CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

The Evolution of Intra-household Consumption Inequality in Nigeria

Bachelor's thesis

Author: Anežka Neckářová

Study program: Economics and Finance

Supervisor: Mgr. Marek Šedivý

Year of defense: 2024

Declaration of Authorship

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.
4. During the preparation of this thesis, the author used the Grammarly premium AI tools in order to improve the clarity and phrasing of the text. After using this tool/service, the author reviewed and edited the content as necessary and takes full responsibility for the content of the publication.

Prague, July 30, 2024

Anezka Neckarova

Abstract

This thesis aims to investigate the presence of consumption inequality among household members on the example of Nigeria and compare the findings to previous research in this field. To assess consumption inequality among household members, it is important to determine the individual consumption levels of each member. As the resource shares allocated to individual household members are not directly observable due to consumption of goods that cannot be assigned to a specific member and lack of sufficient data, they need to be estimated. The methodology we use is based on seemingly unrelated regression estimation of slopes of Engel curves for specific types of goods called assignable, in our case women's, men's, and children's clothes. Our findings indicate that in Nigeria between 2018-2019, the estimated share of consumption is highest for women, which is an unexpected result when compared to previous literature on the topic.

JEL Classification	D13; D63; I32
Keywords	inequality, household, Nigeria, distribution, consumption, assignable goods
Title	The Evolution of Intra-household Consumption Inequality in Nigeria

Abstrakt

Cílem této práce je prozkoumat existenci nerovnosti ve spotřebě mezi členy domácností na příkladu Nigérie a porovnat výsledky s předchozími výzkumy v této oblasti. Pro posouzení tohoto typu nerovnosti je důležité zjistit individuální úrovně spotřeby jednotlivých členů. Vzhledem k tomu, že podíly zdrojů přidělené jednotlivým členům nejsou přímo pozorovatelné z důvodu existence spotřeby zboží, kterou nelze přiřadit konkrétnímu členovi, a nedostatku potřebných dat, je třeba tyto podíly odhadnout. Metodika, kterou používáme, je založena na zdánlivě nesouvislém regresním odhadu (SURE) sklonů Engelových křivek pro konkrétní druh zboží, která se nazývají přiřaditelná, v našem případě to je dámské, pánské a dětské oblečení. Naše zjištění ukazují, že v Nigérii v letech 2018-2019 je odhadovaný podíl spotřeby nejvyšší pro ženy, což je ve srovnání s předchozí literaturou na toto téma neočekávaný výsledek.

Klasifikace JEL	D13; D63; I32
Klíčová slova	nerovnost, domácnost, Nigérie, distribuce, spotřeba, přiřaditelné zbož
Název práce	Vývoj nerovnosti ve spotřebě uvnitř nigerijských domácností

Acknowledgments

The author is especially grateful to Mgr. Marek Šedivýfor his patience and instructions. Furthermore, I would like to thank my parents for their support.

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record

Neckářová, Anežka: *The Evolution of Intra-household Consumption Inequality in Nigeria.* Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 54. Advisor: Mgr. Marek Šedivý

Contents

List of Tables

List of Figures

Acronyms

BCL model the model introduced by Browning *et al.* (2013)

DLP model the model introduced by Browning *et al.* (2013)

SURE Seemingly unrelated regression

Bachelor's Thesis Proposal

Author	Anežka Neckářová
Supervisor	Mgr. Marek Šedivý
Proposed topic	The Evolution of Intra-household Consumption Inequality in Nigeria

Research question and motivation The central research question of my bachelor thesis should be: Is there enough evidence about intra-household inequality in consumption in Nigeria in 2019? (also compared to results from 2013)

Inequality, generally, is one of the biggest problems of present society all over the world and so is discrimination. Nowadays, there is enough evidence that these two problems could come together and one of the possible ways how to prove that is the existence of intra-household consumption inequality (Lechene, 2022), in other words that not all members of the household consume the same proportion of the budget. It is proved that the overall inequality when accounting also for intra-household inequality can increase even by 50% (AlmĂĄs, 2021). This is also connected to new, different views and poverty. It is possible that when a household to live under poverty line, in fact, only some members of the household live under poverty line or the other way around.

Studying particular situations in each country might be useful for understanding the issues on deeper level and also identify countries where some policies identified as mentioned in an article Understanding inequality within households by AlmĂĄs.

Nigeria is considered an emerging economy today, it is also the largest economy in Africa, which is still the poorest continent of the world and that position makes it interesting, however is it also something that could make this country sensitive to problems such as inequality? That is not going to be answered, but it definitely makes more motivation for focusing on this particular country. Also there was high inflow of foreign direct investment in 1992 â€“ 2013, which could also make some difference in every part of life there. (Adeleke Kunle, 2014)

Expected contribution There was already some research focusing partially on Nigeria (and other countries). I want to extend this by focusing on different point of time, compare my results to the existing one and hopefully make some statements about development of this trend during recent period.

The key source for my thesis will be the paper by ValĂ©rie Lechene from 2022, which I want to follow and extend itâ€™s results as mentioned.

There are many papers concerning situations in other particular countries (India by Klasen, 2020), however no work was done concerning particularly Nigeria situation.

Methodology I intend to use microdata dataset provided by the World bank from Living Standards Survey 2018-2019 (there exist many waves of these data for many countries). Key thing in this dataset is that it provides information about some assignable goods, which will be widely used in my thesis. In the previous work they used the latest available wave of this dataset and today I can work with the most recent one also, that means 5 years later, which is not so large period, however in such developing country there might be some differences and if there are not, it might also say something.

The methodology I am going to use is also very similar to the previous paper. It is primarily based on the ordinary least square estimation of the resource shares consumed by each individual in the households using Engel curves and consumption of assignable good by each member of the household as the starting points. The estimation itself will be done at R

Outline

1. Abstract
2. Introduction
    - Why is this topic important
    - Previous work done on intra-household consumption inequality
    - Nigeria economy, why to study inequality here
3. Data and methodology
    - Dataset and processing of the data
    - Presenting of the model used
    - Application of the model on the data
4. Results

- Presenting of results of the estimation
- Interpreting the results
- Comparing the results to existing literature

5. Conclusion

Bibliography LECHENE, ValĂ©rie, Krishna PENDAKUR a Alex WOLF. Ordinary Least Squares Estimation of the Intrahousehold Distribution of Expenditure. Journal of Political Econonmy. 2022, 2022(130).

AlmĂĄs, Ingvild; Ringdal, Charlotte; Hoem Sjursen, Ingrid (2021) : Understanding inequality within households, GLO Discussion Paper, No. 961, Global Labor Organization (GLO), Essen

KLASEN, Stephan a Rahul LAHOTI. How serious is the neglect of intra-household inequality in multidimensional poverty and inequality analyses? Evidence from India. Review of Income and Wealth. 2020, 2021(67).

Bonke, J. and Browning, M. (2009), The Allocation of Expenditures within the Household: A New Survey. Fiscal Studies, 30: 461-481. https://doi.org/10.1111/j.1475-5890.2009.00104.x

M., Adeleke Kunle, Olowe S.O. a Fasesin Oladipo OLUWAFOLAKEMI. Impact of Foreign Direct Investment on Nigeria Economic Growth. International Journal of Academic Research in Business and Social Sciences. 2014, 4(8), Pages 234-242. ISSN 2222-6990. DostupnĂ© z: doi:10.6007/IJARBSS/v4-i8/1092

Chapter 1

Introduction

World Bank (2024a) states that over 700 million people (around 9% of the World's population) live in extreme poverty, meaning that they consume less than \$2.15 per person per day. However, the majority of poverty rate statistics are based on households as the lowest unit of consideration. This approach may produce biased results by overlooking individual members' economic situations. What if only certain household members are living in extreme poverty rather than the entire household, and therefore, the standard poverty measures are inaccurate?

Imagine a household comprising four members: a man, a woman, and two children. We consider the case of per capita measurement implying that, if the statistics commonly provided by national statistics offices or other institutions indicate that the household's daily consumption is valued at \$8, it is assumed that the individual daily consumption is worth \$2, which places all the household members under the extreme poverty line. In reality, it is possible that, for example, the man consumes \$2.5, each child consumes \$2, and the woman only consumes \$1.5. In this case, only three out of the four household members live in extreme poverty. Standard poverty measures could overestimate the real poverty rate by counting all four household members as extremely poor instead of the actual three. On the contrary, if the entire household consumes, for example, \$10 per day, the assumed individual consumption of \$2.5 is above the extreme poverty line, and none of the household members would be considered poor. However, if the resources are distributed unequally among household members, it is possible that some of them consume less than \$2.15 per day and should be, in fact, considered in poverty.

This type of inequality is often overlooked due to the lack of relevant data.

When data is collected at the household level rather than the individual level, it becomes challenging to determine the distribution of resources within the household, which is commonly the case since data collection at the individual level tends to be very complicated.

Intra-household inequality can be studied using collective household models, which are commonly used to analyze household demand. In this thesis, we are building on the work of Lechene *et al.* (2022), who developed a model based on the frameworks established by Browning *et al.* (2013), BCL model, and Dunbar *et al.* (2013), DLP model. BCL and DLP models make use of the collective household models, making them more accessible for estimation. Lechene *et al.* (2022) have estimated the model for 12 countries, including Nigeria, which we chose to focus on and compare our results to those presented by Lechene *et al.* (2022).

According to the World Bank (2024b), Nigeria has Africa's largest economy and population. World Bank (2024b) estimate that the poverty rate will reach 38.9% in 2023, with around 87 million people living below the poverty line, making it the second-largest population in poverty in the world after India.

Lechene *et al.* (2022), in the case of Nigeria, conducted research using data from 2012-2013. We rely on a more recent dataset from 2018-2019. The consumption behavior of households in Nigeria may have changed during that time. However, the main problem in the case presented by Lechene *et al.* (2022) was that the model failed to produce reliable results for the dataset collected in Nigeria in 2012-2013. Given the importance of Nigeria, we replicate the Lechene *et al.* (2022) approach, hoping for reliable estimates.

To sum up, in this thesis, we apply the theory proposed by Lechene *et al.* (2022) on data collected in Nigeria in 2019. First, the development of the theory and the model is introduced together with other research considering this topic applied to data from many countries worldwide. Subsequently, the methodology is explained in detail in Chapter 3. Chapter 4 provides a description of the dataset used and the variables included in the model. Finally, the results of the regression we present in Chapter 5.

Chapter 2

Literature review

The scope of research on intra-household inequality has been expanding since the decade of the 1980s when the first studies accounting for non-equal distribution among household members appeared. Lise & Seitz (2011) states that the standard measures of inequality use adult equivalence scales for dividing consumption among household members. This approach assumes equal consumption-sharing among adult household members, implying no disparity between them. As a consequence, standard inequality and poverty measures take a household as the lowest unit. However, accounting for inequality between household members is also essential in studying poverty, as explained.

Chiappori (1988) proposed the use of collective household models, based on which Browning *et al.* (2013) (BCL model) and Dunbar *et al.* (2013) (DLP model) built methods to uncover within-household consumption distribution.

BCL model addresses within-household consumption by estimating the model based on separate utility functions for each household member, the consumption technology that characterizes to what extent the goods consumed are shared (consumed by more household members together), and the sharing rule, which states how the goods are shared, meaning who gets a bigger or smaller share of the goods that are shared. Using the three components the BCL model seeks to uncover the distribution of bargaining power within households and if it is beneficial for a person to be part of a household compared to living alone and to share consumption of goods that can be shared. They applied the model to data collected in Canada from 1974 to 1992. They found that in households including one man and one woman, more than half of the income is allocated to the woman, and the cohabitation benefits are present. However, the model cannot be applied to household compositions including children. We also need to consider that Canada belongs to more developed countries, and the results might differ for developing countries.

Advancing the theory proposed by the BCL model, the DLP reduced the model and data requirements by imposing several restrictions. Assuming that the preferences of people living alone do not differ from the preferences of people living in a couple, the BCL model uses individual demand functions for men and women living in single households and combines those with data on household demand for the household including a man and a woman living together to capture the variations in prices between single households and households including a couple. That is why it is impossible to include children in the BCL model, as they always live together with adults, implying that collecting data on their individual demands is impossible. If the individual demands are not observable, the BCL model cannot be identified. Therefore, while the BCL model accounted for only childless adult couples, which gave an incomplete picture of the household distribution, the DLP model allows one or more children to be present in a household.

Dunbar *et al.* (2013) states that children might be the most sensitive members to intra-household inequality and emphasizes the importance of focusing on possible children's poverty. However, the DLP model still needs to be revised to count for all possible household structures. It does not enable more adults of the same gender to live in the same households. This is a particularly significant restriction in developing countries, where the households tend to include many members of more generations and be rather complex.

Another restriction imposed by the DLP model compared to the BCL model is using Engel curves for a single good only instead of demands for all the goods consumed. This once again decreases the data requirements by a considerable amount. The proportion of consumption allocated to each household member is then counted from the slopes of Engel curves. How the model works will be explained in detail in the Methodology chapter.

The DLP model is applied to data from Malawi. Malawi is a less developed country than Canada (see, for example, GDP per capita indicator by the World Bank (2023)). Consequently, the results might offer another point of view on the intra-household inequality issue.

The DLP model estimates the resource shares first separately for households with one, two, three, and four children and for all households together as well. They found that, on average, lower resource shares are dedicated to women of all types of household compositions. When estimating for all households together, the resource share allocated to women was, on average, 30%, while for men, the share of household resources consumed by them was 49%. Each child, on average, consumes 10% of household resources. An important finding is that the share consumed by women is lower for families with more children. In households with one child only, the woman consumes 40%, which compared to men's share of 46% makes the inequality within this kind of household not as momentous as in households with more children where the women's share is between 24 and 28% and for men it is 44% and 52%.

The DLP model also focused on poverty rates within households. They multiply the resource share dedicated to each member by the household expenditure and compare the results with the poverty threshold of consumption worth $2 per day set by the World Bank for an adult household member together with the OECD estimate of children needs being 60%. The threshold for children is set at $1.2 per day (60% of the adult's needs). After doing this, for each household, they count the percentage of households where men, women, and children live in poverty. The result is that the share of women living in poverty increases with number of children, from 76% for households with one child to 83% for households with four children. The trend does not apply to men, indicating that women may be more willing to sacrifice their consumption for their children. They also compare these shares to standard poverty rates, with the result being that these systematically underestimate the actual poverty rate. Therefore, more research on intra-household inequality is necessary since its impact on overall inequality might be significant.

Lechene *et al.* (2022) developed the model we intend to apply in this thesis. Following the BCL and DLP models, Lechene *et al.* (2022) propose a linear reframing of the DLP model and estimate the new model for 12 countries across Africa, Asia, and Eastern Europe using less complicated seemingly unrelated regressions (SURE) estimation. This model already allows more adults of the same gender to live within one household. The specific distinctions of the model are detailed in the Methodology chapter.

The data did pass the identification test for 5 of the 12 countries considered. Therefore, the resources allocated to each household member can be estimated for these countries: Albania, Bangladesh, Bulgaria, Iraq, and Malawi. The rest of the datasets failed the identification test, and the results are not reliable. These include Ghana, Tajikistan, Tanzania, Timor-Leste, Uganda, Ethiopia, and Nigeria.

Lechene *et al.* (2022) found that for four out of the five countries left after eliminating those that did not pass the test, the estimated share of resources allocated to men is higher than for women. The only exception is Bulgaria. Here the resource shares dedicated to women are higher by almost 7 percentage points. However, after they consider only the fraction of households that include at least one adult member of each gender (elimination of those that include only men or only women and children), the gender gap, the difference between the share of resources allocated to men and to women, is not statistically significant. In fact, the gender gaps are insignificant in Albania and Malawi as well. For both Bangladesh and Iraq, for which the shares are significant, a higher share of resources is allocated to men with the gender gap being 4.5 percentage points in Bangladesh and 4.1 percentage points in Iraq.

In their study, Lechene *et al.* (2022) discovered that children consistently receive fewer resources than adults in all countries. However, it is important to note that children also have lower needs, and their available resources are compared to a lower poverty line of 60% of the adult poverty line, similar to the approach taken by Dunbar *et al.* (2013). Children in Iraq receive the lowest resource shares among the countries considered. Children in Albania, Bangladesh, Bulgaria, and Malawi are given between 12% and 18% of the total household resources, while in Iraq, this is only 5%. When applying the poverty line for the upper middle-income class (60% of $5.5 for a person per day) Lechene *et al.* (2022) found that 68% of children in Iraq live in poverty. However, they state that there are data issues present in the dataset for Iraq, and subsequently, these results might be misleading. Besides the data issues, the households in Iraq have an average of more children than households in other countries, and the children are younger, meaning that they might have lower needs than older children. Therefore, the resources allocated to children in Iraq might be lower compared to other countries even if no data issues were present.

Lechene *et al.* (2022) show that the difference in estimated resource shares among household members directly results in variation in estimated poverty rates among them. They emphasize the importance of estimating the within-household poverty rates by comparing the results we get by the per capita model, commonly used by the World Bank and other international organizations, to obtain the poverty rates in a country. The per capita model counts the resource shares as one divided by the number of household members and can not account for scale economies in consumption (benefits of living in a household compared to living alone). Lechene *et al.* (2022) suggests that these models leave out a lot of variation in poverty rates between countries since the shares of resources allocated to different household members Lechene *et al.* (2022) obtained differ across countries a lot while the per capita model assumes that the resource shares are equal for all household members and differ with household size only. They conclude that the currently used poverty measures are misleading and underestimate the actual poverty rates within a country by leaving out the within-household inequality.

Verifying its existence in as many cases as possible is necessary to obtain the overall picture of intra-household inequality's presence worldwide. Studies using various methodologies and datasets from many countries have already uncovered the allocation of resources within households. We focus on those using the collective household models since they are the core of the methodology we use as well.

Malawi, mentioned in the description of results obtained by Dunbar *et al.* (2013) and Lechene *et al.* (2022), is an example of a country where we have more results available. We already discussed the model and its results presented by Dunbar *et al.* (2013) for families with different numbers of children (from 1 to 4). For this research, data from the years 2004-2005 were used. Dunbar *et al.* (2021) update these results. They relaxed the similarity of preferences restriction and utilized newer data from the years 2016-2017 for households with 1 - 4 children separately. The first paper found that women receive less of the resources available to the household in all types of households (between 27% and 40%) while between 44% to 52% were allocated to men. Therefore, the gender gap was between 6 and 28 percentage points. The newer paper shows completely different results. For two out of four types of households, the estimated resource shares are higher for women. In households with 4 children, the estimated resource shares are equal between genders. The biggest gender gap is faced by households with three children, where women receive by 4 percentage points more. Also, the resources available to children have risen from 14-29% to 27-40%. Lechene *et al.* (2022) used the exact same dataset from the years 2016-2017 for Malawi, with the difference that they did not distinguish the households by the number of children and found that lower resource shares are allocated to women with the gap being 4 percentage points. This result is different from the one obtained by Dunbar *et al.* (2021). However, it does not show such extreme gender gaps as Dunbar *et al.* (2013). Therefore, Malawi might be actually moving towards lower within-household inequality.

Penglase (2020) also works with data from Malawi. He estimates his model for the dataset collected during the years 2010-2016. Moreover, his research focuses on the difference between children in foster care and those living with their biological parents. He found that households allocate more to foster children than to non-foster. The gender gap between adults here is 7 percentage points, again in favor of men.

With more papers available for this country, we can understand the existence, extent, and development of intra-household inequality in Malawi. The goal in this field of research should be to have this kind of information available for as many countries as possible. That would enable us, for example, to compare the extent of intra-household inequality and potential discrimination between the developed and developing world.

Another reason to study intra-household inequality is the possible consequences of this kind of inequality. Calvi (2020) inspects the differences in mortality rates of men and women in India and the connections of mortality rates to the resource shares allocated to each gender. She indicated a high difference between resource shares allocated to men and women in Indian families having children. In those families, women receive by 15 percentage points lower resource share than men. In families with no children, it is only 6 percentage points. He shows that the asymmetric distribution of resources within Indian households might be an explanation for the higher mortality of women at postproductive age, so-called "missing women".

We have mostly discussed developing countries so far. However, studies on this issue in developed countries offer a different perspective. For example, Bargain & Donni (2012) focused on France, and Bargain *et al.* (2022) focused on the United Kingdom. Both studies found that higher resource shares are allocated to women than to men for all household compositions considered. This contrasts the opposite trend observed in developing countries, although the difference is not as wide.

Chapter 3

Methodology

The methodology we apply in this thesis, presented in this section, is based on previous work done on collective household models in general. Primarily, we build on research done by Lechene *et al.* (2022) introducing a linear model of household Engel curves and their SURE estimation to find the resource shares allocated to each household member type. We use the resource shares to uncover possible inequalities present between these types.

Lechene *et al.* (2022) work with the collective household models introduced by Browning *et al.* (2013) and its extension provided by Dunbar *et al.* (2013) imposing several restrictions set by the BCL model, which enabled them to recover the resource shares. Nevertheless, both of these models are complicated to estimate as they require an estimation of non-linear regression equations. Lechene *et al.* (2022) transform the model to linear form so that the model parameters can be estimated using the SURE model, which is easier to perform.

First, we outline the notation used in the model and provide the necessary definitions. Then, we explain the logic behind the collective household models using an example. Finally, we show the model's linear reframing and the model identification test.

For the estimation, we apply the code provided by Lechene *et al.* (2022). The code can be accessed on the journal's website.

3.1 Model setting

In the model, we distinguish t household member types, i.e., men (m), women (f), and children (c). We denote the number of members of type t living in a household N^t. We set these three particular household member types because of the data availability. We determine the type of each individual using gender and age variables provided by the dataset we use. However, the same estimation could be done with any other groups of household members, which are disjoint and possible to distinguish with the data available. The number of groups can differ as well.

Resource share allocated to type t, η^t, is the percentage of total household expenditure consumed by all the individuals of type t living in the household. The sum of resource shares of all three types should be equal to 1 (whole consumption of the household).

$$\sum_{t=1}^{T} \eta^t = 1 \tag{3.1}$$

In terms of the particular groups we use, this means:

$$\eta^m + \eta^f + \eta^c = 1 \tag{3.2}$$

Another important concept used in this methodology is the Engel curves. Chai & Moneta (2010) states that an Engel curve for a good describes how the share of the household budget spent on the good changes with the size of the overall household budget spent on all goods. W^t represents the Engel curve at the household level for a specific assignable good type t (the assignable good is a specific type of good consumed by household members that will be defined shortly). It depends on the household budget y, spent on all goods by all household members, therefore, we write $W^t(y)$ and determine the share of this budget spent on the assignable good of type t. According to Engel (1857), the slope of Engel curves indicates whether the good of interest is a necessity (downward sloping Engel curve) or a luxury (upward sloping Engel curve).

The individual-level Engel curve function for assignable goods of type t of a person of the same type, denoted w^t, is derived from the amount consumed by one person of that type out of the overall household budget only. The individual level budget of a member of type t, assuming equal sharing among members of the same type within a household, is given by the resource share allocated to all household members of the type t times the household budget y (resources allocated to all members of the type t) divided by the number of members of that type living in the household, $\eta^t(y)y/N^t$. Therefore the individual-level Engel curves are $w^t(\eta^t(y)y/N^t)$.

We set a vector of covariates needed for the model and denote it z^t. The covariates include the number of members of each type t living in the household and demographic variables such as the age of the household members.

3.2 Collective household models

Collective household models are commonly used tools in economic research on household demand. Alderman *et al.* (1995) states that, unlike unitary household models, collective models of household behavior take into account the distribution processes within the household. Unitary models see inequalities in distribution within the household as a willing choice of household members.

The collective household models enable the use of data on one (or more, if available) assignable goods to calculate the overall resource shares consumed by individual household members, depending on their type.

3.2.1 Assignable goods and the "shadow" budgets

Households consume two types of goods, public and private. The private goods are consumed by only one household member at a time. Therefore, if one person is consuming the good, it prevents other people from consuming the good, i.e., clothes or food. Public goods are the opposite. These can be consumed by several people at the same time (can be shared), and if only one person consumes them, another person can join the consumption, i.e. rent or electricity and water supplies. (Almås *et al.* 2020)

Most datasets show the consumption levels of both types of goods on the household level, which implies that we do not know, which member consumes how much of a good regardless of the kind of the good.

Almås *et al.* (2020) states that assignable good is a specific type of private good, for which we have the data on individual consumption available. In other words, its consumption can be assigned to individual household members.

A suitable example of an assignable good is clothes. We can clearly distinguish who is supposed to wear the clothes and how much the household members of each type spent on the purchases of clothes by dividing clothes into three categories, men's clothes, women's clothes, and children's clothes, and collecting data on household expenditure on these categories. The availability of this data is an important argument supporting the use of the household member types we introduced. If we considered other household member types, we might have a problem with the scarcity of data required since it would not be as straightforward to assign a consumption to a particular household member type as it is when we use gender and age to identify the types. We must assume an even distribution of clothes within each category to find the individual-level consumption as mentioned. This means that we assume that if there are, for example, 3 women living in a household, each of them consumes the same share.

Food can also serve as an assignable good, however, the data must be collected directly on an individual level. In the dataset we use, data about food consumption are collected on the household level, therefore food can not be considered an assignable good. We count on household members of type t consuming, besides other things, an assignable good of the same type t, not other types.

For public goods, shares consumed by individual members are difficult to obtain. While for private goods, the value of individual consumption is equal to individual expenditure, that is not true for public goods. Every household member participating in the consumption of a public good within the household consumes the whole household expenditure on that good not just a share. For example, when the whole household is watching TV, the expenditure of each member is just the expenditure spent by the household on watching TV divided by the number of household members. If there are 3 people in the household and the cost of a TV watch is \$30, the individual expenditure is \$10, but everyone consumes the whole TV watch. Therefore, the individual consumption is worth \$30 for each member of the household. This equality of consumption levels among household members holds if we assume that the whole household watches TV together the whole time. If one of the members watches the TV only half of the time compared to others, then her consumption is lower. Therefore, we would need to collect data on which household members are participating in the consumption of the public good and if they are all consuming all of it. In the example, this would mean that we would need to find out if all 3 household members are watching the TV and if all of them are watching it for the whole time.

The example shows that in the case of public goods the value of consumption (\$30 for each) is higher than the expenditure on the good for household members participating in the consumption (\$10 for each). Lechene *et al.* (2022) introduce the term "shadow" budget for this purpose. The "shadow" budget is the sum of the prices of consumption. The sum of the "shadow" budgets of all household members is higher than the real household budget y. In the case of the TV watch, the real household budget is \$30 (household expenditure), while the "shadow" budget is 3 times \$30 (the sum of consumption prices of all household members). For an individual, the real budget is \$10 and a "shadow" budget is the whole cost of a TV watch, \$30 (the value of their consumption).

To show the difference with private goods, if the same household spends in total \$30 on food, every member consumes just a part of the overall expenditure on food and if we sum these parts we get the overall expenditure of \$30. Therefore for private goods, the "shadow" budget is equal to the real expenditure (consumption value is equal to expenditure). The only problem with private goods is that for those, that are not assignable, we do not observe the shares consumed by individual household members.

The unknown size of "shadow" budgets for the consumption of public goods and the fact that we cannot observe the distribution of all private goods are the reasons why getting resource shares consumed by individual members of the household is not straightforward.

We use available data on the consumption of assignable goods to obtain the overall resource shares of overall consumption. However, if we did this by simply assuming that goods, that are not assignable, are consumed by the same proportion as the assignable, we would get biased results. Therefore collective household models are used as a proper tool for estimation of the individual resource shares consumed.

3.2.2 The Model

We show how the model works using an example of a simple household. Suppose a household includes 3 members, one man, one woman, and a child ($N^c = 1, N^f = 1, N^m = 1$). Also suppose that, household members consume together 3 kinds of goods, clothes, food, and housing. On each of these goods, the household as a whole spends part of its budget y. Clothes are a private good and they are also assignable since we can observe the consumption of women's, men's, and children's clothes separately. Food is a private good, but, unless we have data collected on an individual level available, it is not assignable. Housing is a public good since it is consumed jointly by all household members.

Out of the data routinely provided by the statistical offices, we get how much the household spends on each of the 3 goods. We suppose that this household spends \$70 on food, \$100 on housing, and \$30 on clothes. Clothes are an assignable good, therefore we know how much of it each household member consumes. If the household spends \$10 on children's clothes, \$15 on women's clothes, and \$5 on men's clothes, we take as that each (in our case only 1) child consumes \$10, woman \$15 and man \$5 out of the overall expenditure \$30 spent on clothes. The amounts consumed by individual members are not calculable for food and housing as these are not assignable. The sum of the "shadow" budgets of all 3 household members is greater than the overall household expenditure equal to \$200.

Because of the presence of goods, which are not assignable, in household consumption, the resource shares allocated to individual household members are not observable and need to be estimated as explained in the previous section. This is where the Engel curves become useful. Through the estimation of parameters of household-level Engel curves for assignable goods, we can get the resource shares allocated to each type.

The whole household spends \$10, equal to 5% of its budget, on children's clothes. The 5% is the Engel curve function of the household level Engel curve for children's clothes at the value of household budget equal to \$200: $W^c(\$200) = 0.05$. The same applies to all three types of clothes.

Table 3.1: Engel curve functions at household level

Household member type (t)	consumption of clothes	$W^t(\$200)$
c	\$10	0.05
f	\$15	0.075
m	\$5	0.025

These values indicate how much the household allocates on the consumption of clothes of each type t out of the household budget and we can easily get those from data collected at the household level.

However, we can not simply say that the consumption of clothes mirrors the overall consumption shares since household members of different types might differ in preferences for the consumption of different types of goods. In the example, the woman consumes the highest share of clothes purchased by the household. This does not imply that she is consuming more of the other goods as well. She might prefer the consumption of clothes compared to the other two household members and consume less of goods she does not care about that much.

We assume that the household members of the same type t living in the same household have the same preference and, therefore, identical Engel curves.

If more members of the same type live in the same household, they would share identical preferences and consume the same proportions of all goods.

We observe only one value of the Engel curve function of each household for three types of assignable goods (one point of the Engel curve), not their shapes. At this point, with a value of $200 of the household budget, this particular household allocates a higher share of its expenditure on women's clothes. However, the 3 Engel curves might have different slopes, and the proportions of the budget allocated to each type might change differently for each type with a change in income. We do not know how big the shares would be at other values of the household budget. We use individual-level Engel curves to find the shape and location of the household-level Engel curves for assignable goods.

The Engel curve of the whole household for an assignable good of type t is equal to the Engel curve of the particular individual consuming the assignable good of type t, person of type t, times the resource share of all individuals of that type living in the household.

$$W^t(y) = \eta^t(y) w^t(\eta^t(y) y / N^t) \tag{3.3}$$

Out of the equation 3.3, it is possible to obtain the resource share allocated to the household member of type t. However, to do that, the other variables included need to be observable in order for us to count the unobservable resource shares by dividing the household-level Engel curve function for assignable good of type t by the individual-level Engel curve function for the same type of good. The household-level Engel curves are observable and their functions are easy to calculate as shown. Individual-level Engel curves are not observable. As we do not know the size of each person's so-called "shadow" budget, we cannot count the proportion of the budget spent on assignable goods (clothes). We need to use the "shadow" budget since it is the most appropriate measure of the resources available to each household member. Calculating the individual-level budget another way does not show a realistic picture of the resources consumed by individual household members.

Since the equation 3.3 includes two unobservable variables (resource shares and individual-level Engel curves) and only one observable variable (household-level Engel curves), we can not get the resource shares from this equation. To solve this problem, we substitute for the individual-level Engel curves.

We assume that individuals of each type t living in a household, generate

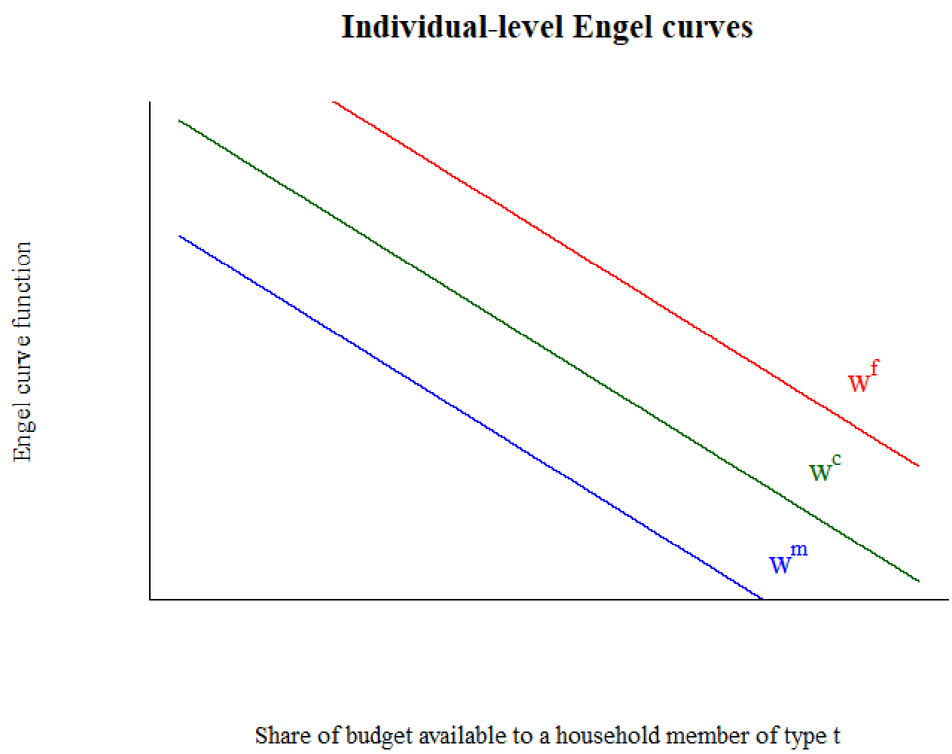

Figure 3.1: Individual level Engel curves

a linear Engel curve and that the slope β of the Engel curve is the same for all three types in a single-member household ($\beta^m = \beta^f = \beta^c = \beta$. Therefore, the Engel curves differ only in their location (given by the intercept parameter α^t), not shape. The picture 3.1 shows an example of how the Engel curves of individual household members can look like. The variation in the intercept parameters among different household member types is given by demographic differences between the types.

$$w^t(\eta^t(y)y/N^t) = \alpha^t + \beta \ln(\eta^t(y)y/N^t) \tag{3.4}$$

The next step is substituting this relationship into the equation 3.3. We have to assume that individuals sharing a household with others of the same or different types have the same preferences as if they were living alone.

$$W^t(y) = \eta^t[\alpha^t + \beta(\ln y + \ln \eta^t - \ln N^t)] \tag{3.5}$$

After the substitution, the household-level Engel curve function depends on individual-level budget share $\eta^t(y)y/N^t$ (not only the household-level budget as before). The result is that household-level Engel curves for assignable goods of type t are a function of resource shares dedicated to that type t, household budget, and number of household members of type t. The household level Engel curves are shown in the picture 3.2 below.

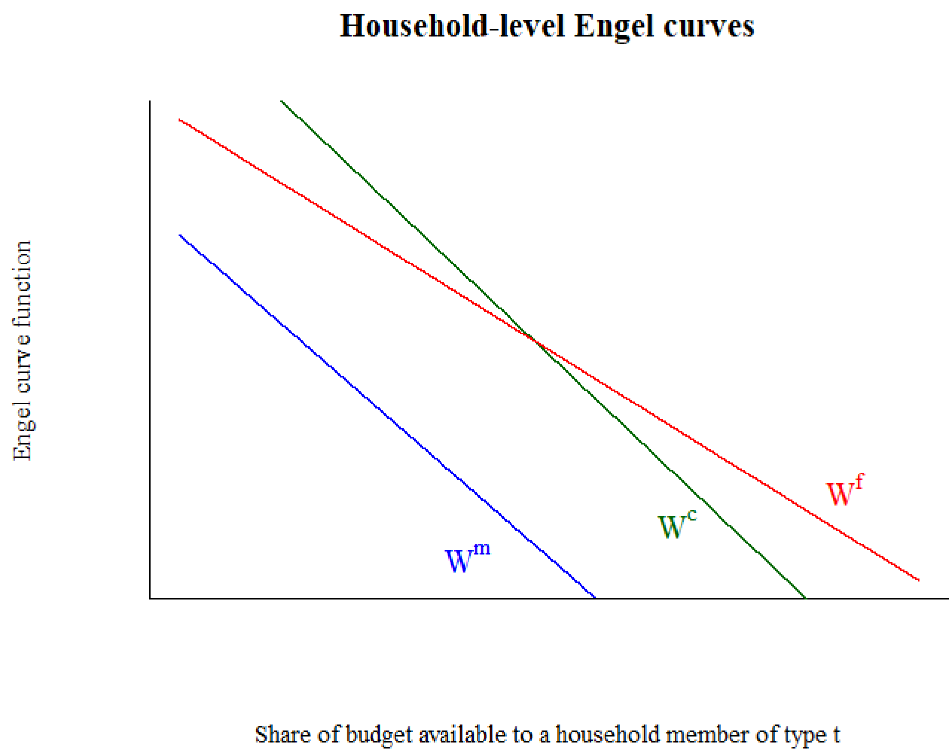

Figure 3.2: Household level Engel curves

The slope of the household-level Engel curves we got after the substitution is the slope of the individual-level curve for the good of type t times the resource share dedicated to the same type t of a household member, $\beta \times \eta^t$, 3.7. We denote the slope parameter of the household-level Engel curves for the assignable good of type t as b^t. Besides that, we know that the resource shares allocated to the 3 types of household members all three types sum to 1, A.6.

$$\sum_{t=1}^{T} \eta^t = 1 \tag{3.6}$$

$$b^t = \beta \times \eta^t \tag{3.7}$$

Putting this together, we have a set of 4 equations containing four unknown variables. These variables are the resource shares allocated to each type t

of household member, η^t, and the slope of the individual-level Engel curves, β. Out of this set of equations, we get that β is the sum of the slopes of the household-level equations of each type t, therefore after performing the estimation, we are going to have all the values required for obtain the resource shares allocated to all types t (for the calculations see the Appendix **??**). The estimated slopes are denoted as $\hat{b}^t$. We get the resource shares by substituting the estimated slopes into the equation A.10.

$$\eta^t = \frac{\hat{b}^t}{\sum_{t=1}^{T} \hat{b}^t} \tag{3.8}$$

For the model to be complete, we include the demographic covariates (vector of covariates z). In the same way as Lechene *et al.* (2022) do, we assume that the resource shares, as well as the slope and intercept parameter of the Engel curves, depend on the demographic covariates. We include z into equation (A.2) and at the same time multiply the variables in bracket with $\eta^t(z)$.

$$W^t(y,z) = \eta^t(z)\alpha^t(z) + \eta^t(z)\beta(z)\ln y + \eta^t(z)\beta(z)\ln\eta^t(z) - \beta(z)\ln N^t \tag{3.9}$$

There are several interaction terms present in the model since, as a result of the substitution for individual-level Engel curves, all the variables of the individual-level Engel curve are multiplied by the resource share. These interaction terms make the model nonlinear and the parameters of it complicated to estimate. The slope estimate of the Engel curve alone, as we need to get the resource shares, would be complicated to get. Lechene *et al.* (2022) propose a linear reframing and approximation of the model, to solve the non-linearity problem effectively.

3.2.3 Linear reframing, approximation, and SURE estimation of the model

Preceding the linear reframing, we give the equation a form of the econometric model with full notation. We denote individual households with an index "h". ϵ_h^t is the error term of the model, a_h^t is the intercept parameter of the Engel curve of household h for assignable good of type t and b_h^t is the slope parameter of it. The linear form of the household-level Engel curves for assignable good of type t then is:

$$W_h^t = a_h^t + b_h^t \ln y_h + \epsilon_h^t \tag{3.10}$$

Both the intercept parameter a_h^t and the slope parameter b_h^t are linear functions of the variables presented (see Appendix B). For the slopes of the Engel curves, it holds (as before the linear reframing):

$$b_h^t = \eta^t(z_h)\beta(z_h) \tag{3.11}$$

Where b_h^t is the slope of household level Engel curve for assignable good of type t and $\beta(z_h)$ is the sum of the slopes of this kind of Engel curve for all 3 types of household members.

We run the SURE regressions of the Engel curve function for assignable good of type t W_h^t on 1 (for the intercept), $\ln N_h^t$, $\ln y_h$ and $z_h \times \ln y_h$ for each type t. As StataCorp LLC (2021) states, the advantage of the SURE regression is that it considers possible relationships between more regression equations, in our case, between equations for each household member type included, and allows the error terms among the regression equations to be correlated. Moreover, it allows for joint testing.

From this regression, we get estimates of the slopes of household-level Engel curves for assignable goods, $\hat{b}^t$ and their sum and use those to get the resource shares allocated to each type of household member.

$$\hat{\eta}^t = \hat{\eta}^t(z_h) = \frac{\hat{b}_h^t}{\sum_{t=1}^{T} \hat{b}_h^t} \tag{3.12}$$

After plugging for all three types we get all the resource shares allocated to each type t. We get an approximate resource share allocated to a person of type t by dividing the resource share by the number of household members of type t living in the household.

3.3 Test of model identification

A potential problem with this approach is the identification of the model. Lechene et al. (2022) address this problem in their work and propose a way of testing if the model is identified for a given dataset.

From (3.12) we know how to compute the resource shares out of the parameters of the estimated model. This would not work if the denominator is close to or equal to zero.

The test is based on estimating the household Engel curve for all types of assignable goods in interest together. In our case, the Engel curve functions here would be the share of the overall household expenditure spent on all three types of clothes we distinguish. We denote this kind of Engel curve, for household h, W_h and we know that the share spent on all three kinds of clothes is equal to the sum of shares spent on each type, $W_h = \sum_t W_h^t$. The same applies to the parameters of those Engel curves and the error term, $a_h = \sum_t a_h^t$, $b_h = \sum_t b_h^t$ and $\epsilon_h = \sum_t \epsilon_h^t$. Therefore in our linear setting, we have an Engel curve for all types of clothes (or common household consumption of other assignable goods) with intercept a_h, slope parameter b_h and error term ϵ_h as follows:

$$W_h = a_h + b_h \ln y_h + \epsilon_h \tag{3.13}$$

The logic behind this model is the same as in the case of the equation for each type of clothes separately, therefore we run the same regression with the only change being that we use the share spent on all three types instead of just one, and the number of members of the household h of all types instead of just members of type t.

We see that an estimate of b_h is in fact the denominator in the equation for determination of resource shares (3.12). This not being equal to zero is our identification restriction in this model.

First, we test if b_h is equal to zero the usual way, using the expected value of this slope parameter at mean values of covariates as the test statistics, with the null hypothesis being that b_h is zero and alternative that it is not. Second, we test this for every observation (every household h) included in the dataset, using corresponding covariates for each observation. If the percentage of the households satisfying this condition is larger than 75%, then the identification condition is satisfied as well and we can use this methodology.

If the data fails the test, then this methodology cannot be used for this dataset and we should search for other possible ways to detect intra-household inequality or consider if the data availability could be the source of the problem.

Chapter 4

Data

In this section, we present the necessary treatment done with the dataset and a few potential problems arising while processing the data. First, we give a brief description of the dataset used, The World Bank General Household Survey. This dataset is used since we continue to follow the research done by Lechene et. al (2022) on this topic and it contains the necessary data on assignable good.

4.1 The World Bank General Household Survey

We use a dataset collected by The World Bank, General Household Survey - Panel (GHS) for the years 2018-2019. The GHS contains detailed data on household expenditure as well as demographic variables included in the model.

The World Bank collects this particular data to obtain statistics on agricultural performance in several countries in Sub-Saharan Africa as a part of the Integrated Surveys on Agriculture (ISA) program. The GHS data exist for more than 40 countries with several rounds for each.

We rely on the 4th wave of this data collection, which is so far the latest one available. The previous rounds were collected in the years 2010-11, 2012-13, and 2015-16. Lechene et al (2022) for their paper used 2012-2013 round for altogether 11 countries, including Nigeria. Therefore we can compare the results obtained in this thesis for the round 2018-2019 to the round 2012-2013.

The dataset contains data about approximately 5000 households from all 36 states of Nigeria and the Federal Capital Territory. It includes information on household expenditure on all kinds of goods the household consumes together with data on housing, education, labor, household assets, and more. There are also data on community levels focusing on prices and labor in the area and specific data on agricultural conditions included, however these are not needed for our research.

The data were collected twice during the period 2018-2019. First, after the planting period in 2018, which is between July and September. Second, after harvest in 2019, which is between January and February. We combine information from both waves in the following way. The assignable expenditure on clothes was captured only in the first collection. However, the data on individual education levels, needed as one of the covariates for the model, is contained only in the post-harvest collection. Therefore we use data on all kinds of household expenditure and most of the demographic characteristics from the post-planting part of the dataset and data on education levels from the post-harvest. We merge it by household ID numbers assigned to individual households included in both parts of the dataset.

4.2 Variables

We base our analysis on the following variables, the dependent variables of the model (the share of the household budget allocated to expenditure on clothes of each type "t"(household level Engel curves)), the independent variables (the overall household budget, the number of household members of each type living in the household). Additionally, we include the covariates, which are dummy variable indicating if the household lives in an urban area or not, the average age of each type of household member present in the household, the minimum age of children if present, and the education level of man and woman living in the household. We also include a few more variables for analysis of the expenditure behavior of the households.

For the model we described in the previous section, we require the data in cross-sectional form with one cross-sectional unit being one household labeled by the household ID ("hhid") and include all the variables necessary for the model. Below we provide a description of the variables, which require more effort during the data processing. These are education and household expenditure.

4.2.1 Education

We use the highest education of the household head and their spouse as an indicator of the education level within the household. Lechene et. al (2022) use two education variables as covariates for the model, one for men and one for women living in the household. If the head of the household is a man, we take the education of the head as an indicator of male education, if it is a woman, we take it as female education in the household. We do the same for the spouse. Therefore if the head and spouse are of opposite gender, we get complete data on the education of both genders within a given household. As there are no households with the household head and their spouse of the same sex, we get education levels for all households for which we have the information about the education of the head and the spouse available.

This way we got higher data coverage than by using within-household averages of education for both genders as Lechene et. al (2022) did. However, the data on education are still highly incomplete. We exclude all variables with missing values of education. Before this exclusion, we had altogether 5148 observations. The data coverage for education was only approximately 22% and after this filtering, we got only 1126 observations.

The World Bank provides the variables indicating education levels as codes representing individual education types, starting with elemental education levels like nurseries, primary and secondary schools, but also several quite specific education types such as nursing school, teacher training, or Quranic education. The codes have values spread from 0 for no education to over 400.

Since we want to examine descriptive statistics of variables these values are not useful, therefore we replace them with a sequence of numbers from 0 to 6, where 0 stands for the lowest education levels and 6 for the highest levels. We include people with preschool education only in the 0 group together with individuals with no education at all. Value 1 stands for primary education, 2 for junior secondary 3 for senior secondary, 4 for higher education following secondary, but without a bachelor's degree, 5 for bachelor degrees, and 6 for postgraduate education.

4.2.2 Household expenditures

Household expenditure contains two types of expenditures, food and non-food. We get the overall expenditure by summing the expenditure on food and non-food goods.

The household non-food expenditure is separated into 3 groups based on the time horizon of their consumption, depending on the characteristics of the individual goods. The value provided in the dataset stands for the household expenditure on a particular good during the last 7 days, 30 days, or 6 months. Expenditure on only several items such as tobacco or gambling is captured over the last 7 days. Mostly regular monthly payments such as electricity are captured by 30 days recall. Expenditure on less frequent purchases and most of the consumer goods such as clothes, books, or electronics is shown over the last 6 months.

We have to choose one time horizon and transform the data so that we get the consumption over the same time unit for all goods. We chose the 6-months recall, therefore we have to multiply the other expenditures by 6 (for 30 days recall) or 26 (for 7 days recall).

Expenditure on food in this dataset is given by expenditure on the most recent purchase of the good, which we take as representative for a week, meaning that we assume that the purchase of food is done once in a week. Therefore we multiply this also by 26 to get an approximate expenditure during half a year. Then we sum it up with the non-food expenditure during the same period.

An important part of the work with the data is the computation of the shares of the overall expenditure spent on clothes. We sum the expenditure on tailored and ready-made clothes with the expenditure on shoes for each type t. Expenditure on clothes is part of the 6 months recall, so we leave it as it is without any multiplication. We divide each of those 3 expenditures we get (on men's, women's, and children's clothes) by the overall expenditure to get the share allocated to this kind of household spending.

4.3 Household compositions

The model we introduced in the previous section has to be estimated separately for different compositions of households. This is necessary because there are different dependent and independent variables and covariates needed for each regression.

We consider altogether four compositions of households. These are households including all three types of members (mfc), households including men and women (mf), households including men and children (mc), and households including women and children (fc). It does not matter how many members of each type the household includes as long as there is at least one of each type belonging to a given composition.

We drop observations not belonging to one of these composition groups we established. For example, a household, which includes a single person or more people of only one type. For this kind of household, it is not meaningful to inspect inequality between shares of expenditure allocated to each type of member as there is only one type living in the household.

We form 4 separate datasets, each including only households belonging to one type of household composition. Then we run the regression on each of the 4 samples. The table 4.1 below shows the number of households in each sample.

household composition	mfc	mf	mc	fc	sum
number of observations	964	161	0	0	1126

Table 4.1: Number of observations for each household composition

In the table, we see that for 2 out of the 4 compositions, mc, and fc, we have no observations at all. The most common type of household composition in our dataset is a household containing all three types of household members. We follow Lechene et al. (2022) and set the boundary of a minimum of 100 observations in the sample to be enough for the estimation to be meaningful. Both samples we obtained satisfy this condition, therefore we run the regression for these two compositions, mf, and mfc.

4.4 Data description

Before running the regression and analyzing the results, we inspect the data in detail, focusing primarily on household expenditures.

The table 4.2 below shows the average share of household expenditure in Nigeria spent on clothes of all three types, food and housing, and the standard deviation of the distribution of those variables. As housing expenditures, we take the sum of rent and mortgage payments. We include statistics about food and housing since we expect the proportions of household budgets on these items to be high as they are necessities for survival. The same statistics are shown for the number of members of each type living in the household, their age, and education of men and women in the table as well.

Households spend on average 1% of their resources on both, men's and women's clothes. These two variables have also approximately the same standard deviation, which is high compared to the mean. On children's clothes, households spend more, which is not surprising taking into account that the average number of children in households is also higher and we do not divide them by gender.

Table 4.2: Summary Statistics

	Mean	(SD)	Mean mfc	(SD mfc)	Mean mf	(SD mf)
c clothes	0.02	(0.03)	0.03	(0.04)	0.01	(0.02)
f clothes	0.01	(0.02)	0.01	(0.02)	0.02	(0.02)
m clothes	0.01	(0.02)	0.01	(0.02)	0.01	(0.02)
age c			8.31	(3.22)		
age m	41.57	(11.86)	40.47	(10.55)	48.20	(16.34)
age f	36.10	(9.97)	34.78	(8.80)	44.03	(12.59)
number of c	2.89	(2.26)	3.37	(2.09)		
number of m	1.73	(1.04)	1.70	(1.00)	1.95	(1.21)
number of f	1.80	(1.03)	1.83	(1.05)	1.66	(0.87)
education m	2.59	(1.41)	2.60	(1.38)	2.57	(1.61)
education f	2.23	(1.25)	2.21	(1.21)	2.36	(1.49)
food	0.60	(0.22)				
housing	0.00	(0.01)				
number of h	1126		964		161	

We know that the household head and their spouse are of the opposite sex for all the observations included, however, the average number of adult household members of both types is close to 2 with a standard deviation of 1. This can be partially caused by the fact that the top boundary of being a child is set at the age of 15 and the children continue living with their parents up to a higher age or the households include 3 generations, not just 2 (grandparents). The minimum age of women in the dataset is 16, which would correspond to the first theory. For men, it is 20, which is not that close to the boundary age of 15. This difference can be caused by the low data availability mentioned, or men may move away from their parents' house earlier and start new households on their own. The highest age is 75.5 for women and 98 for men. Together with a higher average age of men this sign that men in Nigeria live longer.

Also over 80% of heads of households are men. That could mean that men have a higher bargaining power within a household and that is why they might consume higher resource shares within the household and therefore live longer.

Men on average achieve a higher education level, however, the difference is not even 1 point of the scale we set. For both genders, the average education achieved is between junior and senior secondary school.

After summarizing the data on controls, we move to the variable we are primarily interested in, expenditure on clothes.

When focusing on the other expenditures, we can see that the households spend not even 1% of their budget on housing. This does not correspond with our expectations. It might indicate that most of the households live in their own houses or apartments, as they do not spend that much on rent and mortgage payments, however, since there are on average more than 2 adults in the households, as noted these could also be multi-family households.

Food expenditures range from 0 to over 99% of the overall household expenditure with a mean of 60% and a standard deviation of 22%. This is in line with the expectations mentioned. It shows that most of the households included in the dataset allocate over half of their budget to food expenditure.

We also showed the difference between the two samples we ended up with in the table 4.2. This comparison is useful mainly for variables considering children.

The "number of children" variable in the whole dataset has a high standard deviation compared to its mean. The average household there has between 2 and 3 children, however, also the sample of households without children, which makes almost 10% of the whole dataset, is included. The average number of children, if we drop the households without children and look at the composition mfc only, the average increases by approximately 0.5 from 2.89 to 3.37 children per household with an average age close to 8. Expenditure on children's clothes is also noticeably higher in the households belonging to composition mfc.

The ages of both types of adult members are on average higher in households without children. For women, this difference is in round numbers 9 years, and for men 8 years. A possible explanation is that the households of composition mf are formed more frequently by families with children over 15 years old or older couples, whose children already left the households, than by young couples, who do not have children yet.

We also observe a difference between education levels in the two samples. The averages of education levels of men and women are closer to each other in households of composition mfc than mf. However, this difference is only minor considering the low data availability.

We see the opposite trend with the number of men and women. In households without children, there are on average more men and fewer women. That goes in line with the findings about higher average ages in this kind of household and higher ages of men compared to women in the whole dataset. These facts might support the hypothesis that men here live longer mentioned earlier.

Chapter 5

Results

We applied the methodology introduced by Lechene *et al.* (2022) described in Chapter 3 to the two datasets for two separate kinds of household compositions we got after performing the necessary filtering of the data. In this chapter, we will present the results that we have obtained. We start with the result of an identification test.

5.1 Test of model identification

Lechene *et al.* (2022) have run the regression for 12 countries, and for just 5 of the 12 datasets, the model has passed both of the identification tests, testing if the denominator in the equation 3.12 for calculation of the resource shares is statistically different from zero at the mean value of covariates and second if the slope is statistically significant for at least 75% of the observations, using particular covariates for each. Nigeria did not pass the second of the tests. Consequently, Lechene *et al.* (2022) did not present the regression results.

We used a dataset collected six years later than the one used by Lechene *et al.* (2022). Using our data, the model passed the first test for both of the household compositions we considered. The slopes of the Engel curve for all types of clothes are statistically different from zero. They will be discussed in detail in the next section. We did not conduct the subsequent test, so the results should be interpreted cautiously. To be more precise, we would also need to perform the second test.

5.2 Engel curves slopes

As a part of the test, we get the estimated slope of the Engel curve for all types of clothes households of a given composition consume together at the mean value of covariates.

The estimated slopes are -0.135 for households without children and -0.019 for those with all three types of household members. As both slopes are negative, it suggests that clothing is considered a necessity within the Nigerian population sample we have observed.

The table 5.1 below displays the slopes of the Engel curves and the budget shares spent on all types of clothing combined for households comprising only men and women (MF) and for households including children as well (MFC). The table compares these figures to those obtained by Lechene *et al.* (2022). Note that we used data from years 2018-2018 Lechene *et al.* (2022) used data from 6 years before.

	Lechene *et al.* (2022)	MF	MFC
Engel curve for clothes slope	- 0.02	- 0.135	- 0.019
Budget share spent on clothes	1.7%	4%	5%

Table 5.1: Engel curves slopes and budget shares spent on clothes - comparison to Lechene *et al.* (2022)

The slope estimated by Lechene *et al.* (2022) is higher (closer to zero) than the estimates we obtained for both household compositions. While we can observe that the slope has decreased over the six years between the data collections, the budget share spent on clothes has increased. This means that households in 2019 spent more on clothes, and the share spent on it decreases faster with an increase in the budget compared to 2013. This suggests a slight change in the consumption behavior of Nigerian households. Engel curves generally tell us that families that are poorer spend a higher share of their budget on necessities. For example, Engel (1857) shows this phenomenon in food consumption. It is possible that clothing is becoming more of a necessity in Nigeria as poverty rates are still high, as the Bank (2023) states around 40% in 2018, and households are spending more on clothes. However, the changes in both Engel curve slopes, and budget shares spent on clothes are relatively small, the time horizon is short, and the sample is also not large. Therefore, more data would be necessary to draw valid conclusions.

It is important to note that the size of the sample we consider is 1126, while Lechene *et al.* (2022) worked with over 3500 observations. Therefore their results might be more precise.

5.3 Resource shares

The results of the regressions can be found in table 5.2. The estimated resource share allocated to one person of each type t, $\hat{\eta_t}$, is shown separately on the left for households including only men and women (composition MF), and then on the right for households including all three types of household members (composition MFC). Additionally, the table includes the standard errors (SE) of the estimates, as well as the minimum (Min) and maximum (Max) resource share within the sample.

	Composition MF				Composition MFC			
	Mean	(SE)	Min	Max	**Mean**	(SE)	Min	Max
$\hat{\eta_f}$	**0.441**	(0.108)	0.147	0.908	**0.215**	(0.049)	- 0.058	0.592
$\hat{\eta_m}$	**0.255**	(0.101)	- 0.165	0.675	**0.08**	(0.053)	- 0.187	0.329
$\hat{\eta_c}$					**0.222**	(0.042)	0.008	0.726

Table 5.2: Results of the regression

First, note that the resource shares do not add up to one for either of the household compositions, as stated in equation A.6. This occurs for several reasons. Firstly, the estimated resource shares are always presented for one household member of a specific type t, even though households may include multiple members of each type. For example, if one man consumes 8% of the household's resources and there are three men in the household, they consume together 24% of the budget. If we sum the resource shares like this for all three types of household members the resource shares should sum up to one for individual households. However, another problem is that some of the resource shares have negative values (as seen in the Min values). As explained by Lechene *et al.* (2022), the reason for estimates falling outside the interval

(0,1) is that the assignable good of interest is a necessity for one type of household member, and a luxury for another, causing the Engel curves for assignable goods of different types to have opposite signs, leading to negative estimated resource shares for certain types. When some resource shares are negative, they do not make logical sense and they might also cause that the sum of resource shares is not equal to one.

Now we discuss the estimates of the resource shares obtained. We observe that households consisting solely of men and women tend to allocate more resources to women on average. The estimated share of resources consumed by women is 44%, whereas for men, it is slightly above half of that at 25.5%. This results in a gender gap of 18.6 percentage points in favor of women for this kind of household composition. Also, both estimates of resource share are statistically significant at a 5% significance level.

There is not much difference from the first composition in the allocation of resources for households with all three member types (men, women, and children). On average, in these households, the highest proportion of resources, around 22.2%, is allocated to children. The estimated share of resources consumed by women decreased by 22.6 percentage points to 21.5%, and for men, it decreased by 17.5 percentage points. Therefore, the men in this type of household composition are estimated to be left with only 8% of the household's resources each, which is unusually low. However, it is important to note that the estimated resource share for men is not statistically significant, while the shares for women and children are. The gender gap narrowed compared to households without children to 13.5 percentage points.

The obtained results are quite unusual from several points of view. The size of the estimated resource shares itself is surprising. For children, the estimate of 22% per child is slightly above the estimated resource shares found by most of the studies mentioned so far. Out of the estimates presented by Lechene *et al.* (2022) in other countries, which we consider as suitable for comparison mainly since the methodology used is the same, the closest one is the estimated resource share allocated to children in Bulgaria, which is 19%.

A more standing-out finding is the gender gap for both types of household composition. The first phenomenon worth pointing out is the size of the gender gap itself. Women, in the sample we consider, consume extraordinarily large resource shares compared to men even in households with children. This study reveals the largest gender gap in favor of women among all the research cited so far. In cases where a gender gap in favor of women was present, its magnitude was typically small, making this particular gender gap exceptional, especially in developing countries.

The largest gender gap favoring women, from the research mentioned, was estimated by Bargain & Donni (2012) in France for the year 2000. The size was 10 percentage points in households with no children. However, in households with one child, a slightly larger share was allocated to men again. The same trend was observed by Bargain *et al.* (2022) in the UK. But even in these developed countries, the gender gap favoring women was not that large and did not occur in households with children.

When we look at the results obtained by Lechene *et al.* (2022), once again the closest result is in Bulgaria, where the only gender gap in favor of women is present in this paper. The gap is equal to 7 percentage points in households, including all three household member types. The resource shares estimates for Bulgaria are 30% for men and 37% for women, indicating that it is still a very different case. For the rest of the countries studied by Lechene *et al.* (2022), more resources are allocated to men, with the largest gender gap being 4 percentage points.

In Chapter 2, we observed that the gender gap favoring men is becoming narrower, almost disappearing, in the case of Malawi. This trend could signal that the country is progressing toward a better consumption position for women within households. The results we have for Nigeria might indicate a similar trend, although the numbers are still considerably extreme compared to those of Malawi.

Another unusual trend considering the gender gaps we obtained is the change between the household composition types. As was already suggested, women's resource shares usually tend to decrease with an increasing number of children, and therefore, the gender gap moves towards favoring men (see, for example, Bargain *et al.* (2022) or Dunbar *et al.* (2013)) For our estimates, this is not straightforward. Even though the difference between resource shares consumed in households that include children, compared to those without them, is higher for women, it is important to note that men have sacrificed over two-thirds of their resource share, while women have sacrificed just half.

In the previous research, the women are the ones who sacrifice more of their consumption to children, and the gender gap tends to be wider for households with children. For example, the case of India introduced by Calvi (2020) was mentioned, where he found that the gender gap was 6 percentage points in households without children and 15 percentage points in households with children, both in favor of men. The women in the dataset he observed faced 15 percentage points lower resource shares in households with children compared to those without them. For men, the decrease was only 6 percentage points.

Multiple possible explanations for such unexpected results exist. First is the scarcity of data. As already stated, the dataset includes only 1126 households of two kinds of composition. Higher data availability could provide a more complex picture of consumption behavior. Second, is the complexity of household compositions mentioned already in Chapter 4. It is difficult to determine whether the presence of more adults of one gender in a household is due to multiple generations living together, older children still living with their parents or other factors. This uncertainty could result in misclassifying children's consumption. Additionally, it is possible that children are wearing adult clothes or vice versa. Another potential issue is the absence of the second test, which may indicate that the model is not identified for this dataset. Finally, it is possible that households' consumption behavior in Nigeria really changed; however, we still need more data and testing to confirm that.

Chapter 6

Conclusion

In this thesis, we aimed to investigate inequality among household members in Nigeria and compare our findings with those of Lechene *et al.* (2022). We used data collected by the World Bank from 2018 to 2019. We obtained the necessary variables based on the model introduced by Lechene *et al.* (2022) and split the data into four datasets according to household composition. However, two of these datasets had no observations. As a result, we ended up with 964 observations for households consisting of men, women, and children and 161 observations for households with only men and women.

Similar to Lechene *et al.* (2022) for Nigeria, our data passed the first stage of the identification test. While the data failed the second stage for Lechene *et al.* (2022), we did not carry out the second stage and proceeded to analyze the results as they were. This omission could lead to inaccurate results.

The Engel curve for all types of clothing showed a negative slope for both household compositions, indicating that clothing is considered a necessity in Nigeria. We also observed a slight shift in household consumption behavior since 2012/2013, with households spending more on clothing and the slope of the Engel curve moving further from zero. These trends suggest that clothing is becoming even more of a necessity for households in Nigeria.

The results of the resource share estimation provided an unexpected insight into the distribution of resources within households. Resource shares allocated to women were higher for both types of household compositions. Surprisingly, men's resource shares were low, and the gender gap shift between the two household compositions differed from the literature we discussed.

These findings suggest that women's position in the distribution of household bargaining power is shifting for the better and has important implications for inequality and poverty situations in the developing world. As stated in the opening of the thesis, Nigeria, as cited by World Bank (2024b), Nigeria is Africa's largest economy and population. Therefore, the situation here may be more favorable compared to the rest of the continent. However, from its position, Nigeria might be leading the way in improving women's status within households and potentially in society as a whole.

However, it is important to note that our results face several limitations in form for data scarcity or statistical significance. These problems can be possibly both solved by increased data collection.

If the suggested shift is really happening, it could outline trends in inequality and poverty rates that are not currently captured by statistical offices or international organizations. Therefore, it is important to conduct further research. Obtaining these estimates for more countries across Africa can confirm or disprove the hypothesis that the position of women in African households is improving. To enhance this research and its reliability, a higher supply of data is required. The collection of data on person-level consumption of clothing, food, or other assignable goods by individual household members is the foundation for the research and its reliable results. From this, conclusions about the current situation of within-household inequality in Africa, and potentially the whole world, could be made.

The comparison of bargaining power distribution within households in developed and developing countries is another reason to increase research on intra-household inequality. This research could help reduce inequality and improve quality of life in the developing world.

Bibliography

Alderman, H., P.-A. Chiappori, L. Haddad, J. Hoddinott, & R. Kanbur (1995): "UNITARY VERSUS COLLECTIVE MODELS OF THE HOUSEHOLD: IS IT TIME TO SHIFT THE BURDEN OF PROOF?" *The World Bank Research Observer* **10(1)**: pp. 1–19.

Almås, I., C. Ringdal, & I. H. Sjursen (2020): *Understanding Inequality Within Households*, pp. 1–27. Cham: Springer International Publishing.

Bank, W. (2023): "Poverty & equity brief: Nigeria, africa western & central." *Technical report*, World Bank.

Bargain, O. & O. Donni (2012): "Expenditure on children: A rothbarth-type method consistent with scale economies and parents' bargaining." *European Economic Review* **56(4)**: pp. 792–813.

Bargain, O., O. Donni, & I. Hentati (2022): "Resource Sharing in Households with Children: A Generalized Model and Empirical Evidence from the UK." *Journal of the European Economic Association* **20(6)**: pp. 2468–2496.

Browning, M., P.-A. Chiappori, & A. Lewbel (2013): "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power." *The Review of Economic Studies* **80(4)**: pp. 1267–1303.

Calvi, R. (2020): "Why are older women missing in india? the age profile of bargaining power and poverty." *Journal of Political Economy* **128(7)**: pp. 2453 – 2501.

Chai, A. & A. Moneta (2010): "Retrospectives: Engel curves." *Journal of Economic Perspectives* **24(1)**: p. 225â€“40.

Chiappori, P.-A. (1988): "Rational household labor supply." *Econometrica* **56(1)**: pp. 63–90.

Dunbar, G. R., A. Lewbel, & K. Pendakur (2013): "Children's resources in collective households: Identification, estimation, and an application to child poverty in malawi." *American Economic Review* **103(1)**: pp. 438–71.

Dunbar, G. R., A. Lewbel, & K. Pendakur (2021): "Identification of Random Resource Shares in Collective Households Without Preference Similarity Restrictions." *Journal of Business & Economic Statistics* **39(2)**: pp. 402–421.

Engel, E. (1857): "Die productions-und consumtionsverhältnisse des königreichs sachsen." *Zeitschrift des Statistischen Bureaus des Königlich Sächsischen Ministeriums des Innern* **8(9)**: pp. 1–54.

Lechene, V., K. Pendakur, & A. Wolf (2022): "Ordinary least squares estimation of the intrahousehold distribution of expenditure." *Journal of Political Economy* **130(3)**: pp. 681–731.

Lise, J. & S. Seitz (2011): "Consumption Inequality and Intra-household Allocations." *The Review of Economic Studies* **78(1)**: pp. 328–355.

Penglase, J. (2020): "Consumption Inequality Among Children: Evidence from Child Fostering in Malawi." *The Economic Journal* **131(634)**: pp. 1000–1025.

StataCorp LLC (2021): "sureg – seemingly unrelated regression." In "Stata 17 Base Reference Manual," pp. 2888–2896. Stata Press, release 18 edition.

World Bank (2023): "GDP per capita (current US$)."

World Bank (2024a): "Poverty overview." `https://www.worldbank.org/en/topic/poverty/overview`. Accessed: 2024-06-11.

World Bank (2024b): "The world bank in nigeria." https://www.worldbank.org/en/country/nigeria/overview. Accessed: 2024-07-21.

Appendix A

Derivation of the formula for calculation of the resource shares

We have the form of the household-level Engel curves for assignable good of type t.

$$W^t(y) = \eta^t[\alpha^t + \beta(\ln y + \ln \eta^t - \ln N^t)] \tag{A.1}$$

If we multiply by the resource share of type t, we can see the slope of this Engel curve clearly.

$$W^t(y) = \eta^t\alpha^t + \eta^t\beta(\ln y + \ln \eta^t - \ln N^t)] \tag{A.2}$$

The intercept parameter is $\eta^t\alpha^t$, the variables are $\ln y$, $\ln \eta^t$ and $\ln N^t$) and the slope parameter is the resource share times the slope of the individual-level Engel curves, $\eta^t \times \beta$.

After the estimation of the Engel curves we get the estimated slope of all three types of household-level Engel curves. We denote the slope of the household-level Engel curve for an assignable good of type t as b^t and its estimate $\hat{b}^t$. Therefore we have:

$$\hat{b}^f = \eta^f \times \beta \tag{A.3}$$

$$\hat{b}^m = \eta^m \times \beta \tag{A.4}$$

$$\hat{b}^c = \eta^c \times \beta \tag{A.5}$$

We also know that the sum of resource shares is equal to one.

$$\sum_{t=1}^{T} \eta^t = 1 \tag{A.6}$$

The equations A.3, A.4, A.5 and A.6 form the set of four equations.

From the equations A.3, A.4, A.5 we have that the resource share of type t is equal to the estimated slope of the household-level Engel curve for assignable good of type t, which differs for each type, divided by the slope of individual-level Engel curves, which is the same for all three types.

$$\eta^t = \frac{\hat{b^t}}{\beta} \tag{A.7}$$

After substituting this into equation A.6 we get:

$$1 = \frac{\hat{b^f} + \hat{b^m} + \hat{b^c}}{\beta} \tag{A.8}$$

Therefore β is equal to the sum of the slopes of household-level Engel curves for all three types of assignable goods.

$$\beta = \hat{b^f} + \hat{b^m} + \hat{b^c} \tag{A.9}$$

If we substitute this form for β in A.7 we get the formula for calculating the estimated resource shares of each type t after getting estimates of the household-level Engel curves.

$$\eta^t = \frac{\hat{b^t}}{\sum_{t=1}^{T} \hat{b}^t} \tag{A.10}$$

Appendix B

Parameters of the Engel curves after linear reframing and approximation

The parameters of the Engel curves after the approximation are already easy to estimate and linear:

$$a_h^t = a_0^t + a_{\ln N^t}^t \ln N_h^t + a_z^{t'} z_h \tag{B.1}$$

and

$$b_h^t = b_0^t + b_z^{t'} z_h \tag{B.2}$$

If we put these equations together into one, substituting for a_h^t and b_h^t in the household-level Engel curve, we can see the model for estimation:

$$W_h^t = a_0^t + a_{\ln N^t}^t \ln N_h^t + a_z^{t'} z_h + b_0^t \ln y_h + b_z^{t'} z_h \ln y_h + \epsilon_h^t \tag{B.3}$$