A quantitative study on the effects of socioeconomic status on nutritional status in Sub-Saharan African countries dealing with the double burden of malnutrition

Data from the Demographic and Health Surveys (2014-2018)

Student: Jan Vriens (s2753200)
Supervisor: Prof. dr. ir. H. H. Haisma

j.vriens@student.rug.nl
Population Research Centre, University of Groningen
Faculty of Spatial Sciences
GEMTHPOPST
Master's Thesis Population Studies
20-07-2022
Groningen

Abstract

This research paper aims to contribute to the available body of knowledge on the relationship between socioeconomic status [SES] and nutritional status by focusing on women in developing countries dealing with the double burden of (mal)nutrition. Earlier research on this topic found weakening or disappearing relationships in developing countries with increasing obesity levels, and this paper posits that this could be due to the fact that this relationship need not be linear. The countries Cameroon, Lesotho and Zimbabwe were chosen for this purpose, based on prior research on the double burden of (mal)nutrition.

The Capability Approach was used to explain this possible nonlinear relationship theoretically, and to build an empirical model which could be analyzed with Ordered Probit Regression models using data from the Demographic and Health Surveys [DHS]. One separate model for each country, as well as one bigger model for all countries combined, were examined to provide an answer to the research question:

"What are the effects of socioeconomic status on nutritional status for women in Sub-Saharan African countries dealing with the double burden of (mal)nutrition?"

The results showed that there was a significant positive linear relationship between SES and BMI for each country separately, as well as for the combined sample. Indicating that in countries dealing with the double burden of (mal)nutrition, the propensity for individuals to have a higher BMI increases as their SES increases.

In conclusion, this direction of linearity which is normally found in underdeveloped regions indicates that either this research paper misjudged the level of development in Cameroon, Lesotho and Zimbabwe, or that their transition will occur at a later stage of development than what previously has been found in other regions.

Table of contents

List of tables and figures

1. Introduction

1.1. The double burden of (mal)nutrition

All over the world, researchers have found differences in nutritional status between different social groups on the basis of socioeconomic status [SES], especially among women. In the Global North, overweight and obesity rates tend to be higher among those with a lower SES, decreasing stepwise as SES increases (Devaux & Sassi, 2011; Jaacks et al., 2019; Monteiro, Moura, Conde & Popkin, 2004) While many countries in the global south show higher overweight and obesity rates among those with a higher SES, decreasing stepwise as SES decreases (Jaacks et al., 2019). This raises the question of whether another quantitative study on the subject of SES and nutritional status has scientific relevance.

However, while it has been reported that these effects disappear and weaken as countries develop and transition from the second to the first situation. And that this transition is accompanied by rapidly increasing levels of obesity, which cause these countries to suffer from a double burden of (mal)nutrition with undernutrition and overnutrition coexisting among society, little is known about the distribution of nutritional status per socioeconomic group (Jaacks et al., 2019; Sartorius et al., 2015; Abrahams et al., 2013). Is there no more relationship between the two to be found? Or has the relationship simply transitioned into a nonlinear one in these countries dealing with a double burden of (mal)nutrition?

At the base of this double burden of (mal)nutrition lies the nutrition transition theory, first proposed by Popkin (1993). It describes the way in which each region in the world transforms in the way they eat, drink and move, which affects body composition and creates nutritional problems through 5 patterns. (Popkin, Corvalan & Grummer-Strawn, 2020). The first pattern of 'Collecting food', describes a situation in which diets consist mostly of carbohydrates and fibres, with low-fat consumption and high activity levels. Secondly, societies transform into the 'Famine' stage, which refers to a period of food scarcity and low food diversity when societies transition from hunter-gatherers to sedentary lifestyles. Third, the pattern of 'receding famines' describes the carbohydrate decrease and protein and vegetable consumption increase, while physical activity decreases as a result of technological innovation. Fourth, the 'degenerative diseases' pattern describes a transition towards diets high in fat, cholesterol and sugar, accompanied by increasingly sedentary lifestyles. And finally, the fifth stage of 'behavioural change' refers to the emergence of a pattern associated with desires to prolong health and prevent degenerative diseases. (Popkin, 1993; Abrahams et al., 2011; Scott, Ejikeme, Clottey & Thomas, 2012).

However, due to rapid economic expansion and prosperity in developing countries, transitioning countries rarely reach these final stages of the nutritional transition collectively (Abrahams et al., 2011). While a part of society is dealing with the continued burden of undernutrition, another part seems not to be immune to the obesity epidemic and the non-communicable diseases that arise due to obesity. Additionally, this double burden of (mal)nutrition tends to have a disproportionate effect on women, due to differences in wealth and urbanization between men and women (Abrahams et al., 2011; Shetty, 2013). As well as due to an overall preference for overweight body sizes by virtue of its association with wealth, health, strength and fertility. However, this preference transforms into a Western body ideal along with the rest of the shifts in the nutrition transition (Naigaga et al., 2018).

One of these examples of the double burden of (mal)nutrition is currently developing in Sub-Saharan Africa, which consists of low- and middle-income countries where earlier stages of the nutrition transition were prevalent, and more and more societies are now partly entering into the stages where degenerative diseases are becoming more prevalent at a much faster rate than what previously occurred in the industrialized western world, which has led to new problems arising before the old problems had disappeared (Abrahams et al., 2011). Therefore, this research paper has chosen countries in Sub-Saharan Africa as its topic of investigation.

1.2. Double burden in Cameroon, Lesotho & Zimbabwe

Research by Popkin, Corvalan & Grummer-Strawn (2020) identified the double burden of (mal)nutrition at the country level as those countries with a prevalence of wasting of above 15%, a prevalence of stunting of above 30% or a prevalence of women's thinness of above 20%, and a prevalence of adult or child overweight at cutoff points of 20, 30 & 40%. Among 46 Sub-Saharan African studied, they found a double burden of (mal)nutrition at a cutoff point of 40% overweight for 5 countries, 30% overweight for 27 countries and 20% overweight for 29 countries, as is shown in Table 1.

Table 1: Double burden of malnutrition at three different cutoff points for overweight

> 40% overweight	> 30% overweight	> 20% overweight
Botswana, Cameroon, Lesotho, Mauritius & Zimbabwe	Angola, Benin, Burundi, The Central African Republic, Chad, The Comoros, the Democratic Republic of the Congo, Guinea, Liberia, Madagascar, Malawi, Mali, Mozambique, Niger, Nigeria, Rwanda, Sierra Leone, Somalia, Sudan, Tanzania, Uganda & Zambia	Eritrea & Ethiopia

Source: Popkin. Corvalan & Grummer-Strawn (2020)

Since a cross-country analysis of 29 countries is beyond the scope of this research, it will focus only on those countries with > 40% overweight for which data is available in the Demographic and Health Surveys [DHS] (table 2), being Cameroon, Lesotho and Zimbabwe.

The reason that these countries were chosen is that the previously found disappearing relationships between SES and nutritional status were also found in countries with large increases in the prevalence of 'the new problem' of overnutrition (Jaacks et al., 2019). Therefore, the expectation is that these countries are most suitable for testing the hypothesis of (non)linearity of the relationship between SES and nutritional status.

Table 2: Demographic and Health Survey availability

Country	DHS
Botswana	Unavailable
Cameroon	2018
Lesotho	2014
Mauritius	Unavailable
Zimbabwe	2015

Source: The DHS Program (2022)

1.2.1. Development in Cameroon, Lesotho and Zimbabwe

While Popkin, Corvalan & Grummer-Strawn (2020) have documented the double burden of (mal)nutrition in Cameroon, Lesotho and Zimbabwe. And while this is a known indicator that these countries are (partly) entering later developmental stages, other indicators provide better and clearer proof of development.

First and foremost, their development is measured by the Human Development Index [HDI]. The HDI is a multi-dimensional measure, based in part on the capability approach, looking at three dimensions of human development, i.e. length of healthy lives, access to knowledge and education, and the quality of life on a regional or country-level (Robeyns, 2006; United Nations, 2020). Since 1990, when measurement and publishment by the United Nations Development Programme [UNDP] started, all three countries improved by measures of up to 25.7%. The HDI in Cameroon went from 0.448 in 1990 to 0.563 in 2019, HDI in Lesotho went from 0.498 in 1990 to 0.527 in 2019, and HDI in Zimbabwe went from 0.478 in 1990 to 0.571 in 2019 (United Nations, 2020). With these increases, Cameroon and Zimbabwe even passed the threshold for medium-developed countries (United Nations, 2020).

Furthermore, GDP per capita rose from 1045.4 in 1990 to 1537.1 in 2020 for Cameroon, from 350.1 in 1990 to 875.4 in 2020 for Lesotho, and from 842 in 1990 to 1214.5 in 2020 for Zimbabwe (The World Bank, 2020), showing that while HDI was highest in Zimbabwe, GDP is highest in Cameroon.

And finally, life expectancy at birth rose from 52.60 in 1990 to 58.77 in 2015 in Cameroon and from 54.44 in 1990 to 60.83 in 2015 in Zimbabwe (World Population Prospects, 2019). And while life expectancy increases in Lesotho started later, due to a high prevalence of adult HIV at the beginning of this century, they have also seen a steady increase from 44.23 in 2000 to 53.51 in 2015 (United Nations Population Division, 2019; Bor, Herbst, Newell & Bärnighausen, 2013)

In conclusion, while development in Lesotho is lagging behind in regards to the development of Cameroon and Zimbabwe, and even still falls in the low human development category, all countries do show clear signs of development on all three aforementioned measures (United Nations, 2020).

1.3. Research Objective

This research paper aims to contribute to the available body of knowledge on the relationship between SES and nutritional status by focusing on countries in developmental phases where earlier research saw weakening or disappearing relationships (Jaacks et al., 2019; Sartorius et al., 2015). The countries were chosen based on research by Popkin, Corvalan & Grummer-Strawn (2020), who identified a clear case of the double burden of (mal)nutrition, which is known to accompany this development, in Cameroon, Lesotho and Zimbabwe.

Furthermore, in the next chapter, an adaptation of the conceptual framework created by Chiappero-Martinetti & Venkatapuram (2014) is used to explain this phenomenon through the Capability Approach [CA] both theoretically and empirically.

And finally, secondary survey data from the Demographic and Health Surveys [DHS] are used to run several ordered probit regressions that will be analysed to provide an answer to the following research question:

"What are the effects of socioeconomic status on nutritional status for women in Sub-Saharan African countries dealing with the double burden of (mal)nutrition?"

1.4. Structure of the paper

This research paper is divided into five main chapters. This first chapter put forth the societal and scientific relevance, the population under investigation, and the research objective and research question. The second chapter provides a theoretical framework, a conceptual model, and the hypotheses that will be tested. The third chapter describes the source of the data, the operationalization of this data, as well as the research methods used and the ethical considerations that were taken into account. The fourth chapter describes the datasets and shows the results of the analyses. And the final chapter discusses the findings, strengths and limitations, and recommendations, and provides an answer to the research question.

2. Theoretical Framework

2.1. Capability Approach

Amartya Sen first created the capability approach [CA] in the 1970s as a new and more coherent method for evaluating well-being, in terms of people's capabilities to achieve the kind of life they deem valuable. According to Sen, previous frameworks like resourcism and utilitarianism did not suffice, since these only evaluated social welfare on the basis of subjective well-being or the availability of means for a good life (Wells, 2013; Robeyns, 2003; Chiappero-Martinetti & Venkatapuram, 2014).

The CA goes beyond basic principles such as income or resources, negative liberties, basic needs, or utility, and claims that a person reaches the highest level of social welfare when they have the highest degree of capabilities to achieve desired combinations of intrinsically valued functionings that make up a valuable and prosperous life (Wells, 2013; Chiappero-Martinetti & Venkatapuram, 2014).

The first concept in the CA is that of the capability itself, which is best described as the opportunities that a person has to convert resources such as income or employment into valuable beings and doings. Or in other words, what an individual is able to be and do in their daily life (Wells, 2013; Chiappero-Martinetti & Venkatapuram, 2014).

The second concept in the CA is that of the functionings, which can be seen as the fruition of these capabilities into end achievements of what each individual has chosen to value and seek after (Wells, 2013; Chiappero-Martinetti & Venkatapuram, 2014).

Third, there is the concept of agency, which is the extent to which individuals are limited to or able to pursue different functionings (Chiappero-Martinetti & Venkatapuram, 2014).

And finally, there is the concept of conversion factors. These reflect different individual characteristics (age, gender, nationality, etc.) which positively or negatively affect the ability of individuals to turn resources (income, education, etc.) into capabilities (ability to work, ability to eat, etc.) (Chiappero-Martinetti & Venkatapuram, 2014).

The combination of these concepts is then used to conclude that an individual reaches the highest level of social welfare when they have the highest degree of capabilities, which they have the agency to convert into desired combinations of functionings (Wells, 2013; Chiappero-Martinetti & Venkatapuram, 2014).

2.2. A capability perspective on nourishment

It would be tempting to draw a utilitarian conclusion by applying the CA in such a way that having a higher socioeconomic status automatically leads to a healthier nutritional status as it appears to do in the Global North (Devaux & Sassi, 2011; Jaacks et al., 2019). Since with the right conversion factors, the right choices, and the agency to do so, these extra resources can be converted into an ability to take better care of yourself and a functioning of well-nourishment.

However, these conversion factors, choices and agency cause such a conclusion on the basis of the CA to be premature. For instance, the country where an individual is born can act as a conversion factor of resources into over nourishment by having over nourishment as a beauty standard that is associated with wealth, health, strength and fertility (Naigaga et al., 2018). Living alone can cause resources to be converted into (mal)nourishment since eating is not only a physical activity but also a social activity (Odencrants, Bjuström, Wiklund & Blomberg, 2013). And age can cause resources to be converted into (mal)nourishment since older individuals can be impaired in their ability to acquire or digest food (Forster & Gariballa, 2005).

Therefore, while health and well-nourishment are functionings that all individuals are likely to value, it is not reasonable to assume that they are the only functionings that individuals value or choose to pursue. In other words, a higher SES can also be transformed into different functionings, while neglecting the need for health and well-nourishment.

Research by Banerjee & Duflo (2011) on this topic even found that on a household level, as household budgets first rose above poverty levels, they tended not to buy higher quantities of the cheap nutritious food that they consumed while in poverty, but lower quantities of the more expensive and tastier options such as grains, fats and sugars. Indicating that while they had the capability to pursue well-nourishment as a functioning, they instead opted to pursue a functioning of indulgence and a pleasant life that they deemed more valuable.

Additionally, on an individual level, as incomes first improve, individuals tend to decrease their physical activity levels and increase their portion sizes and frequency of fast food intake (Sartorius et al., 2015; Sobngwi et al., 2002; Fezeu et al., 2008; Witkowski, 2007). Further strengthening the belief that having a higher socioeconomic status does not necessarily lead to well-nourishment.

It is only after having experienced this luxury for a while that behaviours among these richer SES groups change, with their functioning shifting to health in accordance with the nutritional transition model (Popkin, 1993; Abrahams et al., 2011; Scott, Ejikeme, Clottey & Thomas, 2012).

Therefore, a more thought-out conclusion on the basis of the CA might be that in the first stages of development, the higher SES groups rise above poverty levels first, causing them to pursue functionings that cause over nourishment. While in the latter stages, in which different socioeconomic groups are undergoing different transitions, the higher SES groups shift their focus towards the 'behavioural change' phase of the nutritional transition, while poorer or middle SES groups will pursue indulgence that accompanies the 'degenerative diseases phase' of the nutritional transition.

Therefore, this research paper posits that keeping all other conversion factors, capabilities and resources constant, the relationship between SES and nutritional status need not be linear. It takes the stance that it is likely that since citizens escaping poverty pursue indulgence and pleasure, these groups suffer more from over-nourishment than the higher SES groups, who shift their focus towards wellbeing. This would then cause a nonlinear relationship between SES and nutritional status in countries dealing with the double burden of (mal)nutrition.

2.3. Transformation for empirical research

Despite the CA having already been used in numerous qualitative and quantitative studies in fields such as sociology, population studies, etc., doubts about its practicality for implementation in empirical research remain. Since defining capabilities as a single metric often warps their theoretical construction, losing the values and reasoning on which they are based (Chiappero-Martinetti & Venkatapuram, 2014).

Chiappero-Martinetti & Venkatapuram (2014) therefore state that empirical research is required to define its concepts with enough detail to make them applicable. Furthermore, they put forward the Demographic and Health Surveys [DHS] as an example of a survey that has defined their concepts successfully. Stating that: "DHS data, provide several data points about preferences, choices and decisions that can be used as a proxy for measuring capabilities."

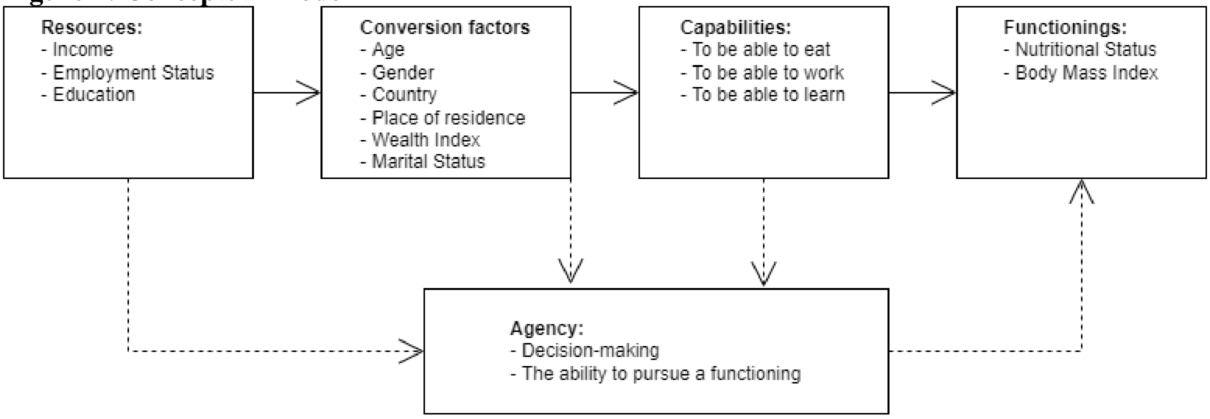
This research paper will therefore use DHS data to operationalize variables that are suited for empirical analysis.

2.4. Conceptual Model

In conclusion, the empirical analysis in this research paper is built on the following conceptual model that is visualized in figure 1. The dependent variable is the functioning of nutritional status, measured as the BMI of the respondents in categories defined by the World Health Organization (2010). And the main predictor is not the resource of income, so as to not generate distortion, but the conversion factor of income; the wealth index. This is a composite measure of a household's assets, which converts income into a much clearer measure of capabilities (The DHS Program, 2020).

Additionally, other resources such as employment status and level of education, and other conversion factors such as age, type of place of residence, marital status, and country will be used as control variables to ensure that the main effect between SES and nutritional status is not a spurious one, since earlier research has given reason to believe that these factors could also influence nutritional status (Forster & Gariballa, 2005; Odencrants et al., 2013; Djuikom & van de Walle, 2022; Wassie et al., 2015; Jaacks et al., 2019).

Figure 1: Conceptual Model



2.5. Hypotheses

Based on the theoretical framework and the expected similarities surrounding nutritional status in Cameroon, Lesotho and Zimbabwe, this research paper hypothesized that:

H1: There is a relationship between socioeconomic status and nutritional status in countries dealing with the double burden of (mal)nutrition when controlling for other resources, conversion factors, and capabilities.

H2: The relationship between socioeconomic status and nutritional status in countries dealing with the double burden of (mal)nutrition when controlling for other resources, conversion factors, and capabilities is nonlinear.

H3: The same relationship between socioeconomic status and nutritional status in countries dealing with the double burden of (mal)nutrition when controlling for other resources, conversion factors, and capabilities is found in Cameroon, Lesotho and Zimbabwe.

3. Data and methods

3.1. Source of data

This explanatory and descriptive study uses secondary survey data from the Demographic and Health Surveys [DHS] (2014 - 2018). The DHS are nationally representative household surveys that provide data on the topics of population, health, and nutrition. It was originally developed in 1984 by the United States Agency for International Development [USAID] and has since been conducted in over 90 countries through more than 400 surveys in overlapping five-year phases (The DHS Program, 2022).

The primary data collection in each country was conducted through four model questionnaires. A Household questionnaire, to collect data on the characteristics of the household, as well as to identify household members who are eligible for an individual interview. Eligible individuals were then interviewed through separate Woman's and Man's questionnaires. And finally, Biomarker questionnaires were conducted to collect objective medical signs, as opposed to self-perceived health indicators. Furthermore, some countries had a special need for a fifth custom questionnaire on specific topics that were not included elsewhere. All interviews that were conducted were only conducted under the condition that the respondents provided full voluntary informed consent (The DHS Program, 2022).

To ensure that information across all countries is comparable, a DHS program was developed to create similar questionnaires and survey procedures in each country. Fieldwork manuals were implemented to explain the standard approach, and to train field staff, interviewers, supervisors and editors. A sampling manual was created to present the general DHS approach to sampling issues. Tabulation manuals were designed to provide model tables that clearly visualise the major findings for policymakers and program managers. And finally, Statistical and Methodological Documentation guidelines were constructed as a reference guide for how to interpret the DHS survey indicators. Thanks to this program, all questionnaires contain essentially the same information, which makes cross-country analysis possible (The DHS Program, 2022).

The survey results are then published in preliminary reports, a more detailed final report and a key findings report. As well as major output being distributed through the DHS Data Archive in the form of micro-level datasets, which will be used for the analysis part of this study (The DHS Program, 2022).

3.2. Operationalization of variables

The definition and operationalization of variables derived from the micro-level DHS datasets are tabulated in Table 3. The control variables were chosen on an empirical basis, where earlier research found significant relationships between them and the dependent variable nutritional status (Forster & Gariballa, 2005; Odencrants et al., 2013; Djuikom & van de Walle, 2022; Wassie et al., 2015; Jaacks et al., 2019).

The selection of each variable was carefully considered, as the inclusion of irrelevant variables could cause the regression coefficients to be less precise estimators of the population. And the exclusion of relevant variables could lead to the wrong conclusions, due to unobserved spurious relationships and omitted variable bias.

The expectation is therefore that the combination of this dependent variable, with these predictors, can best explain the predicted non-linear association.

Table 3: Definition of variables in the study and their operationalization for the analysis

Concepts/variables	*Definition*	*Operationalization*
Dependent variable:		
Nutritional Status	Body Mass Index = $\frac{Weight \in KG}{(Length \in M)^2}$	Recoded as a categorical ordered variable with four increasing options: 0. Underweight (<18.5) 1. Normal weight (18.5 - < 25) 2. Overweight (25 - < 30) 3. Obese (>30) Flagged cases (coded as 9998) were dropped.
Independent variable:		Keep:
Socioeconomic status	Wealth Index: Composite measure of a household's cumulative living standard.	Scale with 5 options: 1. Poorest 2. Poorer 3. Middle 4. Richer 5. Richest Recoded into dummy variables with poorest as the reference category.
Control variables:		
Age	The current age of respondent	Keep: 15-50
Level of education	Highest attained level of education	Keep: Scale with 4 options: 1. No education 2. Primary 3. Secondary 4. Higher Recoded into dummy variables with no education as the reference category
Marital Status	Current marital status of the respondent	Keep: Scale with 6 options 0. Never in union 1. Married 2. Living with partner 3. Widowed 4. Divorced 5. Separated Recoded into dummy variables with never in union as the reference category
Type of place of residence	Rural or urban	Recode to: 0. Urban 1. Rural
Employment status	Whether or not the respondent is currently working.	Keep: 0. No 1. Yes
Country	Which country the respondent is from. (This variable was only used in the cross-country analysis.)	Coded as: 0. Cameroon 1. Lesotho 2. Zimbabwe

3.3. Type of analysis

3.3.1. Descriptive analysis

This research paper first analyses the separate countries on the basis of descriptive statistics such as percentage distributions for the categorical variables, means and standard deviations for continuous variables, and modes and a total N for all variables. It does this to provide an overview of what the sample of Cameroon, Lesotho and Zimbabwe look like as a whole, and what the samples look like for each category of nutritional status. These descriptive statistics can provide an initial overview of possible relationships before they are empirically tested.

Additionally, these descriptive statistics are also used to report on the randomness of the missing values.

3.3.2. Explanatory analysis: Ordered Probit Regression

Given that the dependent variable in the analysis (nutritional status) is a categorical variable for which the categories have an increasing order, based on an underlying continuous measure (BMI), the best estimation of an individual's nutritional status can be achieved using an ordered probit regression model (Gupta & Bansal, 2020; Becker & Kennedy, 1992).

The reason for this is that Ordered Probit Regression methods recognize the increasing difference between the categories, without making the assumption that the distance between categories needs to be equal. Whereas Ordinary Least Squares regression methods assume that the difference between underweight and normal weight can be treated as equivalent to the difference between normal weight and overweight, which is inappropriate in a situation like this (Becker & Kennedy, 1992).

Mathematically, Ordered Probit Regression uses the following function which suggests that nutritional status [y*] is determined by a constant [a], a vector of covariates [$\sum \beta_i \chi_i$], and a disturbance term with a standard normal distribution [ε]:

$$y^* = a + \sum \beta_i \chi_i + \varepsilon$$

Or in the case of this research paper:

$$y^* = a + \beta_1 * WIpoorer + \beta_2 * WImiddle + \beta_3 * WIricher + \beta_4 * WIrichest + \beta_5 * rural + \beta_6 * age + \beta_7 * employed + \beta_8 * EduPrim + \beta_9 * EduSec + \beta_{10} * EduHigh + \beta_{11} * Married + \beta_{12} * Cohabiting + \beta_{13} * Widowed + \beta_{14} * Divorced + \beta_{15} * Separated + \varepsilon$$

Since the dependent variable has four categories, we then observe the following possibilities, in which the **δ**'s are cut-off points that are estimated along with the other coefficients (Becker & Kennedy, 1992):

1. *Nutritional status = underweight if* $y^* \leq \boldsymbol{\delta_0}$
2. *Nutritional status = normal weight* $\boldsymbol{\delta_0} < y^* \leq \boldsymbol{\delta_1}$
3. *Nutritional status = overweight if* $\boldsymbol{\delta_1} < y^* \leq \boldsymbol{\delta_2}$
4. *Nutritional status = obese if* $\boldsymbol{\delta_2} < y^*$

In other words, the probability that an individual in our dataset is underweight is the probability that y* falls below the cutoff point between underweight and normal weight.

Finally, estimation in Ordered Probit Regression is an example of Generalized Linear Models that use Maximum Likelihood Estimation. This means that the linear function mentioned earlier is connected to the dependent variable with a link function. Therefore, the probability that a given observation falls into a certain category of the dependent variable is a nonlinear function of the predictors (Daykin & Moffat, 2002).

3.4. Ethical considerations

The DHS are free-to-use, public datasets, for which users must register and request access by explaining their intended use. This system ensures that all users who gain access to the data have an understanding of the ethical standards in place, and have agreed to these ethical standards.

These ethical considerations include not sharing the data with others without the consent of DHS and ensuring that no effort will be made to try to identify households or individuals from the sample data (The DHS Program, 2022).

4. Results

4.1. Descriptive statistics and missing values report

Generally, all surveys used a similar selection process where all eligible participants included women of reproductive ages. However, in Cameroon, a maximum cutoff point of 50 was used, while Lesotho and Zimbabwe used a cutoff point of 49 (The DHS Program, 2018; The DHS Program, 2014; The DHS Program, 2015). Since these selection processes were similar, and there were only 5 fifty-year-old participants in the Cameroonian sample, no additional sample selection was needed to run the ordered probit regression analyses.

Furthermore, while all surveys reported the Body Mass Index [BMI] of respondents, there are many missing values on this variable in the surveys for Cameroon and Lesotho. However, there are two reasons to assume that these missings aren't due to another variable and that we can treat them as missing completely at random [MCAR]. First of all, the remaining percentage distributions are quite similar to the data by Popkin, Corvalan & Grummer-Strawn (2020) from the most recent Unicef and NCD-Risc surveys, where Cameroon had 6% women's thinness and 43% overweight, and Lesotho had 4.3% women's thinness and 55.3% overweight. Second, the final columns for each of these countries in table 4 visualize the difference between the distributions for those respondents with a value for BMI and those respondents with missing values for BMI. These differences are so small that it is safe to assume that these missing values would not have influenced the analysis. Therefore, the default method of listwise deletion, where all respondents with a missing value on a variable are not included in the analysis, is unlikely to lead to biased results.

Additionally, the number of remaining respondents for each survey is high enough that the analyses are still representative of the total population.

4.1.1. Descriptive statistics of the general samples

Table 4 shows the general distributions of the background characteristics for female respondents in the Cameroon (2018), Lesotho (2014) & Zimbabwe (2015) Demographic and Health Surveys.

A few noteworthy observations about the total samples are that while in the Cameroonian sample almost a third of respondents (27.7%) received no education, this number was much lower for the Lesotho (2.3%) and Zimbabwe (2.3%) samples. While the distributions of people who had higher education were not as far apart (3.3%; 5.4% & 6.3%). This is because a large share of respondents in Lesotho had only primary education (57.1%) and an even larger share in Zimbabwe had achieved secondary education (57.7%), as opposed to a relatively equal distribution of no, primary and secondary education in Cameroon.

Second, the wealth distributions in all countries are somewhat similar, with no percentage distributions surpassing 25%, and only one percentage distribution dropping below 15% *(Richest, Cameroon).* The other groups are all somewhat evenly distributed around the 20% per group mark.

Third, the average age in all countries are also somewhat similar, with Lesotho being the highest *(μ = 35.31; SD = 8.36)*, followed by Cameroon *(μ = 34.64; SD = 8.11)* and Zimbabwe *(μ = 34.60; SD = 7.71)*.

Fourth, while the rural population is the mode or majority in all countries, the distributions still show large differences. With almost three-quarters of the population in Lesotho living in rural areas (74.4%), almost two-thirds of the Zimbabwe population (63.3%) living in rural areas. And only a little over half of the Cameroon population living in rural areas (56.3%).

Fifth, large differences are shown in the distribution of employed women in our countries, with almost three-quarters of the women in Cameroon being employed (74.9%), while most women in Lesotho (60.5%) and Zimbabwe (52.3%) are unemployed.

Sixth, the majority of women in each of the observed countries are married (64.0%; 74% & 78.2%). However, relatively speaking this number is quite a bit lower in Cameroon than in the others, while the number of women who are cohabiting with their partner without being married is much higher in Cameroon (17.0%) than in Lesotho (1.0%) and Zimbabwe (3.4%).

And finally, while all countries have low distributions of underweight women (5.2%; 2.1% & 4.0%), and all countries also have females with normal weight as the mode (51.7%; 43.2% & 50.5%). The total distribution of overweight and obese women is still quite a bit higher in Lesotho (28.1% + 26.6% = 54.7%) than in Cameroon (25.6% + 17.6% = 43.2%) and Zimbabwe (27.7% + 17.9% = 45.6%)

4.1.2. Descriptive statistics by nutritional status

Table 5 and figures 2, 3 and 4 show the distributions of the background characteristics for female respondents in the Cameroon (2018), Lesotho (2014) & Zimbabwe (2015) Demographic and Health Surveys by nutritional status.

First and foremost, these figures and the table with the distributions showcase that for each of the countries, BMI appears to increase as SES increases. Apart from a higher distribution of richest individuals in the underweight category for Lesotho.

Other similarities between all countries are that the obese and overweight categories have higher average ages than the normal weight and underweight categories. That the distribution of individuals in urban areas increases along with BMI, apart from a higher distribution of urban residents in the underweight category for Lesotho. That the distribution of employed women increases as BMI increases. And that the distribution of women who had secondary or higher education increases as BMI increases, apart from a higher distribution of secondary education in the underweight category for Lesotho.

The only major difference in their shifts in distribution is found in their changes in marital status. With the distribution of married individuals decreasing as BMI increases in Cameroon and increasing as BMI increases in Lesotho and Zimbabwe.

Table 4: Descriptive statistics of the variables

	Cameroon				***Lesotho***				***Zimbabwe***		
Variable	*Mean (SD) and % distributions*	*Mode*	*Total N*	*Missing values on BMI*	*Mean (SD) and % distributions*	*Mode*	*Total N*	*Missings on BMI*	*Mean (SD) and % distributions*	*Mode*	*Total N*
Socioeconomic Status (Scale, 5 items: {1=Poorest, 2=Poorer, 3=Middle, 4=Richer, 5=Richest)	18.7 % 24.4% 24.2% 18.4% 14.3%	2	33988	18.0% 24.9% 24.4% 17.2% 15.5%	23.7 % 19.1% 20.1% 19.7% 17.4%	2	11705	24.3% 17.7% 20.1% 19.1% 18.0%	20.3% 17.7% 17.1% 23.3% 21.6%	2	20774
Age	34.64 (8.11)	30	33988	34.67 (8.13)	35.31 (8.36)	34	11705	35.31 (8.35)	34.60 (7.71)	34	20774
Level of education (Scale, 4 items: 0=No education; 1=Primary education; 2=Secondary education; 3=Higher education)	27.7% 36.6% 32.4% 3.3%	1	33988	27.7% 36.0% 32.8% 34.4%	2.3% 57.1% 35.2% 5.4%	1	11705	2.4% 57.3% 35% 5.3%	2.3% 33.7% 57.7% 6.3%	1	20774
Type of place of residence (0=Urban; 1=Rural)	43.7% 56.3%	1	33988	43.8% 56.2%	25.6% 74.4%	1	11705	26.5% 73.5%	36.4% 63.6%	1	20774
Employment Status (0=No; 1=Yes)	25.1% 74.9%	1	33988	24.8% 75.2%	60.5% 39.5%	1	11705	61.9% 38.1%	52.3% 47.7%	1	20774
Marital Status (Scale, 6 items: 0=Never in union, 1=Married, 2=Living with partner, 3=Widowed, 4=Divorced, 5=Separated)	6.8% 64% 17% 5% 1.8% 5.4%	1	33988	6.9% 64.7% 16.4% 4.9% 1.9% 5.3%	6.2% 74% 1% 12.4% 1.8% 4.6%	1	11705	6.6% 73.9% 0.7% 12% 2.1% 4.8%	2.6% 78.2% 3.4% 6.8% 5.4% 3.6%	1	20774
Body Mass Index (Scale, 4 items: 0=Underweight, 1=normal weight, 2=overweight, 3=obese)	5.2% 51.7% 25.6% 17.6%	1	17583	-	2.1% 43.2% 28.1% 26.6%	1	6066	-	4.0% 50.5% 27.7% 17.9%	1	20199

Table 5: Descriptive statistics of the variables by nutritional status

	Cameroon				Lesotho				Zimbabwe			
Variable	Underweight	Normal weight	Overweight	Obese	Underweight	Normal weight	Overweight	Obese	Underweight	Normal weight	Overweight	Obese
Socioeconomic Status (Scale, 5 items: {Poorest, Poorer, Middle, Richer, Richest)	49.1% 26.7% 12.9% 9.6% 1.8%	26.7% 28.4% 23.1% 14.5% 7.2%	9.7 % 21.7% 27.2% 24.0% 17.5%	3.6 % 13.1% 25.1% 30.5% 27.7%	40.9 % 15.0% 15.0% 10.2% 18.9%	34.1% 22.7% 17.7% 15.5% 9.9%	12.9% 21.1% 22.6% 18.4% 19.0%	8.2% 16.4% 19.2% 30.5% 25.6%	33.5% 22.2% 18.8% 17.7% 7.7%	27.3% 21.1% 18.1% 20.0% 13.5%	16.1% 16.1% 18.5% 27.0% 22.2%	5.6% 10.8% 13.3% 27.2% 43.1%
Age	33.95 (7.80)	33.47 (8.24)	35.18 (7.90)	37.27 (7.21)	34.68 (10.07)	33.37 (8.51)	35.99 (8.11)	37.84 (7.41)	34.03 (8.40)	33.19 (7.93)	35.30 (7.12)	37.33 (6.71)
Level of education (Scale, 4 items: No education, Primary education, Secondary education, Higher education)	66.2% 20.9% 12.5% 0.3%	35.1% 36.0% 27.3% 1.6%	16.3% 42.1% 36.8% 4.8%	11.1% 38.2% 44.5% 6.3%	0% 64.6% 34.6% 0.8%	3.7% 62.4% 31.1% 2.9%	1.6% 57.1% 34.9% 6.3%	0.6% 48.0% 43.1% 8.8%	2.5% 41.5% 54.6% 1.4%	2.6% 39.0% 55.0% 3.5%	2.2% 31.7% 59.8% 6.3%	1.1% 23.9% 62.3% 13.8%
Type of place of residence (0 = Urban; 1 = Rural)	25.4% 74.6%	33.6% 66.4%	49.9% 50.1%	68.9% 31.3%	26.0% 74.0%	20.7% 79.3%	25.1% 74.9%	30.6% 69.4%	22.2% 77.8%	26.9% 73.1%	38.6% 61.4%	58.0% 42.0%
Employment Status (0 = No; 1 = Yes)	39.9% 60.1%	24.1% 75.9%	25.1% 74.9%	24.7% 75.3%	70.9% 29.1%	65.8% 34.2%	56.9% 43.1%	49.7% 50.3%	64.3% 35.7%	57.2% 42.8%	50.5% 49.5%	39.9% 60.1%
Marital Status (Scale, 6 items: Never in union, Married, Living with partner, Widowed, Divorced, Separated)	1.5% 80.9% 8.3% 3.6% 4.1% 1.7%	6.2% 63.5% 18.1% 5% 1.5% 5.7%	8.6% 59.7% 18.7% 5.1% 1.9% 6.1%	6.7% 62.8% 16.6% 6.5% 1.5% 5.8%	10.2% 57.5% 5.5 16.5% 2.4% 7.9%	7.6% 71.4% 1.5% 13.0% 1.6% 4.9%	5.4% 74.8% 0.9% 13.4% 1.4% 4.1%	3.4% 78.8% 1.1% 11.3% 1.6% 3.9%	5.4% 67.5% 3.5% 11.0% 9.4% 3.4%	3.4% 75.9% 4.2% 6.8% 5.4% 4.4%	1.8% 80.5% 2.8% 6.1% 5.5% 3.4%	1.3% 84.0% 1.8% 6.3% 4.3% 2.3%

Figure 2: SES by nutritional status - Cameroon

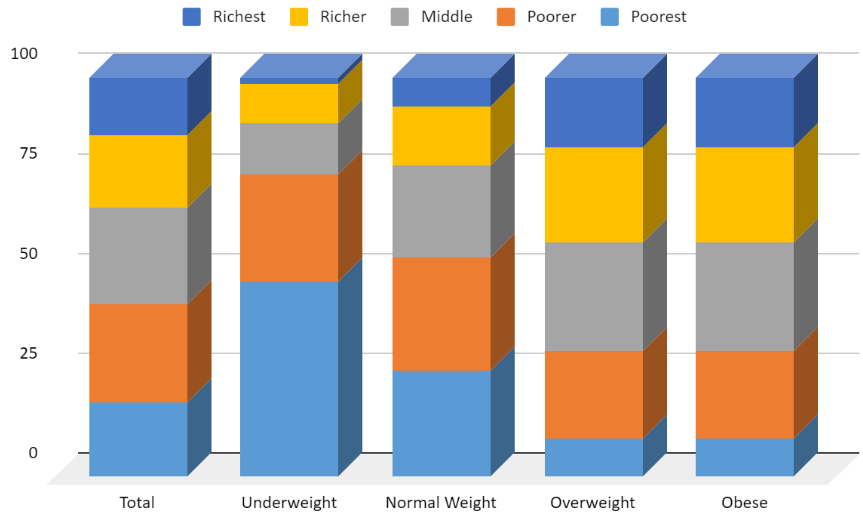


Figure 3: SES by nutritional status - Lesotho

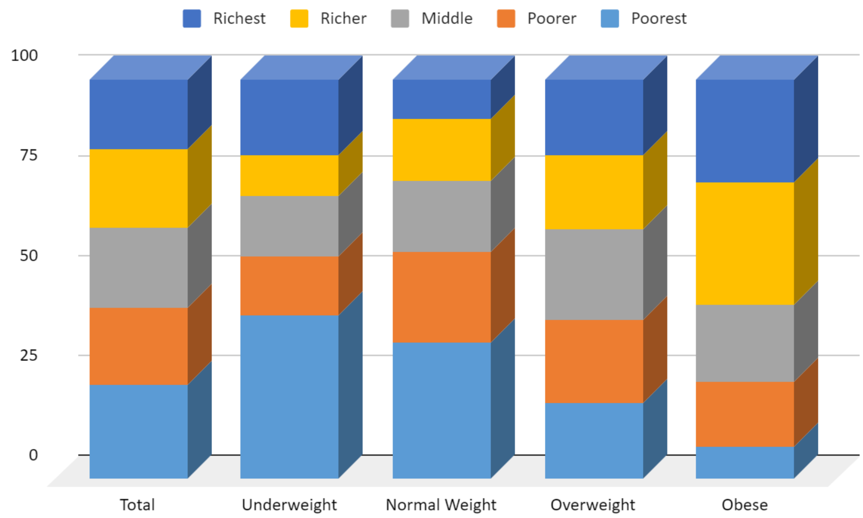
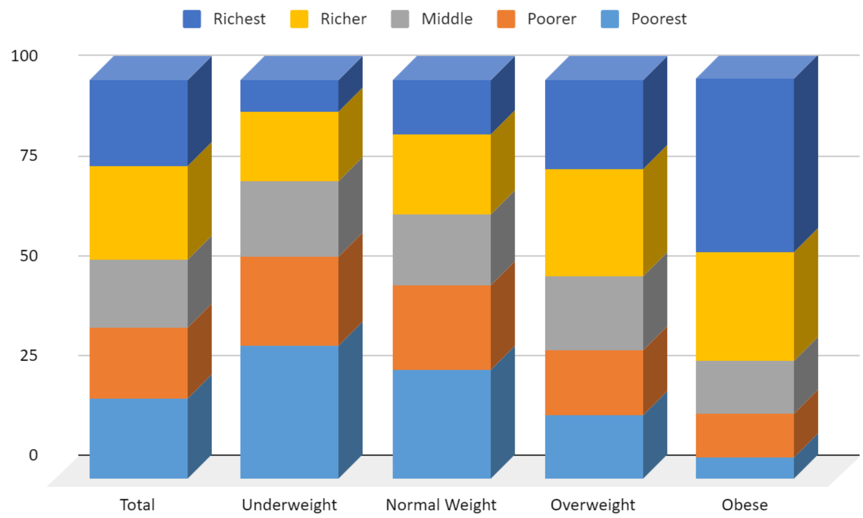


Figure 4: SES by nutritional status - Zimbabwe



4.1.3. Correlations

4.1.3.1. Cameroon

The spearman correlation coefficients for the Cameroon Demographic and Health Surveys (2018) are tabulated in table 6. Most noteworthy is that all variables are significantly associated with each other at the 0.01 alpha-level, except for age and place of residence. Meaning that there is no co-linearity between the age of the respondents and their type of place of residence.

Furthermore, the independent variables SES *(χ^2 = 0.393),* level of education *(χ^2 = 0.292),* age *(χ^2 = 0.159),* and employment status *(χ^2 = 0.025),* are all positively correlated with the dependent variable. Indicating that when either variable increases, the other variable increases in the same direction. The only negative correlation with the dependent variable was found for type of place of residence *(χ^2 = - 0.273),* indicating that BMI is lower in rural areas than in urban areas.

Table 6: Correlations of the continuous variables (Cameroon)

	SES	*Education*	*Place of Residence*	*Age*	*Employment Status*	*BMI*
SES	-	-	-	-	-	-
Education	0.543*	-	-	-	-	-
Place of residence	-0.638*	-0.315*	-	-	-	-
Age	0,025*	-0.108*	-0.009	-	-	-
Employment Status	-0,094*	0.057*	0.095*	0.145*	-	-
BMI	0,393*	0.292*	-0.273*	0.159*	0.025*	-

** Statistically significant at the 0.01 level*

4.1.3.2. Lesotho

The spearman correlation coefficients of the Lesotho Demographic and Health Surveys (2014) are tabulated in table 7. As visualized in this table, all correlation coefficients are statistically significant at the 0.01 alpha-level, apart from the correlation coefficient between age and place of residence, and employment status and place of residence. Meaning that there is no co-linearity between the age of the respondent and their place of residence, and between employment status and place of residence.

Furthermore, the independent variables SES *(χ^2 = 0.305),* level of education *(χ^2 = 0.164),* age *(χ^2 = 0.212),* and employment status *(χ^2 = 0.140),* are all positively correlated with the dependent variable. Indicating that when either variable increases, the other variable increases in the same direction. The only negative correlation with the dependent variable was found for type of place of residence *(χ^2 = - 0.088),* indicating that BMI is lower in rural areas than in urban areas.

Table 7: Correlations of the continuous variables (Lesotho)

	SES	*Education*	*Place of Residence*	*Age*	*Employment Status*	*BMI*
SES	1	-	-	-	-	-
Education	0.458*	1	-	-	-	-
Place of residence	-0.543*	-0.266*	1	-	-	-
Age	0.085*	-0.123*	0.019	1	-	-
Employment Status	0.316*	0.201*	-0.255	0.132*	1	-
BMI	0.305*	0.164*	-0.088*	0.212*	0.140*	1

** Statistically significant at the 0.01 level*

4.1.3.3. Zimbabwe

The spearman correlation coefficients of the Zimbabwe Demographic and Health Surveys (2014) are tabulated in table 8. As visualized in this table, all correlation coefficients are statistically significant at the 0.01 alpha-level.

Furthermore, the independent variables SES *($\chi^2 = 0.316$),* level of education *($\chi^2 = 0.168$),* age *($\chi^2 = 0.189$),* and employment status *($\chi^2 = 0.132$),* are all positively correlated. Indicating that when either variable increases, the other variable increases in the same direction. The only negative correlation with the dependent variable was found for type of place of residence *($\chi^2 = - 0.234$),* indicating that BMI is lower in rural areas than in urban areas.

Table 8: Correlations of the continuous variables (Zimbabwe)

	SES	*Education*	*Place of Residence*	*Age*	*Employment Status*	*BMI*
SES	1	-	-	-	-	-
Education	0.466*	1	-	-	-	-
Place of residence	-0.782*	-0.382*	1	-	-	-
Age	0.023*	-0.079*	0.024*	1	-	-
Employment Status	0.250*	0.158*	-0.199*	0.111*	1	-
BMI	0.315*	0.167*	-0.234*	0.189*	0.131*	1

** Statistically significant at the 0.01 level*

4.2. Model evaluation

All ordered probit regression models in this section were executed using forward selection methods, which start with an empty model and add in one variable at a time in order of highest improvement to the model (Sutter & Kalivas, 1993)

This order of addition is the order in which these variables are reported in the regression tables, meaning that predictors at the top provide the best explanation for the sample differences in nutritional status, and the bottom predictors provide the worst explanation.

Additionally, all predictors that had a p-value of > 0.1 were omitted from the regression model completely.

Furthermore, all tables and figures with predicted probabilities were calculated by holding all predictors apart from SES ceteris paribus at their mean.

4.2.1. Regression results for Cameroon

Table 9 with the Ordered Probit Regression coefficients and table 10 (and figure 5) with the predicted probabilities for nutritional status by socioeconomic status clearly visualize that the probability for individuals to be overweight or obese increases linearly as the SES of the respondents increases, and that the probability for individuals to be underweight or normal weight decreases linearly as the SES of the respondents increases.

However, SES is not the most important predictor of nutritional status in Cameroon. More importantly, rural residents have a lower propensity of falling into higher categories of nutritional status *($B = -0.136; p < 0.001$),* and older residents have a higher propensity of falling into higher categories of nutritional status *($B = 0.024; p < 0.001$)*

Furthermore, the relationship between educational status and nutritional status is non-linear, since the coefficient for primary education *($B = 0.508; p < 0.001$)* is higher than the coefficient for secondary education *($B = 0.486; p < 0.001$).* Indicating that the propensity for individuals to belong to higher categories of nutritional status is higher for individuals who have only completed primary education than for people who have also completed secondary education.

And finally, the weakest significant effects were found for employment status *($B = 0.058; p = 0.004$),* indicating that employed individuals have a higher propensity of falling into higher categories of nutritional status. And for widowed respondents in comparison to the reference group (Marital Status: Never in Union), indicating that widowed respondents were more likely to fall into higher categories of nutritional status.

While many significant effects of our predictors on the dependent variable were found using this ordered probit method, the explained variance was still very low *(Pseudo $R^2 = 0.099$)*, indicating that only 9.9% of the variance in nutritional status is explained using this set of predictors. So while this may still be interpreted as a significant improvement over the null model with only the dependent variable *($p < 0.001$),* it is reasonable to assume that this model is not able to capture the full complexity of the differences in nutritional status.

Table 9: Ordered Probit Regression Results for Cameroon

	B(SE)	*P*
Place of residence	-0.136 (0.022)	<0.001*
Age	0.024 (0.001)	<0.001*
Wealth Index		
Richest	1.065 (0.042)	<0.001*
Richer	0.844 (0.036)	<0.001*
Middle	0.618 (0.030)	<0.001*
Poorer	0.327 (0.028)	<0.001*
Level of education		
Primary	0.508 (0.024)	<0.001*
Secondary	0.486 (0.027)	<0.001*
Higher	0.607 (0.055)	<0.001*
Employment Status	0.058 (0.020)	0.004*
Marital status		
Widowed	0.099 (0.039)	0.011*
Underweight | Normal Weight		-0.166 (0.049)
Normal Weight | Overweight		1.916 (0.050)
Overweight | Obese		2.792 (0.052)
Pseudo R^2		0.099
Total N		17583
Wald χ^2		3980.51
P		<0.001**

* Significant at $P < 0.05$

Reference categories: Poorest, No Education & Never in union

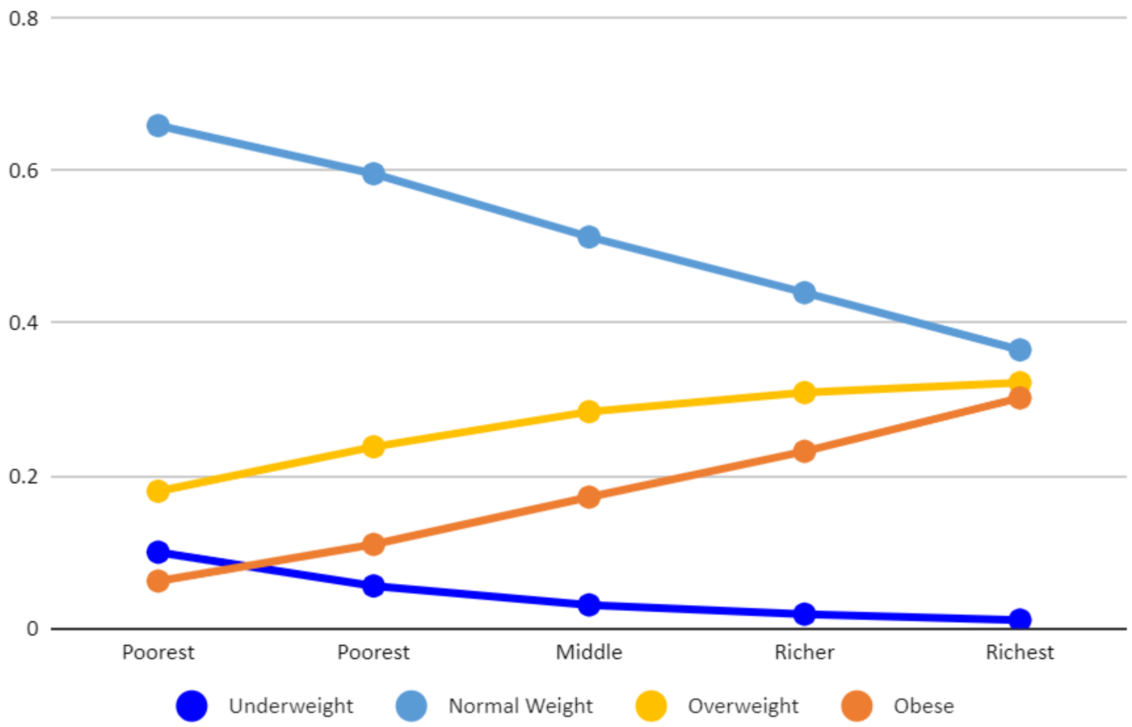
Omitted predictors: Married, Living with partner, Divorced & Separated

Table 10: Predicted Probabilities for nutritional status (Cameroon)

	Poorest	*Poorer*	*Middle*	*Richer*	*Richest*
Underweight	0.100	0.056	0.031	0.019	0.011
Normal Weight	0.659	0.596	0.513	0.440	0.365
Overweight	0.180	0.238	0.284	0.309	0.322
Obese	0.062	0.110	0.172	0.232	0.302

**All predicted probabilities are statistically significant at the 0.01 level*
**All other predictors were held ceteris paribus at their mean*

Figure 5: Predicted probabilities by nutritional status - Cameroon



4.2.2. Regression results for Lesotho

Table 11 with the Ordered Probit Regression coefficients shows a stepwise increase in nutritional status as SES increases. However, unlike the predicted probabilities for Cameroon, the predicted probabilities for Lesotho in table 12 do not show a linear increase or decrease for all categories of nutritional status.

While both the underweight and normal weight categories show a gradual decrease, and the obese category shows a gradual increase as SES increases. The same cannot be said for the overweight category. This category first increases from poorest *(p^ = 0.231)* to poorer *(p^ = 0.292)* to middle *(p^ = 0.303)* to richer *(p^ = 0.309)* before decreasing to richest *(p^ = 0.307)*. Indicating that the probability for an individual to be overweight (but not obese) is highest for the richer SES category. However, this appears to be mainly due to the high probability to be obese, since the probability of the richest category to be overweight or obese (0.690) is much higher than their probability to be underweight or normal weight (0.308)

Furthermore, where place of residence was the most important predictor for Cameroon, indicating that rural citizens were less likely to fall in the higher categories of nutritional status. For Lesotho, place of residence *(B = 0.209; p < 0.001)* is only the sixth predictor and actually shows the opposite effect, where living in a rural place makes it more likely to fall in the higher categories of nutritional status. This negative coefficient is very counterintuitive when looking back at the descriptive statistics in table 5, and the correlation coefficient in table 7. This most likely indicates that there is omitted variable bias, where the effect of type of place of residence on nutritional status can be explained through a variable that was not included in the model. This is something to keep in mind when interpreting the results.

Most importantly, the best predictor of nutritional status in Lesotho was the predictor of age *(B = 0.028; p < 0.001)*. Indicating that nutritional status increases as age increases, which is an almost identical effect to the effect in Cameroon, where age was the second best predictor of nutritional status.

And finally, there is a stepwise positive increase between educational status and nutritional status. Indicating that the propensity for individuals to belong to higher categories of nutritional status is higher for individuals with higher levels of education. As could be expected following the descriptive analysis.

While many significant effects of our predictors on the dependent variable were found using this ordered probit method, the explained variance was still very low *(Pseudo $R^2 = 0.070$)*, indicating that only 7.0% of the variance in nutritional status is explained using this set of predictors. So while this may still be interpreted as a significant improvement over the null model with only the dependent variable *($p < 0.001$),* it is reasonable to assume that this model is not able to capture the full complexity of the differences in nutritional status.

Table 11: Ordered Probit Regression Results for Lesotho

	B(SE)	*P*
Age	0.028 (0.002)	<0.001*
Wealth Index		
Richest	0.929 (0.060)	<0.001*
Richer	0.871 (0.051)	<0.001*
Middle	0.593 (0.047)	<0.001*
Poorer	0.463 (0.045)	<0.001*
Marital Status		
Married	0.276 (0.034)	<0.001*
Place of residence	0.209 (0.041)	<0.001*
Level of education		
Higher	0.695 (0.123)	<0.001*
Secondary	0.498 (0.106)	<0.001*
Primary	0.368 (0.102)	<0.001*
Employment Status	0.081 (0.032)	0.011*
Marital Status		
Living with partner	-0.268 (0.130)	0.039*
Underweight | Normal Weight		0.131 (0.131)
Normal Weight | Overweight		2.216 (0.132)
Overweight | Obese		3.045 (0.133)
Pseudo R^2		0.070
Total N		6066
Wald χ^2		972.91
P		<0.001**

* Significant at $P < 0.05$

Reference categories: Poorest, No Education & Never in union

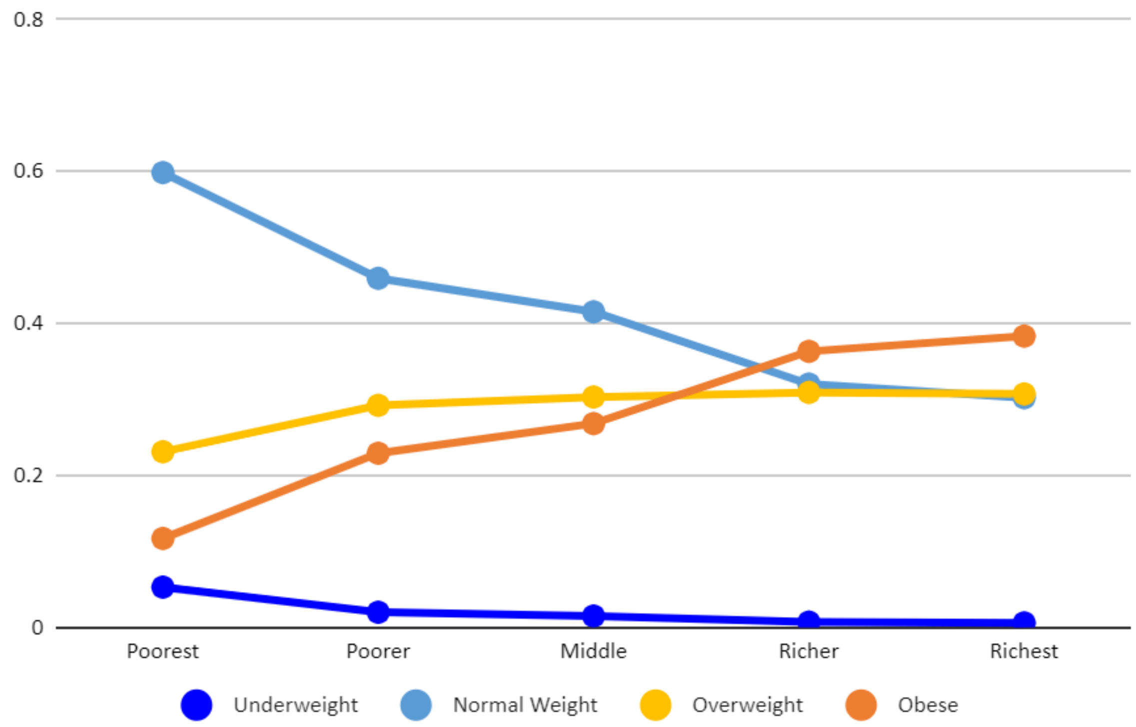
Omitted predictors: Widowed, Divorced & Separated

Table 12: Predicted probabilities of nutritional status (Lesotho)

	Poorest	*Poorer*	*Middle*	*Richer*	*Richest*
Underweight	0.053	0.020	0.015	0.007	0.006
Normal Weight	0.598	0.459	0.415	0.320	0.302
Overweight	0.231	0.292	0.303	0.309	0.307
Obese	0.117	0.229	0.268	0.363	0.383

**All predicted probabilities are statistically significant at the 0.01 level*
**All other predictors were held ceteris paribus at their mean*

Figure 6: Predicted probabilities by nutritional status - Lesotho



4.2.3. Regression results for Zimbabwe

Table 13 with the Ordered Probit Regression coefficients for Zimbabwe and table 14 (and figure 7) with the predicted probabilities for nutritional status by SES clearly visualize that the probability for individuals to be overweight or obese increases stepwise as the SES of the respondents increases, and that the probability for individuals to be underweight or normal weight decreases stepwise as the SES of the respondents increases. As was expected following the descriptive analysis. Additionally, in this analysis, the Richest category is even the best predictor for nutritional status among all used predictors.

Furthermore, similarly to the previous two models, age *(B = 0.028; p < 0.001)* is one of the most important predictors of nutritional status. Indicating once again that an increase in age is accompanied by an increase in BMI.

Likewise, both married respondents *(B = 0.312; p < 0.001)* and respondents who are living with their partner *(B = 0.092; p = 0.052)* have a higher propensity of belonging to the higher groups of nutritional status. Although only the married category is significant at the 0.05 alpha-level.

But yet again, the effect of type of place of residence *(B = 0.108; 0.022)* is opposite of that in Cameroon, where it was the most important (negative) predictor. This negative coefficient is very counterintuitive when looking back at the descriptive statistics in table 12, and the correlation coefficient in table 13. This most likely indicates that there is omitted variable bias, where the effect of type of place of residence on nutritional status can be explained through a variable that was not included in the model. This is something to keep in mind when interpreting the results.

Finally, the absence of primary education is noteworthy, since each level of education has been a significant predictor in both previous models.

While many significant effects of our predictors on the dependent variable were found using this ordered probit method, the explained variance was still very low *(Pseudo $R^2 = 0.069$)*, indicating that only 6.9% of the variance in nutritional status is explained using this set of predictors. So while this may still be interpreted as a significant improvement over the null model with only the dependent variable *($p < 0.001$)*, it is reasonable to assume that this model is not able to capture the full complexity of the differences in nutritional status.

Table 13: Ordered Probit Regression Results for Zimbabwe

	B(SE)	P
Wealth Index		
Richest	1.046 (0.042)	<0.001*
Age	0.028 (0.001)	<0.001*
Wealth Index		
Richer	0.721 (0.033)	<0.001*
Marital Status		
Married	0.310 (0.021)	<0.001*
Wealth Index		
Middle	0.344 (0.026)	<0.001*
Poorer	0.216 (0.026)	<0.001*
Employment Status	0.104 (0.017)	<0.001*
Level of education		
Higher	0.258 (0.039)	<0.001*
Place of residence	0.109 (0.032)	0.001*
Level of education		
Secondary	0.055 (0.018)	0.002*
Marital Status		
Living with partner	0.092 (0.047)	0.052
Underweight | Normal Weight		-0.049 (0.055)
Normal Weight | Overweight		1.969 (0.055)
Overweight | Obese		2.874 (0.057)
Pseudo R^2		0.069
Total N		20182
Wald χ^2		3169.08
P		<0.001**

* Significant at P < 0.05

Reference categories: Poorest, No Education & Never in union

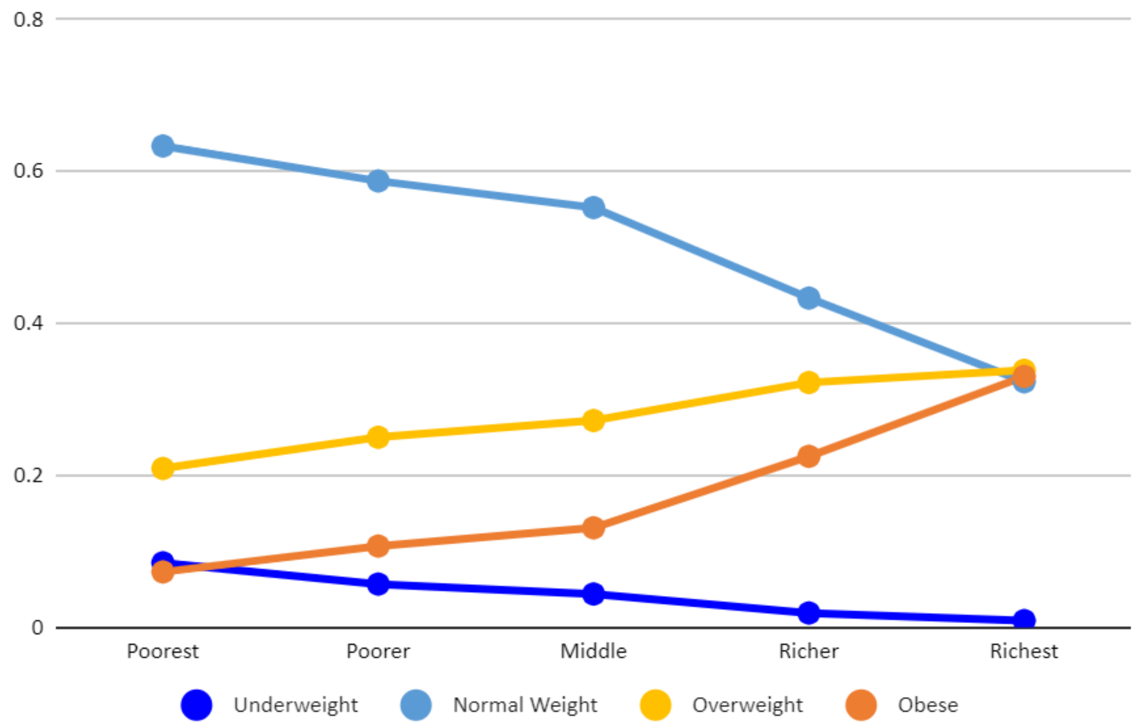
Omitted predictors: Widowed, Divorced, Separated & Primary education

Table 14: Predicted probabilities of nutritional status (Zimbabwe)

	Poorest	*Poorer*	*Middle*	*Richer*	*Richest*
Underweight	0.085	0.057	0.044	0.019	0.009
Normal Weight	0.633	0.587	0.552	0.433	0.323
Overweight	0.209	0.250	0.272	0.322	0.338
Obese	0.073	0.107	0.131	0.225	0.330

**All predicted probabilities are statistically significant at the 0.01 level*
**All other predictors were held ceteris paribus at their mean*

Figure 7: Predicted probabilities by nutritional status - Zimbabwe



4.2.4. Regression results for Cross-country comparison

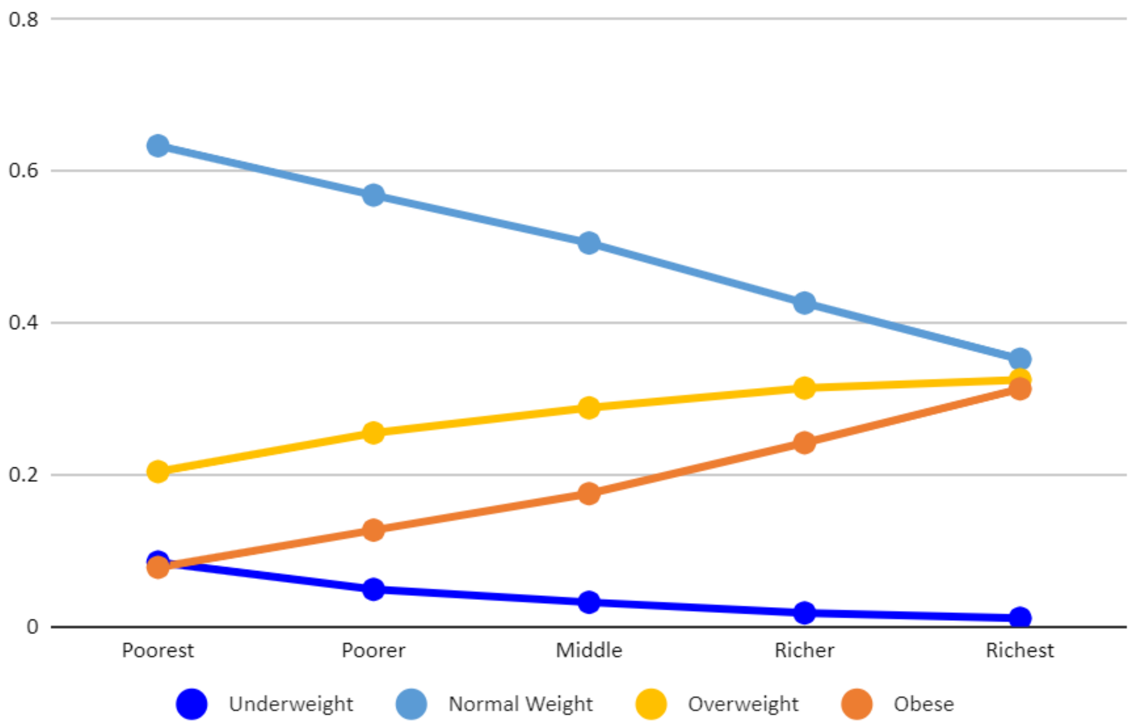
The three Ordered Probit Regression models for each of the countries of observation already showed a lot of similarities, and some differences, about which conclusions can be drawn. Nevertheless, to really figure out what the effects are of SES on nutritional status in countries dealing with the double burden of (mal)nutrition, and to see if the samples show significant differences on the dependent variable, this research paper also ran an Ordered Probit Regression model for all countries combined.

Important to note here is that when examining the results, it must be taken into account that the relative weight of each survey is proportional to their sample sizes, meaning that for instance the survey of Zimbabwe has 20199/6071 = 3.33 times as much predictive power as the survey of Lesotho and 20199/17583 = 1.15 times as much predictive power as the survey of Cameroon. Moreover, these surveys were held in 2018 (Cameroon), 2015 (Zimbabwe), and 2014 (Lesotho), therefore the reference time is a bit vague, causing an ill-defined sample population. Regardless, the pooling of these samples in itself is unproblematic, since they use the same set of predictors.

First, table 15 showcases the main results of the combined Ordered Probit Regression model, and table 16 (and figure 8) visualize the predicted probabilities that are calculated from these main results.

What these results show is that like in all the separate models, there is a stepwise increase of probability to fall into the higher categories of nutritional status for each stepwise increase in SES. The smallest coefficient is found for Poorer *($B = 0.289$; $p < 0.001$)*, followed by Middle *($B = 0.503$; $p < 0.001$)*, Richer *($B = 0.746$; $p < 0.001$)*, and finally Richest *($B = 0.969$; $p < 0.001$)*. When transforming these coefficients into predicted probabilities, the linearity of the probabilities is easily visible.

Figure 8: Predicted probabilities for nutritional status in all countries combined



Likewise, the effects of age *(B = 0.027; p < 0.001)* and employment status *(B = 0.063; p < 0.001)* also showed effects in a similar direction to the separate models, where significantly positive coefficients were also always found for these predictors.

However, there were also some noticeable differences. First of all, a significantly negative coefficient for Marital Status: Divorced *(B = -0.082; p = 0.011)* was found, while this predictor had not previously given significant results for any of the separate analyses.

Place of residence yielded a significantly negative coefficient *(B = -0.040; p = 0.013),* even though only the Cameroonian results gave a negative coefficient previously. Although the positive result that was found for Zimbabwe and Lesotho was rather counterintuitive, this is still remarkable.

And finally, while primary education as a predictor was absent for Zimbabwe, and had a higher coefficient than the predictor for secondary education in Cameroon, the results for the combined regression showcased a stepwise increase in the propensity of falling into higher categories of nutritional status for each increase in education. With the coefficient for Primary *(B = 0.493; p < 0.001)* being the lowest, Secondary *(B = 0.526; p < 0.001)* being second and Higher *(B = 0.715; p < 0.001)* being the highest.

Second, table 15 also shows the differences in Nutritional Status between Cameroon, Lesotho and Zimbabwe (as well as their order of admission into the model). What these coefficients show is that individuals in Lesotho are most likely to fall in higher categories of nutritional status, with Cameroon *(B = -0.221; p < 0.001)* and Zimbabwe *(B = -0.329; p < 0.001)* having a significantly negative coefficient in comparison to the reference category Lesotho.

Furthermore, out of Zimbabwe and Cameroon, Zimbabwe appears to score lower on nutritional status, having a significantly negative coefficient *(B = -0.107; p < 0.001)* in comparison to the reference category of Cameroon. However, these differences in nutritional status do not appear to have caused drastically different results between the different analyses.

While many significant effects of our predictors on the dependent variable were found using this ordered probit method, the explained variance was still very low *(Pseudo R^2 = 0.080)*, indicating that only 8% of the variance in nutritional status is explained using this set of predictors. So while this may still be interpreted as a significant improvement over the null model with only the dependent variable *($p < 0.001$),* it is reasonable to assume that this model is not able to capture the full complexity of the differences in nutritional status.

Table 15: Ordered Probit Regression Results for cross-country comparison

	B(SE)	*P*	*Order of admission*
Wealth Index			
Richest	0.969 (0.025)	<0.001**	
Richer	0.746 (0.021)	<0.001**	
Age	0.027 (0.001)	<0.001**	
Wealth Index			
Middle	0.503 (0.018)	<0.001**	
Poorer	0.289 (0.017)	<0.001**	
Level of education			
Higher	0.717 (0.033)	<0.001**	
Primary	0.493 (0.020)	<0.001**	
Secondary	0.525 (0.021)	<0.001**	
Marital Status			
Married	0.159 (0.015)	<0.001*	
Employment Status	0.062 (0.012)	<0.001**	
Marital Status			
Living with partner	0.097 (0.023)	<0.001**	
Divorced	-0.082 (0.032)	0.011*	
Place of residence	-0.040 (0.016)	0.014*	
Reference category: Cameroon			
Lesotho	0.221 (0.019)	<0.001**	6
Zimbabwe	-0.107 (0.014)	<0.001**	10
Reference category: Lesotho			
Cameroon	-0.221 (0.019)	<0.001**	9
Zimbabwe	-0.329 (0.017)	<0.001**	11
Reference category: Zimbabwe			
Cameroon	0.107 (0.014)	<0.001**	6
Lesotho	0.329 (0.017)	<0.001**	10
Underweight | Normal Weight		0.065 (0.036)	
Normal Weight | Overweight		2.11 (0.037)	
Overweight | Obese		2.978 (0.376)	
Pseudo R^2		0.080	
Total N		43831	
Wald χ^2		8036.86	
P		<0.001**	

**Significant at P <0.01; * Significant at P < 0.05

Reference categories: Poorest, No Education & Never in union

Omitted variables: Separated

Table 16: Predicted probabilities of nutritional status (Combined)

	Poorest	*Poorer*	*Middle*	*Richer*	*Richest*
Underweight	0.085	0.049	0.032	0.018	0.011
Normal Weight	0.633	0.568	0.505	0.426	0.352
Overweight	0.204	0.255	0.288	0.314	0.325
Obese	0.078	0.127	0.175	0.242	0.313

**All predicted probabilities are statistically significant at the 0.01 level*
**All other predictors were held ceteris paribus at their mean*

4.3. Hypothesis testing

As previously stated in Chapter 2: Theoretical Framework, there are three hypotheses that this research paper aims to test.

H1: There is a relationship between socioeconomic status and nutritional status in countries dealing with the double burden of (mal)nutrition when controlling for other resources, conversion factors, and capabilities.

Based on the results from the ordered probit regressions for Cameroon, Lesotho and Zimbabwe, as well as the results from the combined regression analysis. There is enough evidence to reject the null hypothesis stating that there is no relationship between socioeconomic status and nutritional status. Therefore, we can conclude that the first hypothesis is supported by the results of the analyses.

H2: The relationship between socioeconomic status and nutritional status in countries dealing with the double burden of (mal)nutrition when controlling for other resources, conversion factors, and capabilities is nonlinear.

Based on the results from the ordered probit regressions there is no indication that the relationship between socioeconomic status and nutritional status is nonlinear. In all three countries, as well as in the cross-country analysis, a stepwise increase in BMI was visible for each stepwise increase in SES. The only possible example of nonlinearity that was found was that of the predicted probabilities to be overweight in Lesotho. However, this was likely caused by the large distribution of obese people in the richest category, since the overweight and obese categories combined still had a much larger probability than the normal weight and the underweight categories. Therefore, we can conclude that the second hypothesis is not supported by the analyses, and may be rejected.

H3: The same relationship between socioeconomic status and nutritional status in countries dealing with the double burden of (mal)nutrition when controlling for other resources, conversion factors, and capabilities is found in Cameroon, Lesotho and Zimbabwe.

Every single one of the separate Ordered Probit Regression analyses found a stepwise increase in BMI for each stepwise increase in SES. Additionally, the combined model also found this same effect. Therefore, we can conclude that the third hypothesis is supported by the results of the analyses.

5. Discussion and conclusion

5.1. Summary of findings

In conclusion, with only two of the three proposed hypotheses being supported, the answer to the research question *"What are the effects of socioeconomic status on nutritional status for women in Sub-Saharan African countries dealing with the double burden of (mal)nutrition?"* partly contradicts the expectations described in the first two chapters.

According to our empirical findings, the relationship of socioeconomic status on nutritional status for women in Sub-Saharan African countries dealing with the double burden of (mal)nutrition is linear, with the propensity to fall into higher categories of nutritional status increasing for each stepwise increase in SES.

But while this direction of linearity is not unheard of in academic research, the theory suggested that it belonged to countries with lower levels of development, while rising obesity levels in Cameroon, Lesotho and Zimbabwe suggested that they already belonged to countries in the second stage of the developmental transition (Jaacks et al., 2019).

5.1.1. Reflection on literature

While the first two chapters of this research provided a literary basis from which the unsupported nonlinearity hypothesis was derived, the fact that it was not supported does not indicate that this theoretical framework is necessarily false, but rather that this paper might have misjudged the developmental stages in Cameroon, Lesotho and Zimbabwe.

Since while there is a gap in research on the relationship between socioeconomic status and nutritional status in developing countries that this paper tried to bridge, the beginning and final stages of this relationship have been thoroughly documented.

The beginning stages are characterized by higher overweight and obesity rates among those with a higher SES, decreasing stepwise as SES decreases in economically underdeveloped countries (Jaacks et al., 2019).

The final stages are characterized by higher overweight and obesity rates among those with a lower SES, decreasing stepwise as SES increases (Devaux & Sassi, 2011; Jaacks et al., 2019; Monteiro, Moura, Conde & Popkin, 2004).

And the middle phase, which is central to this research paper, is characterized by disappearing relationships (Jaacks et al., 2019).

When taking the beginning and end stages as certainties, somewhere in between during this development, the linear effect has to disappear or change. And while this change could be much more gradual than this paper posited with its nonlinearity stance, the linear relationship simply cannot completely change direction overnight.

Therefore, unless Cameroon, Lesotho and Zimbabwe are an exception to the existing body of knowledge, where regardless of their developmental status, higher SES groups also have higher BMI's. They are still in the earliest developmental stages of the obesity transition.

In conclusion, either this paper was correct in concluding that these countries are in similar developmental stages as those in which disappearing relationships were found, but this phenomenon does not occur at the same developmental stage in each society. Or this paper was incorrect in concluding that Cameroon, Lesotho and Zimbabwe were in similar developmental stages, and they had not yet seen enough development at the time of data collection for this phenomenon to occur.

5.1.2. Reflection on the conceptual model

While conclusions can be drawn about the effects of the observed resources and conversion factors, very little can be said about the agency and capabilities of the individuals. Since these could not be operationalized in the analysis. This incomplete operationalization should be kept in mind when interpreting the following conclusions about the conceptual model.

First of all, the Wealth Index that was constructed by the DHS converted the resource of income into higher categories of nutritional status. Each stepwise increase on the scale of the Wealth Index, from poorest to richest, meant an increase in the propensity to fall into higher categories of the scale of Nutritional Status. This conclusion can not only be drawn from the full model which combined all countries under investigation, but also from each of the individual analyses.

Second, as individuals get older, they tend to convert their resources into higher categories of nutritional status. Each year's increase in age led to an increase in the propensity to fall into higher categories of nutritional status. And like the Wealth Index, this conclusion can be drawn from all four of the models in this research paper.

Third, the full model as well as the model for the Cameroonian population, showed that a rural place of residence as opposed to an urban place of residence tended to convert resources into lower categories of nutritional status. However, the opposite conclusion had to be drawn based on the Lesotho and Zimbabwe models. So while the expectation is that these opposite effects are caused by omitted variable bias in the model, it is still not possible to draw a definitive conclusion about the conversion factor of type of place of residence.

Fourth, the full model, as well as the model for Lesotho, showed a stepwise increase in nutritional status for the resource of education. With each increase in the level of education increasing the propensity to fall into higher categories of nutritional status. And while this relationship was non-linear in Cameroon and the primary education category had no effect in Zimbabwe, we can still conclude that in general, having received education increases your likelihood of falling into higher categories of nutritional status in comparison to having received no education.

Fifth, various different conclusions about the conversion factor of marital status can be drawn based on the different analyses. But the full model showed a higher propensity of falling into higher categories of the functioning nutritional status for married individuals and cohabiting individuals, and a lower propensity of falling into higher categories for divorced individuals.

And finally, the country of origin of the respondents had an effect on the functioning of nutritional status in the sense that the cross-country analysis showed significant differences between all three countries. With the highest scores of nutritional status being found in Lesotho, followed by Cameroon and finally by Zimbabwe.

5.2. Strengths and limitations

This research paper simultaneously has strengths that made answering the research question and testing the hypotheses possible, as well as limitations which force the reader to be cautious about the interpretation of the conclusions that were drawn.

Starting on a positive note, the analysis found similar results across the different countries involved. And while this was to be expected after concluding in the introduction that all of the observed countries were in comparable situations regarding the double burden of (mal)nutrition and development, it still increases the credibility of the model that was used and decreases the odds that the linear relationship that was found differs from the situation in the actual populations.

Furthermore, an important strength of the DHS program was that it ensured similar questionnaires and survey procedures, which not only made it possible to build one theoretical model and apply it to all three countries involved. But also made it possible to do a cross-country analysis with all countries in a single model. However, in the interpretation of this cross-country analysis, it must be taken into account that the different countries had different weights on the final results and that the samples were from different years, causing an ill-defined population.

As a third positive, the forward selection method was useful both in providing additional information in the form of the relative importance of the added variables, through their order of admission. As well as in improving the accuracy of the estimation coefficients by only including those variables that had a significant effect of $p < 0.01$. In doing so, non-significant effects do not unnecessarily distort the estimation of the predicted probabilities. However, forward selection methods do come with some risks. As the possibility exists that the inclusion of a new variable makes a previously added variable non-significant (Chowdhury & Turin, 2020).

On a critical note, an important limitation was found in the link between the CA and the Demographic and Health Surveys datasets. Since the DHS does not allow for a full implementation of the conceptual model due to its inability to operationalize agency and capabilities. Therefore, while this paper was able to analyze how resources were converted into functionings, it was not able to account for an individual's agency or ability to pursue this functioning.

And finally, all four models had very low predictive powers with a pseudo R^2 of below 0.1. This indicates that while the models were an improvement over the null model with only the dependent variable, they were not able to capture the full complexity of what determines nutritional status. Another example of this lack of predictive power was visible in the sign of the coefficient for type of place of residence in Lesotho and Zimbabwe, which was counterintuitive when looking at the descriptive statistics and spearman correlations. This likely indicated omitted variable bias, where an important predictor of nutritional status that was not included in the model influenced the sign of type of place of residence.

5.3. Policy and practise recommendations

Based on the findings in this paper, some recommendations can be given for future researchers as well as for policy-makers in countries dealing with the double burden of (mal)nutrition.

First of all, while the study has successfully answered the research question, it is important to acknowledge that there are still gaps that could be filled. For instance, future researchers could build on this paper by trying to expand the predictive power of the model through a search for extra variables that could be of influence on nutritional status, such as religion, ethnicity, and whether or not they had a child, etc. Or the research topic could be transformed into a qualitative study, where in-depth interviews could provide a more detailed understanding of what the ideal functioning of nourishment is for women in different socioeconomic groups.

Second, the same model could be ran for the other 26 countries dealing with the double burden of (mal)nutrition in Sub-Saharan Africa, for which data is available in the DHS database, using the commands in *"Appendix A: Stata Do-File"*. In doing so differences and similarities can be explored between countries with different degrees of overweight.

And finally, the same model can be executed for countries that have dealt with the double burden of (mal)nutrition longer, or for Cameroon, Lesotho and Zimbabwe in 5 or 10 years. To examine whether this current analysis was not able to find a non-linear relationship due to them not yet having reached the second stage of the obesity transition.

As for recommendations for policy-makers in countries dealing with the double burden of (mal)nutrition goes, the main idea should be to try to stimulate the transition into the behavioural change phase of the nutrition transition as quickly as possible.

A few options that could accelerate this transition are to educate individuals about the risks of degenerative diseases that accompany over-nourishment. Since currently, higher educated individuals have a higher risk of being overweight, implementing these lessons into their education could spur them to change their behaviour towards fats and sugars.

Second, the attractiveness of healthier alternatives could be promoted with subsidies that could lower prices further. In doing so, the higher SES groups could still opt for the more unhealthy options, but the incentive to opt for healthy options would become increasingly larger.

6. References

Abrahams, Z., Mchiza, Z., & Steyn, N. P. (2011). Diet and mortality rates in Sub-Saharan Africa: Stages in the nutrition transition. *BMC Public Health*, *11*(1). https://doi.org/10.1186/1471-2458-11-801

Banerjee, A. V., & Duflo, E. (2011). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. *Population and Development Review*, *37*(4), 796–797. https://doi.org/10.1111/j.1728-4457.2011.00462.x

Becker, W. E., & Kennedy, P. E. (1992). A Graphical Exposition of the Ordered Probit. *Econometric Theory*, *8*(01), 127–131. https://doi.org/10.1017/s0266466600010781

Bor, J., Herbst, A. J., Newell, M. L., & Bärnighausen, T. (2013). Increases in Adult Life Expectancy in Rural South Africa: Valuing the Scale-Up of HIV Treatment. *Science*, *339*(6122), 961–965. https://doi.org/10.1126/science.1230413

Chiappero-Martinetti, E., & Venkatapuram, S. (2014). The Capability Approach: A Framework for Population Studies. *African Population Studies*, *28*(2), 708. https://doi.org/10.11564/28-2-604

Chowdhury, M. Z. I., & Turin, T. C. (2020). Variable selection strategies and its importance in clinical prediction modelling. *Family Medicine and Community Health*, *8*(1), e000262. https://doi.org/10.1136/fmch-2019-000262

Daykin, A. R., & Moffatt, P. G. (2002). Analyzing Ordered Responses: A Review of the Ordered Probit Model. *Understanding Statistics*, *1*(3), 157–166. https://doi.org/10.1207/s15328031us0103_02

Devaux, M., & Sassi, F. (2011). Social inequalities in obesity and overweight in 11 OECD countries. *The European Journal of Public Health*, *23*(3), 464–469. https://doi.org/10.1093/eurpub/ckr058

Djuikom, M. A., & Van de Walle, D. (2022). Marital status and women's nutrition in Africa. *World Development*, *158*, 106005. https://doi.org/10.1016/j.worlddev.2022.106005

Fezeu, L. K., Assah, F. K., Balkau, B., Mbanya, D. S., Kengne, A. P., Awah, P. K., & Mbanya, J. C. N. (2008). Ten-year Changes in Central Obesity and BMI in Rural and Urban Cameroon. *Obesity*, *16*(5), 1144–1147. https://doi.org/10.1038/oby.2008.44

Forster, S., & Gariballa, S. (2005). Age as a determinant of nutritional status: A cross sectional study. *Nutrition Journal*, *4*(1). https://doi.org/10.1186/1475-2891-4-28

Gupta, S., & Bansal, S. (2020). Does a rise in BMI cause an increased risk of diabetes?: Evidence from India. *PLOS ONE*, *15*(4), e0229716. https://doi.org/10.1371/journal.pone.0229716

Jaacks, L. M., Vandevijvere, S., Pan, A., McGowan, C. J., Wallace, C., Imamura, F., Mozaffarian, D., Swinburn, B., & Ezzati, M. (2019). The obesity transition: stages of the global epidemic. *The Lancet Diabetes & Endocrinology*, *7*(3), 231–240. https://doi.org/10.1016/s2213-8587(19)30026-9

Kushitor, S. B., Owusu, L., & Kushitor, M. K. (2020). The prevalence and correlates of the double burden of malnutrition among women in Ghana. *PLOS ONE*, *15*(12), e0244362. https://doi.org/10.1371/journal.pone.0244362

Monteiro, C. A., Moura, E. C., Conde, W. L., & Popkin, B. M. (2004). Socioeconomic status and obesity in adult populations of developing countries: a review. *Bulletin of the world health organization*, *82*(12), 940–946. https://doi.org/10.1590/S0042-96862004001200011

Naigaga, D. A., Jahanlu, D., Claudius, H. M., Gjerlaug, A. K., Barikmo, I., & Henjum, S. (2018). Body size perceptions and preferences favor overweight in adult Saharawi refugees. *Nutrition Journal*, *17*(1). https://doi.org/10.1186/s12937-018-0330-5

Odencrants, S., Bjuström, T., Wiklund, N., & Blomberg, K. (2013). Nutritional status, gender and marital status in patients with chronic obstructive pulmonary disease. *Journal of Clinical Nursing*, *22*(19–20), 2822–2829. https://doi.org/10.1111/jocn.12222

Popkin, B. M. (1993). Nutritional Patterns and Transitions. *Population and Development Review*, *19*(1), 138. https://doi.org/10.2307/2938388

Popkin, B. M., Corvalan, C., & Grummer-Strawn, L. M. (2020). Dynamics of the double burden of malnutrition and the changing nutrition reality. *The Lancet*, *395*(10217), 65–74. https://doi.org/10.1016/s0140-6736(19)32497-3

Robeyns, I. (2003). Sen's Capability Approach and Gender Inequality: Selecting relevant capabilities. *Feminist Economics*, *9*(2–3), 61–92. https://doi.org/10.1080/1354570022000078024

Robeyns, I. (2006). The Capability Approach in Practice*. *Journal of Political Philosophy*, *14*(3), 351–376. https://doi.org/10.1111/j.1467-9760.2006.00263.x

Sartorius, B., Veerman, L. J., Manyema, M., Chola, L., & Hofman, K. (2015). Determinants of Obesity and Associated Population Attributability, South Africa: Empirical Evidence from a National Panel Survey, 2008–2012. *PLOS ONE*, *10*(6), e0130218. https://doi.org/10.1371/journal.pone.0130218

Scott, A., Ejikeme, C. S., Clottey, E. N., & Thomas, J. G. (2012). Obesity in sub-Saharan Africa: development of an ecological theoretical framework. *Health Promotion International*, *28*(1), 4–16. https://doi.org/10.1093/heapro/das038

Shetty, P. (2013). Nutrition Transition and Its Health Outcomes. *The Indian Journal of Pediatrics*, *80*(S1), 21–27. https://doi.org/10.1007/s12098-013-0971-5

Sobngwi, E., Mbanya, J. C. N., Unwin, N., Kengne, A., Fezeu, L., Minkoulou, E., Aspray, T., & Alberti, K. (2002). Physical activity and its relationship with obesity, hypertension and diabetes in urban and rural Cameroon. *International Journal of Obesity*, *26*(7), 1009–1016. https://doi.org/10.1038/sj.ijo.0802008

Sutter, J., & Kalivas, J. (1993). Comparison of Forward Selection, Backward Elimination, and Generalized Simulated Annealing for Variable Selection. *Microchemical Journal*, *47*(1–2), 60–66. https://doi.org/10.1006/mchj.1993.1012

The DHS Program. (2014). *The DHS Program - Lesotho: Standard DHS, 2014 Dataset*. https://dhsprogram.com/data/dataset/Lesotho_Standard-DHS_2014.cfm?flag=1

The DHS Program. (2015). *The DHS Program - Zimbabwe: Standard DHS, 2015 Dataset*. https://dhsprogram.com/data/dataset/Zimbabwe_Standard-DHS_2015.cfm?flag=1

The DHS Program. (2018). *The DHS Program - Cameroon: Standard DHS, 2018 Dataset*. https://dhsprogram.com/data/dataset/Cameroon_Standard-DHS_2018.cfm?flag=0

The DHS Program. (2020). *The DHS Program - Wealth-Index-Construction*. https://dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm#:%7E:text=The%20wealth%20index%20is%20a%20composite%20measure%20of%20a%20household's,water%20access%20and%20sanitation%20facilities.

The DHS Program. (2022) - *Quality information to plan, monitor and improve population, health, and nutrition programs.* Demographic and Health Surveys. https://dhsprogram.com/

The World Bank. (2020). *GDP per capita (current US$) | Data*. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

United Nations. (2020). *Human Development Index | Human Development Reports*. Human Development Reports. https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

United Nations Population Division. (2019). *World Population Prospects - Population Division - United Nations*. World Population Prospects. https://population.un.org/wpp/

Wassie, M. M., Gete, A. A., Yesuf, M. E., Alene, G. D., Belay, A., & Moges, T. (2015). Predictors of nutritional status of Ethiopian adolescent girls: a community based cross sectional study. *BMC Nutrition*, *1*(1). https://doi.org/10.1186/s40795-015-0015-9

Wells, T.R., (2013). *Reasoning about Development: Essays on Amartya Sen's Capability Approach.* Retrieved March 8, 2022 from https://repub.eur.nl/pub/40509. Rotterdam: Erasmus School of Philosophy.

Witkowski, T. H. (2007). Food Marketing and Obesity in Developing Countries: Analysis, Ethics, and Public Policy. *Journal of Macromarketing*, *27*(2), 126–137. https://doi.org/10.1177/0276146707300076

World Health Organization. (2010). *A healthy lifestyle - WHO recommendations*. https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations#:%7E:text=The%20recommended%20levels%20are%20adapted,24.9%20as%20a%20normal%20BMI

Appendix A: Stata DO-file

```
*Do-File for Master Thesis / Jan Vriens / s2753200
*Changing working directory:
cd [redacted]
*Syntax for Cameroon, Lesotho and Zimbabwe.
use Cameroon
use Lesotho
use Zimbabwe

*Keep only necessary variables
keep V714 V106 V012 V190 V000 V025 V501 V445
*generating country variables
*For Cameroon
gen country = 0
*For Lesotho
gen country = 1
*For Zimbabwe
gen country = 2

*recoding v025 so that it is binary with 0 and 1.
gen V025_recode = V025
recode V025_recode (2=1)(1=0)

*Recoding BMI into categories
drop if V445==9998
gen BMIcat = V445 / 100
recode BMIcat (0/18.5=0)(18.5/25=1)(25/30=2)(30/100=3)

*Descriptive statistics in general.
tab1 V190 V106 V025_recode V501 V714
sum V012

*Descriptive statistics per BMI category.
tab1 V190 V106 V025_recode V501 V714 if BMIcat==0
tab1 V190 V106 V025_recode V501 V714 if BMIcat==1
tab1 V190 V106 V025_recode V501 V714 if BMIcat==2
tab1 V190 V106 V025_recode V501 V714 if BMIcat==3

sum V012 if BMIcat ==0
sum V012 if BMIcat ==1
sum V012 if BMIcat ==2
sum V012 if BMIcat ==3

*Descriptive statistics for missing variables of BMI.
tab V190 BMIcat, missing
tab V714 BMIcat, missing
tab V501 BMIcat, missing
tab V106 BMIcat, missing
tab V025_recode BMIcat, missing

*Calculating the spearman correlations
spearman V190 V106 V025_recode V012 V714 BMIcat, stats (rho p)
```

```
*Ordered Stepwise Probit model with poorest as a locked category to
calculate predicted probabilities for it.
stepwise, pe(.1) lockterm1: oprobit BMIcat (1b.V190) i.V190 i.V106
i.V025_recode i.V501 V012 V714

*Calculation of predicted probabilities.
margins, at((mean) V190=(1)) at((mean) V190=(2)) at((mean) V190=(3))
at((mean) V190=(4)) at((mean) V190=(5))
*Plotting the predicted probabilities
marginsplot, nolabels


*save and replace files after running to switch between files.
save Cameroon, replace
save Lesotho, replace
save Zimbabwe, replace

*Additional syntax for the combined file.
use Cameroon
append using [redacted]
Lesotho.dta"
append using [redacted]
Zimbabwe.dta"
save Combined, replace

use Combined

*ordered stepwise probit model coefficients country comparison.
*Cameroon as the reference category
stepwise, pe(.1) lockterm1: oprobit BMIcat (1b.V190) i.V190 i.V106
i.V025_recode i.V501 V012 V714 ib0.country
*Lesotho as the reference category
stepwise, pe(.1)lockterm1: oprobit BMIcat (1b.V190) i.V190 i.V106
i.V025_recode i.V501 V012 V714 ib1.country
*Zimbabwe as the reference category
stepwise, pe(.1) lockterm1: oprobit BMIcat (1b.V190) i.V190 i.V106
i.V025_recode i.V501 V012 V714 ib2.country

*Calculating predicted probabilities
margins, at((mean) V190=(1)) at((mean) V190=(2)) at((mean) V190=(3))
at((mean) V190=(4)) at((mean) V190=(5))
*Plotting the predicted probabilities
marginsplot, nolabels

save Combined, replace
```