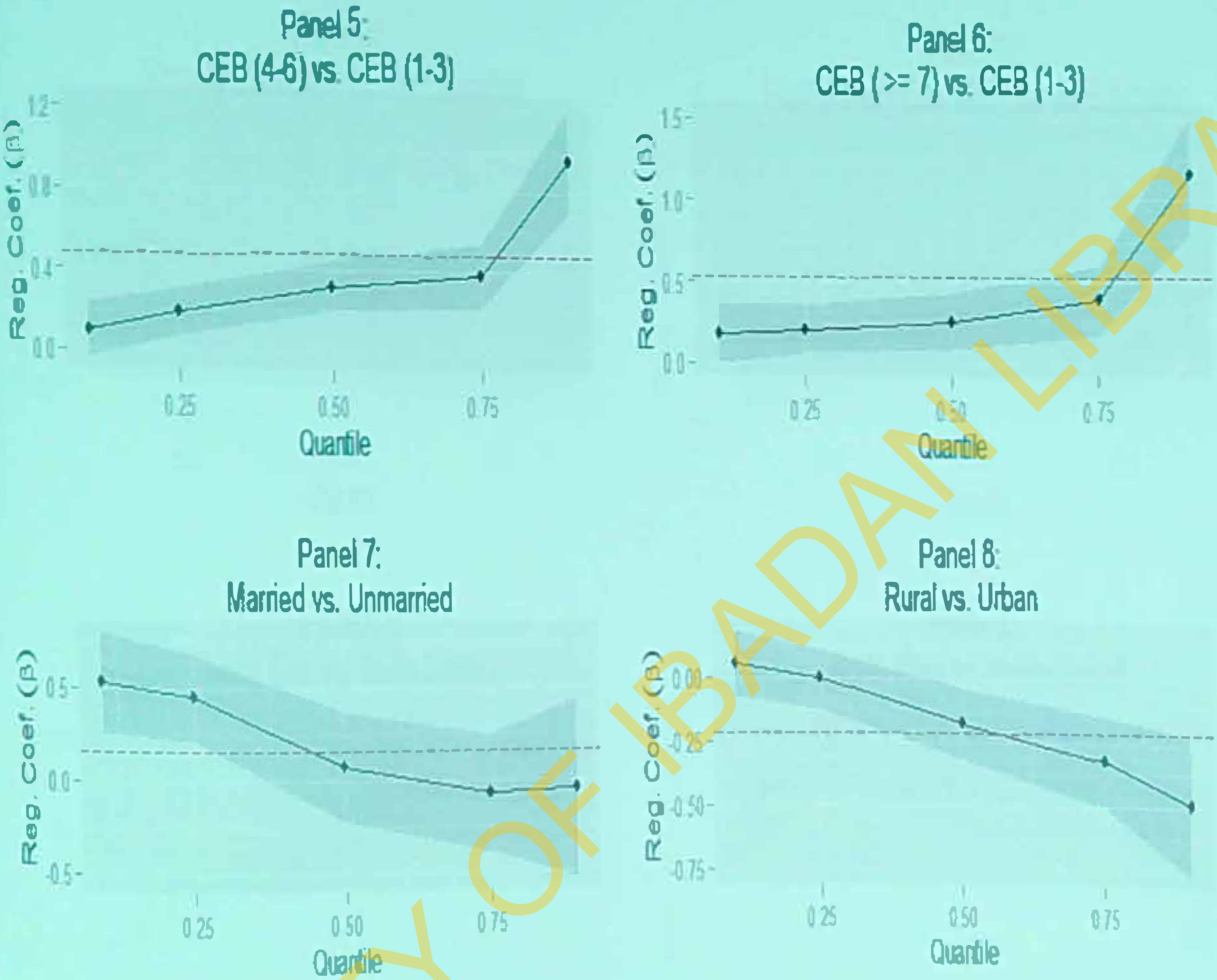


Figure 4.3: Quantile regression models demonstrating effects of children ever born (compared with <5), marital status and place of residence on BMI.

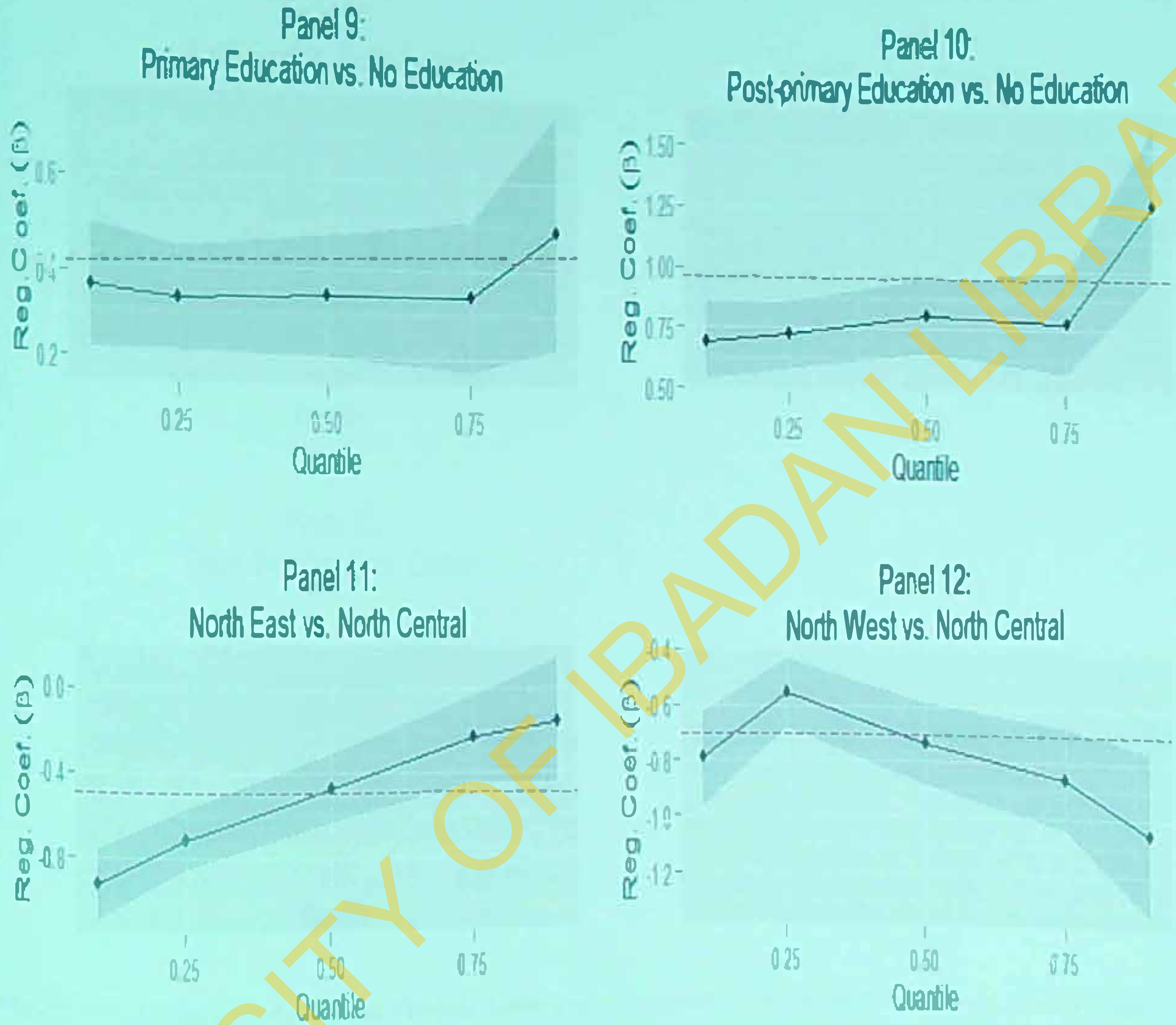


Figure 4.4: Quantile regression models demonstrating effects of education (compared with no education), region-North East and North West (compared with North Central) on BMI.

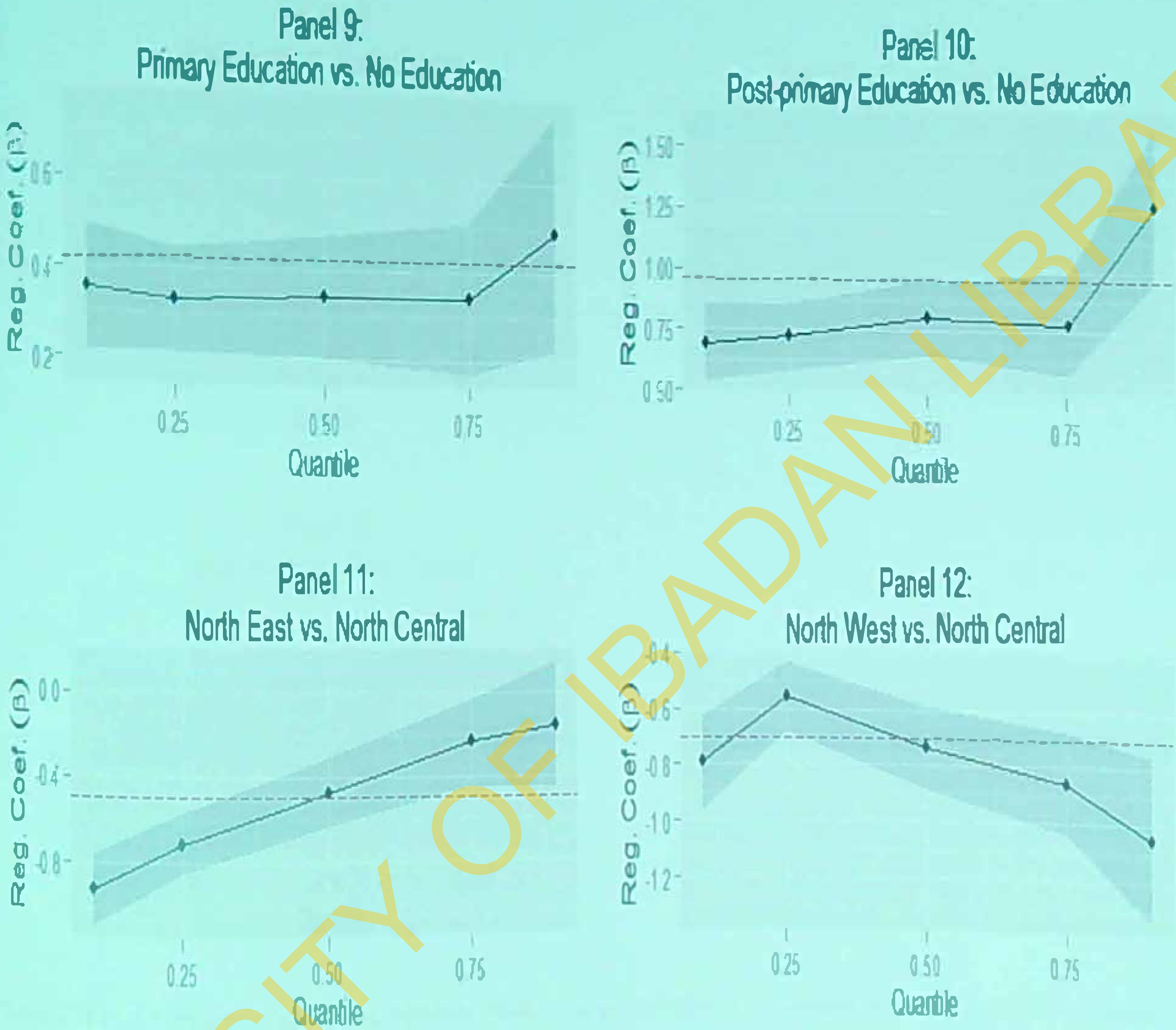


Figure 4.4: Quantile regression models demonstrating effects of education (compared with no education), region-North East and North West (compared with North Central) on BMI.

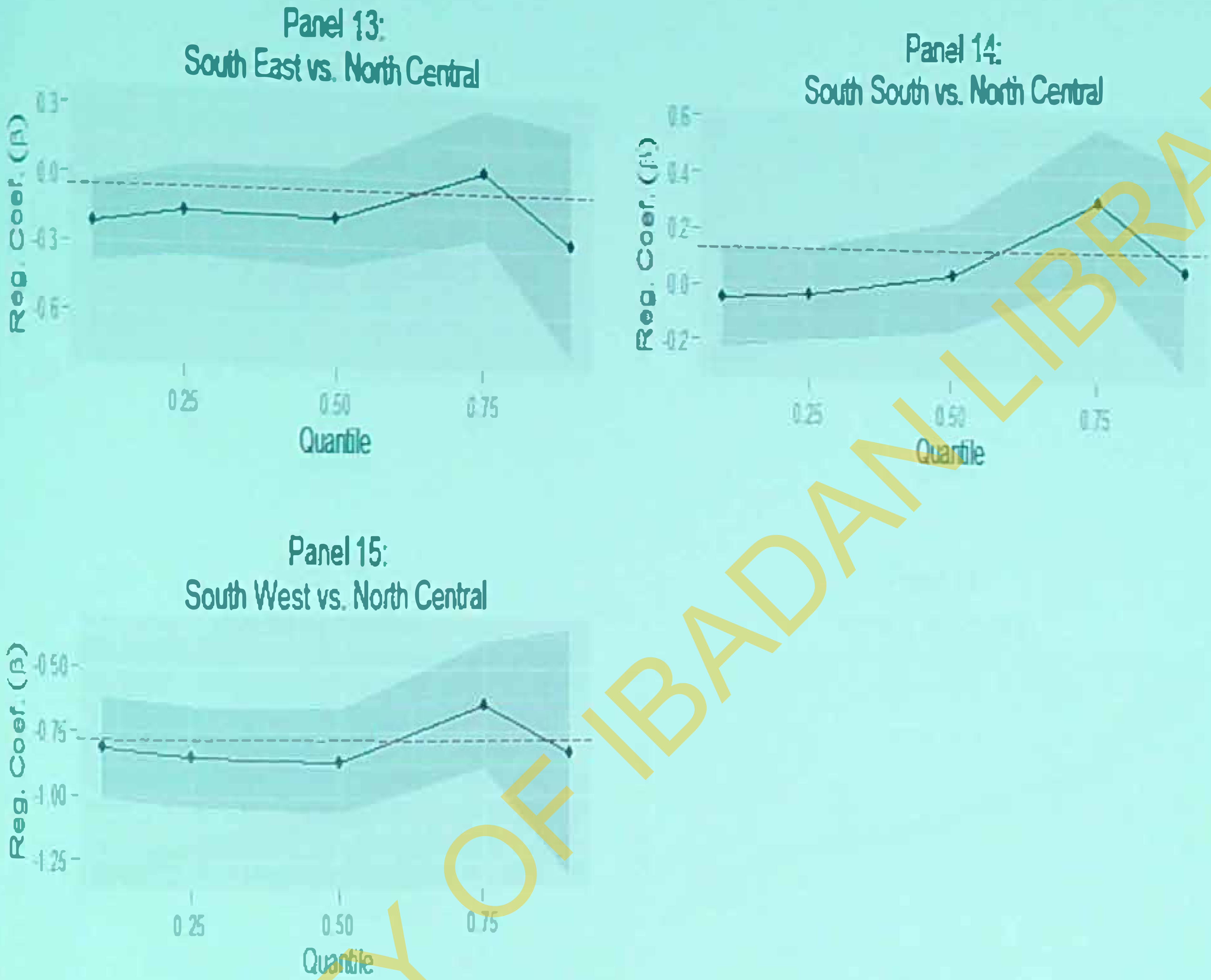


Figure 4.5: Quantile regression models demonstrating effects of region-South East, South South and South West (compared with North Central) on BMI.

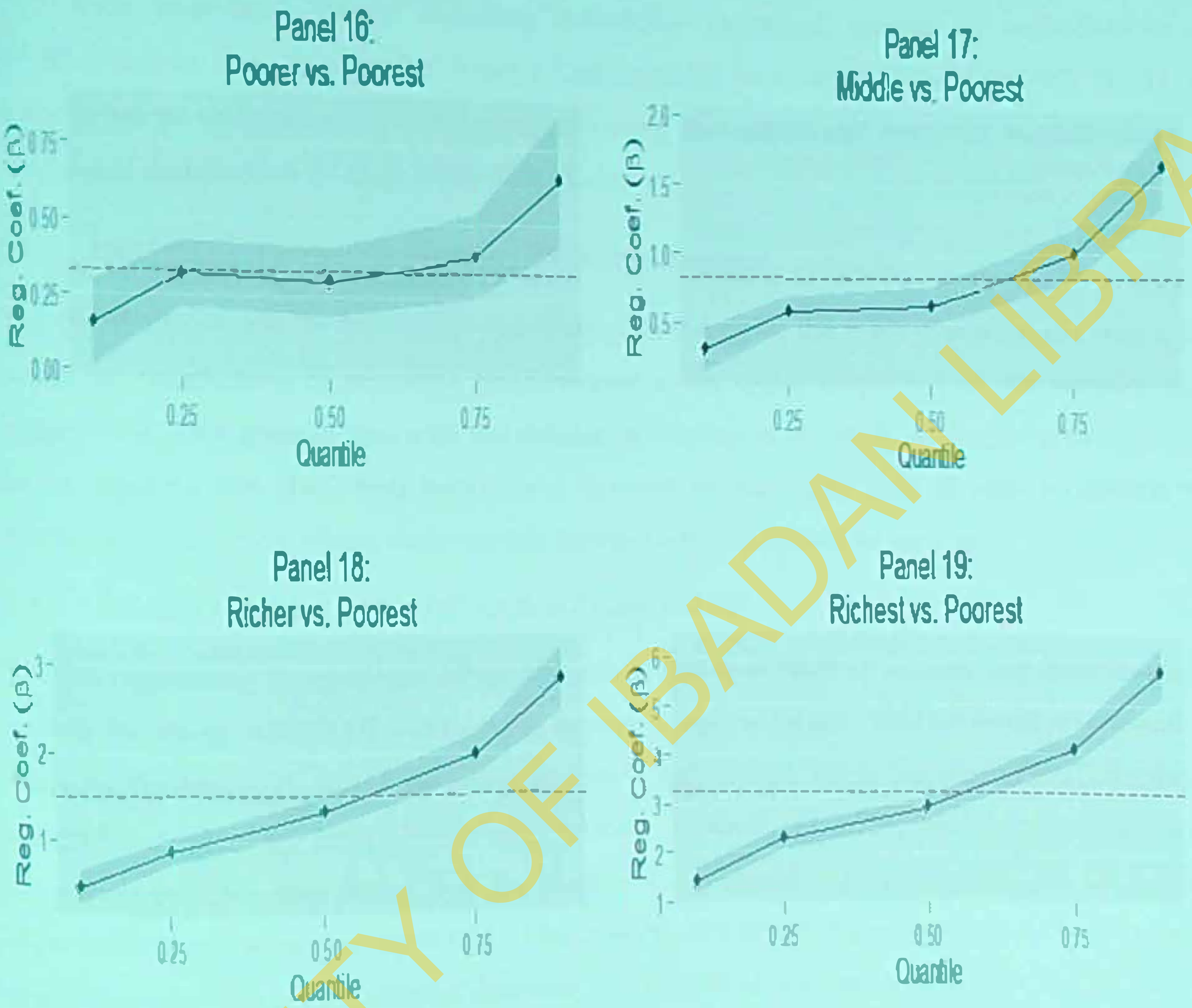


Figure 4.6: Quantile regression models demonstrating effects of wealth index (compared with poorest) on BMI.

CHAPTER FIVE

5.0 DISCUSSION

This study examined factors affecting nutritional status of women of reproductive age (15-49 years) in Nigeria from the Nigeria Demographic and Health Survey (NDHS 2013). The study helps in understanding the factors affecting the nutritional status of women using the conditional distribution of their Body Mass Index.

5.1 DISTRIBUTION OF THE BODY MASS INDEX

The distribution curve in this study was skewed to the right; the main characteristic was a shift where the distribution of women's BMI became wider and a proportion of the samples had a higher BMI. This goes in line with the dataset in Ouyang et al. 2015, where the curve shift was to the right i.e the BMI distribution was skewed to the right. This is also consistent with Ramokolo et al., 2015 where their sample distribution was shifted to the right.

5.2 ORDINARY LEAST SQUARE REGRESSION

In OLS regression, a significant effect was found between BMI of women and their respective ages in the study. Also BMI increased as women's age increased. Similar result was found in a study by Peixoto et al., 2007 regarding women's age; in that, linear regression analysis showed a significant effect on BMI of women. Previous research (Shankar, 2010) estimating a linear regression on BMI data found that for rural area compared with urban area, the OLS results showed BMI to be higher by 0.002. The comparable linear regression (OLS) results in this study showed a somewhat smaller decrease in the BMI of women in rural area. Intuitively, one may expect a rural dweller to be associated with lower BMI values and urban dwellers to have higher increases for higher BMIs.

There was a rise in the BMI of women who were married in a study conducted by Nagata et al., 2009 though marriage was associated with an increase in women's BMI. However, though this study found out that the mean BMI of married women increased, it was not statistically significant.

The Wealth quintiles had statistically significant effects on BMI and as the SES increases, the BMI of women increases in linear regression. This findings were consistent with a study conducted by Nagata et al., 2009 who found out that higher SES respondents had higher BMIs

The level of education has been shown to have a significant effect on women's BMI. The higher the education of women, the higher her BMI. However Shankar, 2010 found out a statistical insignificance education variable which revealed the more educated a woman is, the more she may be expected to better comprehend information relating to health. This difference in both studies may be due to Shankar, 2010 use of a continuous education variable

5.3 COMPARISON OF THE RESULTS OF THE ORDINARY LEAST SQUARE REGRESSION WITH THE QUANTILE REGRESSION ANALYSIS

The present study shows that the BMI of women, in OLS and all the selected quantiles were significantly influenced by their age groups, the BMI of women increased as their age increased. Compared to women in the youngest age group (aged 15-24 years), the positive age effect is smaller among thinner women than among obese ones for both age groups. This finding is similar to a study that reported women in the oldest age group were more likely to be overweight and obese than the youngest women. (Okoh, 2013; Uthman, 2009).

For women with a family size of 5-8, results from OLS showed that family size did not have a significant effect on women's BMI, while at the lowest quantile, family size had a significant effect on BMI. This shows that a woman with 5-8 children compared to those with <5 children are more likely to be thinner. Also for women with more than 8 household members, the effect on BMI were not significant for OLS and in any quantile, but the magnitude of the changes differs; the negative family size effect is smaller among obese women. Family size did not contribute significantly to the BMI effect.

Results from both OLS and QR analyses show that children ever born had a significant effect on women's BMI except at the lowest quantile (10th) where children ever born had no significant effect on women's BMI. The QR results show that the magnitude of the changes differ as the location of the women in the BMI distribution changes. The effect is smaller in underweight women but much bigger in obese women. This agrees with findings from Yahaya et al, 2007 who reported that the higher the number of children, the more difficult for a mother to attain optimum nutritional requirements.

For OLS, though married individuals have higher BMI compared to unmarried women, it showed that marriage is not a determinant for increased BMI which is consistent with a study where unmarried individuals have a lower BMI compared to married individuals (Esma Asil et al, 2014). It was found in the QR results of this study that married individuals have a significantly higher BMI compared to unmarried individuals at the 10th and 25th quantile while at the upper quantiles, unmarried women tend to be overweight. These 10th and 25th QR results go in line with another study carried out by Okoh, 2013 which showed that married women were significantly more likely to be overweight or obese than never married women.

For OLS result, the effect is significantly influenced by place of residence while for the selected quantiles, only the ones below the median were not significantly influenced by their residential place. In the QR results, the magnitude of the changes differs as the quantile increases. There is a bigger decrease among obese women. This indicates that women who live in rural areas tend not to add BMI compared to urban women. This is similar to a study conducted by Hill et al, 2014 who found out that the severity of obesity is worse for those living in town limits compared to those living in rural area.

Women's BMI was also found out to be significantly influenced by the level of education for both OLS and QR results. Women with primary education tend to be overweight compared to those with no education. Results showed that women who have post primary education are more likely to be overweight/obese compared to women who have no education. This agrees with findings from Uthman, 2009 who reported that those with secondary or higher education were less likely to be underweight compared to women who had never been to school. The positive education effect is higher among obese women.

In the OLS result, the effect of BMI on North East decreases compared to North Central, also for North West, there is a higher decrease of BMI while in the QR results, the magnitude of the changes differs. The negative region effect is smaller among thinner women in the North East while it is higher among obese women in the North West. The effect of women's BMI is found out to be significantly influenced by women in South West for OLS and QR results; also there is a smaller decrease of the effect of BMI at the 75th quantile.

The findings in this study revealed that as the wealth quintile increases, the BMI increases. Also the OLS and the 50th percentile results are consistent with a previous study in Ghana a developing country in West Africa, where higher/middle class subjects exhibited higher BMI values (Amoah, 2003) and also consistent with a Nigeria study which revealed that higher socio economic status and occupational level were significantly associated with the risk of obesity (Akarolo-Anthony et al, 2014). This result is in agreement with findings that have been reported by studies in Akwa Ibom state of Nigeria and in Morrocco (Ojofeitimi et al. 2007; Goedecke et al, 2006). Compared to poorest women, the wealthiest women were more likely to be overweight and obese (Uthman, 2009). This finding is consistent with this study which reveals that increased wealth increases BMI. Women who are richer and richest had significantly higher BMI throughout the distribution than those with who are poorest.

QR regression analysis was able to show the amount of both underestimation and overestimation produced by the OLS regression.

5.4 MODEL FIT

It can be seen that the lowest (10^{th}) quantile has the lowest value of pseudo R^2. The goodness of fit for BMI is poorer at the lower tail while the 90^{th} quantile model better predicts the BMI distribution among the quantiles.

5.5 STRENGTHS AND LIMITATIONS

Quantile regression method allowed a complete examination of the relationship between the factors/independent variables and Body Mass Index across its entire distribution in that the variables were considered in different percentiles, which facilitated a deeper and fuller analysis. The strength of this study was the use of continuous response variable in regression rather than categorized response variables as in the case of binary or multinomial regressions. Use of continuous response variable led to more flexibility in choosing the level of response to consider in regression. In other words, binary or multinomial regressions are limited to the assigned response categories whereas continuous quantile regression is more flexible in such that one can choose any response level by choosing appropriate values of τ which represent desirable response quantiles to regress.

Another strength of this study is the use of a large sample size.

One of the limitations is that the dataset used did not include pregnant women and women that had just given birth (maternal) nutritional status in the study.

Another limitation is that the survey or dataset did not collect information on dietary intake.

5.6 CONCLUSION

This study identified the determinants of nutritional status of women using BMI. In particular, the study employed a unique statistical method that showed the effects of these factors in each BMI quantile to identify the effect of each variable.

The quantile regression analysis showed that the magnitude of the changes differed depending on the location of the woman in the BMI distribution

While OLS regression can identify the factors influencing Body Mass Index, Quantile regression helps to understand differences across the conditional distribution or levels of BMI. Quantile regression can be used to model specific parts of the BMI distribution and should be preferred to OLS regression if the original scale of the outcome variable was continuous with a non-normal distribution.

5.7 RECOMMENDATION

For variables that are grossly skewed, it will be wrong to use OLS and even the median. Therefore, quantile regression method is suggested.

Also a policy is needed to correct the shift in the BMI distribution so as to tackle overweight/obesity in the country.

REFERENCES

Abreveya, J., 2002. The effects of demographics and maternal behavior on the distribution of birth outcomes. *Empirical Economics* 26: 247–257.

Adams, A. S., Trinacty, C. M., Zhang, F., 2008. Medication Adherence and Racial Differences in A1C Control. *Diabetes Care* 31.5: 916-921.

Adebowale, A. S., Palamuleni, M, E. and Odimegwu C. S., 2015. Wealth and undernourishment among married women in two impoverished nations: evidence from Burkina Faso and Congo Democratic Republic. *BMC Research Notes* 8:34.

Adedoyin, R. A., Mbada, C. E., Awofolu, O. O. and Oyebami, O. M., 2005. The influence of socioeconomic status on casual blood pressures of the adult Nigerians. *European Journal of Cardiovascular Prevention and Rehabilitation* 12:271 – 273.

Ahmed, W. A. M., Ahmed E. A., Arafa, K. A.O., El-Amin, E. I. S, Alostaz, Z. M. and Khalid, K. E., 2014. Nutritional status of mothers and its determinants in Alemtidad area, Khartoum. *Food and Nutritional Sciences*. 5.22: 1-5

Ajieroh, V, 2010. A Quantitative Analysis of Determinants of Child and Maternal Malnutrition in Nigeria. *International Food Policy Research Institute.*

Agee, M. D., 2010. Reducing Child Malnutrition in Nigeria: Combined Effects of Income Growth and Provision of Information about Mothers' Access to Health Care Services. *Soc Sci Med* 71.11: 1973–1980.

Agetta, 2010. Factors Influencing the Nutritional Status of Women of Reproductive Age in Erute Internally Displaced Persons' Camp-Lira District.

Ashton, D. M., Lawrence, H. C., Adams, N. L. and Fleischman, A. R., 2009. Prevention of Preterm Birth. *Obstet Gynecol* 113:925–930.

Akpa, M. R. and Mato, C. N., 2008. Obesity in Nigeria: Current Trends and Management. *Nigerian Medical Practitioner* 54.1: 11-15.

Asirvatham, J., 2009. Examining Diet Quality and Body Mass Index in Rural Areas using a Quantile Regression Framework. *The Review of Regional Studies* 39.2: 149–169.

Austin, P., Tu, J., Daly, P., and Alter, D. (2005). The Use of Quantile Regression in Health Care Research: A Case Study Examining Gender Differences in the Timeliness of Thrombolytic Therapy. *Statistics in Medicine*, 24: 791–816.

Ball, K. and Crawford, D., 2005. Socioeconomic Status and Weight Change in Adults: A Review. *Social Sci. and Med* 60: 1987-2010.

Barnes, D., 2001. A model of the Effect of Strand Length and Strand Thickness on the Strength Properties of Oriented Wood Composites. *Forest Prod. J* 51.9: 36-46.

Benson, T. and Shekar, M. 2006. Trends and Child Under-Nutrition, Disease and Mortality in Sub- Saharan Africa. 2nd edition, Washington, DC: World Bank.

Berg, F. M., 1995. Obesity costs reach $45.8 billion. *Healthy Weight Journal* 9: 67.

Beyerlein, A., Fahrmeir, L., Mansmann, U., Toschke, A. M., 2008 Alternative Regression Models to Assess Increase in Childhood Body Mass Index. *BMC Med Res Methodol* 8:59.

Carlsson, A. C., Wandell, P. E., Gigante, B., Leander, K., Hellenius, M. L. and De Faire, U., 2013. Seven Modifiable lifestyle factors predict reduced risk for ischemic cardiovascular disease and all-cause mortality regardless of body mass index. A cohort study. *Int. J. Cardiol.* 168: 946-952.

Che, J., 2002. Underweight Canadians: Canadian Social Trends. *Statistics Canada Catalogue* 11.8.

Chinedu, N. S. and Emiloju, C. O., 2014. Underweight, overweight and obesity amongst adults in ota, Nigeria. *Journal of Public Health and Epidemiology* 6.7: 235-238.

Cizek, P., 2011. SemiParametrically Weighted Robust Estimation of Regression Models, 55.1: 774–788.

Davino, C., Furno, M. and Vistocco, D., 2014. Quantile Regression Theory and Application.

De Benoist, B., McLean, E., Egli, I. and Cogswell, M., 2009. Worldwide Prevalence of Anaemia. *Public Health Nutrition*. 12.4: 444-454.

Dharmalingam, A., Navaneetham, K., Sturm, R., and Datar, A., 2005. Body Mass Index in Elementary School Children, Metropolitan Area Food Prices and Food Outlet Density. *Public Health* 119: 1059–1068.

Dressler, H. and Smith C., 2013. Food Choice, Eating Behavior and Food Liking Differs between Lean/Normal and Overweight/Obese. Low-Income Women. *Appetite* 65: 145-152.

Esma, A., Metin, S., Surucuoglu, F., Pinar, C., Asli, U., Ayse, O. O., Mustafa, V.Y. and Lale S.A., 2014. Factors that Affect Body Mass Index of Adults. *Pakistan Journal of Nutrition* 13.5: 255-260.

Ene-Obong, H.N., Enugu, G.I. and Uwaegbute, A.C., 2001. Determinants of Health and Nutritional Status of Rural Nigerian Women. *J Health Popul Nutr* 19.4: 320-330.

ESHRE Capri Workshop Group, 2006. Nutrition and Reproduction in Women. *Hum Reprod Update* 12: 193-207.

Ezzati, M., Lopez, A. D., Rodgers, A., Vander Hoorn, S., and Murray, C.J., 2002. Comparative Risk Assessment Collaborating Group. Selected major risk factors and global and regional burden of disease. *Lancet* 360.9343: 1347–1360.

Finer, N., 2003. Obesity. *Clin Med* 3:23-7.

Finucane, M. M., Stevens, G. A. and Cowan, M. J., 2011. Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index).

Flegal, K. M., Graubard, B. I., Williamson, D. F., Gail, M. H., 2007. Cause Specific Excess Deaths Associated with Underweight and Obesity. *JAMA* 293.15:1861-1867.

Flegal, K. M., Carrol, M. D., Cynthia, L., Ogden, C. L. and Curtin, L. R., 2010. Prevalence and Trends in Obesity among US Adults. *JAMA* 303.3: 235-241.

Gannoun, A., Saracco, J. and Yu, K., 2003. Nonparametric prediction by conditional median and quantiles. *Journal of Statistical Planning and Inference* 117: 207–223.

Girma, W., and Genebo, T., 2002. Determinants of Nutritional Status of Women and Children in Ethiopia. *ORC Macro.*

Hao, L. and Naiman, D. Q., 2007. Quantile Regression: Sage Publications, Inc.

He, X. and Hu, F., 2002. Markov Chain Marginal Bootstrap. *Journal of the American Statistical Association* 97.459: 783–795.

Hill, J. L., You, W. and Zoellner, J. M., 2014. Disparities in Obesity among Rural and Urban Residents in a Health Disparate Region. *BMC Public Health* 14: 1051

Ho, P.M., Rumsfeld, J.S., and Masoudi, F.A., 2006. Effect of Medication Non Adherence on Hospitalization and Mortality among Patients with Diabetes Mellitus. *Arch Intern Med* 166.17: 1836-1841.

Hoss F. and Fischbeck P.S., 2014. Performance and robustness of probabilistic river forecasts computed with quantile regression based on multiple independent variables in the North Central USA. *Hydrol. Earth Syst. Sci. Discudd* 11: 11281-11333.

Igiri, A., Ekong, M., Ogan, C. and Odey, P., 2009. Body Mass Index Measure of Young Adult Nigerians Residents in the Calabar Metropolis. *The Internet Journal of Biological Anthropology* 2:2

Islam, M. Z., Akhtaruzzarman, M. and Lamberg-Allardt, C., 2004. Nutritional status of women in Bangladesh: Comparison of Energy Intake and Nutritional Status of Low Income Rural Group with a High Income Urban Group.

Jiten, H., Irani, S. and Partha J. H., 2012. Risk Factors of Undernutrition among Women in the Reproductive Age Group of India: Evidence from NFHS-3. *American-Eurasian Journal of Scientific Research* 7.1: 05-11.

Kamal, S. M., Hassan C. H., Alam G. M., 2015. Dual burden of underweight and overweight among women in Bangladesh: Patterns, prevalence and socio demographic correlates 33.1: 92-105.

Kiday, H., Afework, M. and Meron 2013. Feeding Practices, Nutritional Status and Associated Factors of Lactating Women in Samre Woreda, South Eastern Zone of Tigray, Ethopia. *Nutriton Journal* 12: 28.

Kocherginsky, M., 2003. Extensions of the Markov Chain Marginal Bootstrap. PhD Thesis, University of Illinois Urbana-Champaign.

Kocherginsky, M., He, X. and Mu, Y., 2005. Practical Confidence Intervals For Regression Quantiles. *Journal of Computational and Graphical Statistics* 14.1: 41–55.

Koenker, R., 2005. Quantile Regression. Cambridge: Cambridge University Press.

Kumanyika, S., Jeffery, R. W., Morabia, A., Ritenbaugh, C., and Antipas, V. J., 2002. Obesity Prvevention: the Case for Action. *Int J Obes Relat Metab Disord* 26.3: 425-436.

Letamo, G. and Navaneetham, K., 2014. Prevalence and Determinants of Adult Under-Nutrition in Botswana. *PLoS ONE* 9.7: e102675.

Lingxin, H.. and Naiman, D. Q., 2007. Quantile regression. Press/CRC.

Mohamad, A. P., Mohsen, V., Asma, P., Bijan M., Azadeh, S., Elham, M., Fatemeh, G., and Mohammad, R. Z., 2009. Comparing linear regression and quantile regression to analyze the associated factors of length of hospitalization in patients with gastrointestinal tract cancers. *Italian Journal of Public Health* 6.2: 136-139

Montenegro, C.E., 2001. Wage Distribution in Chile: Does Gender Matter? A Quantile Regression Approach, Working Paper Series 20, the World Bank, Development Research Group.

Mora, J. O., and Nestel, P. S., 2000. Improving Prenatal Nutrition in Developing Countries: Strategies, Prospects and Challenges. *American Journal of Clinical Nutrition* 71: 1353–1363.

Mummery, W. K., Schofield, G. M., Steele, R., Eakin, E.G. and Brown, W. J., 2005. Occupational Sitting Time and Overweight and Obesity in Australian Workers. *Am. J. Preventive Med* 29: 91-97.

Must, A., Spadano, J., Coakley, E. H., and Field, A. E., and Colditz, G., 1999. The Disease Burden Associated with Overweight and Obesity. *The Journal of the American Medical Association* 282: 1523–1529.

Nandi, C., 2000. Maternal Pregravid Weight, Age, and Smoking Status as Risk Factors for Low Birth Weight Births. *Public Health Rep* 107: 658–662.

National Plan of Action on Food and Nutrition, 2005. Goals of the National Food and Nutrition Policy and Plan of Action. National Planning Commission. WHO Global Nutrition Policy Review.

National Population Commission, Nigeria and ICF International, 2014. Nigeria Demographic and Health Survey 2013, Abuja, Nigeria, and Rockville, Maryland, USA: NPC and ICF International.

Nemati, R.A. and Mogadam, P.A., 2008. Assessment of Nutritional Status of Women of Reproductive Age in Ardebil Province, Iran. *Research Journal of Ecological Sciences* 3.1: 116-119

Nube, M. and Van Den Boom, G. J., 2003. Gender and Adult Under-Nutrition in Developing Countries. *Ann Hum Biol* 30: 520–537.

Ogundipe S. and Obinna C., 2010. Obesity: An Underestimated 'Silent Killer' Vanguard Online Edition Special Reports. http://www.vanguardngr.com/2010/10/obesity-an underestimated -silent-killer.

Okwu, G. N., Ukoha, A. I, Nwachukwu, N., and Agha, N. C., 2008. Studies on the Predisposing

Factors of Protein Energy Malnutrition among Pregnant Women in a Nigerian Community. *Journal of Health Allied Science* 6: 1-10.

Oluwatayo, I. B. 2008, Explaining Inequality and Welfare Status of Households in Rural Nigeria: Evidence from Ekiti State. *Humanity and Social Science Journal* 3:70-80.

Onyedikachi O., John. 2015. Robustness of Quantile Regression to Outliers. *American Journal of Applied Mathematics and Statistics* 3.2: 86-88.

Ouyang, Y., Wang, H., Su, C., Wang, Z., Song, Y., Xiao, Y., Du, W., and Zhang, B., 2015. Use of quantile regression to investigate changes in the body mass index distribution of Chinese adults aged 18–60 years: a longitudinal study. *BMC Public Health* 1-10

Owen P., Mtambo, L., Salule, J., Masangwi, Lawrence N., and Kazembe, M., 2014. Analysis of Childhood Stunting in Malawi Using Bayesian Structured Additive Quantile Regression Model. *International Journal of Statistics and Applications* 4.3: 161-171.

Pawlinska-Chmara, R., Wronka, I., Suliga, E. and Broczek, K., 2007. Socio-Economic Factors and Prevalence of Underweight and Overweight among Female Students in Poland. *Journal of Comparative Human Biology* 58: 309–318.

Popkin, B. M., 2001. The Nutrition Transition and Obesity in the Developing World. *J.Nutr* 131: 871-873.

Ramokolo.V., Lombard. C., Chhag M. I., Engebretsen, M. S., Doherty T., 2015. Effects of early feeding on growth velocity and overweight/obesity in acohort of HIV unexposed South African infants and children. *International breasfeeding journal* 10:14

Rotimi, C., Okosun I., Johnson L., Owoaje E., Lawoyin T., Asuzu M.,Kaufman J., Adeyemo A., and Cooper R., 1999. The Distribution and Mortality Impact of Chronic Energy Deficiency among Adult Nigerian Men and Women. *European Journal of Clinical Nutrition* 53: 734-739.

Svedberg, P., 2000. Poverty and Undernutrition: Theory. Measurement and Policy.

Shankar, B., 2010. Obesity in China: the differential impacts of covariates along the BMI distribution. *Silver Spring* 18: 1660–1667.

Sharma, S. V., Gernard, A.D., and Day, R. S., 2008. Nutrition Knowledge Predicts Eating Behaviour of All. *Clin. Nutr.* 55: 805-81.

Siminnialayi, I. M., Emem-Chioma P. C. and Dapper, D. V., 2008. The prevalence of obesity as indicated by BMI and waist circumference among Nigerian adult attending Family Medicine clinics as outpatients in Rivers state. *Niger J Med.* 17:340-345.

Singh, K., Bloom, S., and Brodish, P., 2011. Influence of Gender Measures on Maternal and Child Health in Africa. Measure Evaluation Technical Report.

Sobal, J., Rauschenbach, B. and Frongillo, E. A., 2003. Marital Status Changes and Body Weight Changes: A U S Longitudinal Analysis. *Social Sci. and Med* 56: 1543-1555.

Stevens, G.A., Singh, G.M. and Lu, Y., 2012. National, Regional, and Global Trends In Adult Overweight and Obesity Prevalences. *Popul Health Metr* 10: 22.

Teller, H. and Yimar, G., 2000. Levels and Determinants of Malnutrition in Adolescent and Adult Women in Southern Ethiopia. *Ethiopian Journal of Health Development* 14.1: 57-66.

Terry, M. B., Wei, Y., Esserman, D., 2007. Maternal, Birth, and Early-Life Influences on Adult Body Size in Women. *Am J Epidemiol* 166: 5–13.

Tikimoto, H., Yoshike, N., Kaneda, F., and Yoshita, K., 2004. Thinness among Young Japanese Women. *American Journal of Public Health* 94: 1592–1595.

UNICEF. 2005. Nutrition Status of Mothers in Ethopia. http://dhsprogram.com/pubs/pdf/FA39

Uthman, O. A., 2009. Patterns, distribution, and determinants of under and overnutrition among women in Nigeria: a population-based analysis. *Journal of Public Health* 17.5: 289-299.

Van de Poel, E., Hosseinpoor, A.R., Speybroeck, N., Van Ourti, T., and Vega, J., 2008. Socioeconomic Inequality in Malnutrition in Developing Countries. *Bull World Health Organ* 86.4: 282–291.

Wei, Y., Pere, A., Koenker, R., and He, X. (2006). Quantile Regression Methods for Reference Growth Charts. *Statistics in Medicine* 25, 1369–1382.

WHO. 1998. Obesity: Preventing and Managing the Global Epidemic. Report of WHO Consultation on Obesity, Geneva.

WHO. 2000. Obesity: Preventing and Managing the Global Epidemic Report of a WHO Consultation. WHO Technical Report Series 894.Geneva.

WHO. 2013. Health Topics. Poverty by World Health Organization.

Winkvist, A., Rasmussen, K. M., and Lissner, L., 2003. Associations between Reproduction and Maternal Body Weight: Examining the Component Parts of a Full Reproductive Cycle. *European Journal of Clinical Nutrition* 57: 114–127

Winkvist A, Habicht, J.P. and Rasmussen, K.M., 1998. Linking Maternal and Infant Benefit of a Nutritional Supplement during Pregnancy and Lactation. *Am. J. Clin. Nutr* 68: 656–661.

Yahaya, M. K., Nabinta, R. T. and Olajide, B. R., 2007Nutritional Status of Female Farmers on Agricultural Intensification in Selected Communities of Gombe State, Nigeria. J. Hum. Ecol. 22.4: 285-289.

Zaman, A., Rousseeuw, P. J., and Orhan, M., 2001. Econometric Applications of High-breakdown Robust Regression Techniques. *Economics Letters* 71: 1-8.