Gender Difference in Indian Consumption Expenditure

A thesis submitted to Indian Statistical Institute
in partial fulfillment of the thesis requirements for the degree of
Doctor of Philosophy in Quantitative Economics

Author:
Mannu DWIVEDI

Supervisor:
Prof. Manoranjan PAL

Economic Research Unit
Indian Statistical Institute
203, B. T. Road, Kolkata,
West Bengal, India - 700 108.

Dedicated to

My Parents

“Gender equality is more than a goal in itself. It is a precondition for meeting the challenges of reducing poverty, promoting sustainable development and building good governance”

— Kofi Annan

Acknowledgements

I would like to express my deepest gratitude to my academic supervisor, Prof. Manoranjan Pal, for his enthusiasm, guidance, and unrelenting support in this great journey and the preparation of this work. He has routinely gone beyond his duties to help me overcome my worries, concerns, and anxieties, and for instilling great confidence in both myself and my work. I am especially thankful to him for the work we have done on the topic of gender difference which was an inspiration to me.

I would also like to extend my appreciation for the invaluable suggestions and support to Prof. Amita Majumder, Prof. Ishita Mukhopadhyay, Dr. Samarjit Das, Prof. Nityananda Sarkar and Prof. Indraneel Dasgupta. Their influence and encouragement kept my work and this thesis in a proper direction.

My sincere appreciation is extended to the administration of the Indian Statistical Institute, Kolkata for funding my research by providing me with a research fellowship and the necessary infrastructure for doing this work.

I also want to acknowledge the encouragement, influence, input, support, and expert knowledge of all the research fellows, in my department and others, who came out of their way to make this journey memorable to me. The names I will never forget, Somnath Chattopadhyay, Trishita Roy Barman, Conan Mukherjee, Sattwik Santra, Debasmita Basu, Srikanta Kundu, Sandip Sarkar, Kushal Banik Chowdhury, Priya Brata Dutta, Rajit Biswas, Debojyoti Mazumder, Parikshit De, Chandril Bhattacharyya, Tanmoy Das, Mahamitra Das, Arindam Pal, GopaKumar A C, Aditya Bhan, Chayanika Mitra, Dripto Bakshi, Ranajoy Guha Neogi, and Soumyarup Sadhukhan. Not to mention those whose names I have not mentioned here.

Special thanks to Nazma Patel, Nitu Jaiswal, Sameen Zafar, Babita Yadav, Shashank Singh and Sedighey Mirza who believed in me when I was losing belief in myself and for giving me the company I need in the difficult times. Their encouragement and emotional support kept me motivated throughout this research.

I would also like to extend my deepest gratitude to my family. Without their encouragement, I would not have dreamed of finishing a research like this at Indian Statistical Institute.

Finally, I would like to express my love and gratitude to my partner, Nitin Aggarwal, whose unremitting encouragement prevented me from all the potholes in this great journey. He proofread whole of this thesis and provided many valuable suggestions. Thank you for being there throughout this difficult journey.

Kolkata, Feb 16, 2018 — Mannu Dwivedi

Contents

List of Tables

List of Figures

1 Gender Difference –A Brief Survey of Literature

1.1 Introduction

The goal of any development policy is to increase the living standard of the people in the society. In any developing country, the efficiency of such policy depends on how families reallocate the resources among themselves. If the distribution is even within the family then only it can be said that the policy was fruitful. Therefore, a family is an important part of any policy and the allocation within the family should be just. But, the distribution of the resources among the family members are not always equal. Inequality always exists! Discrimination against girls or women persists in approximately all the developing countries. The attitude towards women and the progress of a country both socially and economically are directly linked. The status of women is central to the health of the society. If one part suffers so does the other (Moser (2012)). Therefore, the issue of gender discrimination, over the last several decades, have attained increased prominence in the debates over development policy. Also gender awareness links policy and projects to equitable, efficient, and sustainable development. Sometimes it is also seen that not only females but also males are discriminated against in the resource allocation for some resources (see Fuwa (2014)).

India is a vast country with diverse cultural and ethnic groups. It is emerging as a major global market, and it is the sixth-largest economy in terms of the nominal GDP. The average growth rate is approximately 7% over the last two decades. Although it is doing exceptionally good at the macro level, one cannot ignore the major drawbacks such as poverty, gender inequality, pollution, income inequality, unemployment, poor educational standards, poor infrastructure, inefficient agriculture, inequality within regions etc. Out of all these problems, the gender difference in intra-household resource allocation of resources is the most prominent yet mostly unnoticed as it is present in the very foundation of the society i.e. how the basic human necessity depends on the perception of the society that one gender is more entitled than the other. For most of the households, a major part of their income goes towards the food expenditure which includes fruits, vegetables, milk, spices, cereals and pulses, etc. and this expenditure is important for the health of the members of the households. There are instances that reveal gender difference in the food expenditure (Lancaster et al. (2008)). Also, gender difference has been observed in the areas of health, education, consumption, labour market outcomes etc., all over the world.

To get a more clear picture of the gender issues in India we can look into a few indices and compare the neighbouring countries with ours and look for the extent to which there is discrimination. These indices help us to understand the respective position and condition of the nation with respect to others. One of these measures is 'Gender Inequality Index value'. This 'Gender Inequality Index' is a composite measure reflecting inequality in achievements between women and men in three dimensions: reproductive health, empowerment and the labour market. According to this measure, India has a value of 0.563 and rank 127, according to the Human Development Report, 2013. For the entire world, this value is 0.450. There is another measure known as 'Gender-related development Index' (GDI). According to this measure, India has rank 132 in 2013. This may be compared with the ranks of the neighbouring countries like China(88), Pakistan (145), Bangladesh

(107) and Srilanka (66). On top of this female to male ratio in 2013 was 0.828. Lastly, according to Gender Gap Index, India's rank was 114 in 2014, and in 2013 it was 101. From these indices, it is clear that India is already in a state of an alarming situation and we cannot ignore the gender issues completely if we have to pursue our development objectives.

To further highlight the issue in another dimension, this work is devoted to the study of gender difference present at the very foundation of the Indian society i.e. within the families. The intra-household allocation of resources among family members is a difficult thing to study mainly due to data issues and it is rarely done with the target to understand how the perception of the society governs which gender is entitled what and which is not? This gender difference persists through generations as children observe the biasedness present in their surroundings and grow to learn to be biased in such environment. If one can implement policies to control this, then only the policy can help to eradicate this from the grass-root level.

The thesis opens up a long forgotten regression model, which not only enables us to find out within house gender bias in consumption pattern but also among different demographic groups based on age and sex. It also gives a semiparametric form of Engel curve which is more general than parametric curves taken so far for the purpose of finding gender differences. Such elaborate and minute level studies have not been made so far, especially for regression decomposition method. Regression decomposition technique provides us with a method to decompose household level information into the individual level as National Sample Survey(NSS) provides data at household level on consumption expenditure. It does not give individual level consumption expenditure data. This regression decomposition approach enables one to investigate gender discrimination at the individual level within households. Also, item-wise gender discrimination has only been studied by Subramanian and Deaton (1991) who only considered rural Maharashtra. The result of the study opened up a few insight into the situation of rural Maharashtra. This thesis is to scale up the whole result for the entire country and that's why it should contribute to the existing literature.

In this chapter, a brief review of the existing literature on empirical studies on gender bias, especially in the contexts which are prominent in India are discussed in Section 1.2. The objective of this work is discussed in Section 1.3. Finally, the format of the thesis is given in Section 1.4.

1.2 A Brief Review of the Literature

In India, discriminatory attitude towards men and women have existed for generations and affected the lives of both genders. Although the Constitution of India has granted men and women equal rights, gender disparity still remains. All too often, women are discriminated against in health, education, consumption, labour market outcomes etc. The issue of gender difference is prominent not just in India but also in a number of other countries as is pointed out by the literature. The literature of gender difference is vast and covers dimensions like education, health care etc. Let us first observe the literature in a few of these dimensions.

Gender discrimination in educational expenditure are addressed by many authors like Kingdon (2002, 2005), Zimmermann (2012), Lancaster et al. (2008), Himaz (2010) etc. Kingdon (2005) used the Engel curve approach to find gender bias in educational expenditure although it often fails to detect the gender bias in previous works [such as Deaton (1997) and Case et al. (2002)]. She tried to search for the reason why this approach often fails to detect the biasedness. Kingdon (2005) finds out that the enrollment rates for girls are significantly lower than the boys in the Indian States. For those states where there are lower enrollment rates there exists lower educational expenditure on girls. But among the enrolled children, very little gender bias is found. She tested both types of data i.e. individual level as well as household level. The household-level expenditure data fails to show any gender discrimination. Her tests show that Engel curve approach fails partly because of its incorrect functional form. Lancaster et al. (2008) uses a three-stage least square(3SLS) estimation techniques for a set of system equations. They found evidence of gender bias in adult consumption of several items. Mostly education expenditure in the backward regions of rural India shows significant gender bias. For rural Kerela, they actually found pro-girl bias in education expenditure in the age group 10-16 years. They observed that in the backward areas of eastern India the gender bias exists in favour of boys in educational spending among the illiterate households, but weakens or disappears as one move to literate households, suggesting a vicious cycle that keeps the female child in the former households in a state of perpetual educational backwardness and illiteracy. Zimmermann (2012) uses the Engel curve approach to detect gender bias in education expenditure using India Human Development Survey(IHDS) from 2005. She found strong results of gender discrimination in educational expenditure for aged 5-9 years and the discrimination increases with age, leading to pretty widespread gender bias once children reach 15-19 years of age. She concludes that Engel curve does not always fail to detect gender bias. When data is large and discrimination is strong, the problems that are usually encountered in small data cases like data aggregation and usage of different statistical models become irrelevant.

Gender discrimination in health-related outcomes like excess female mortality and morbidity are considered by authors like Arnold et al. (1998), Gupta (1987), Miller (1997) etc. Neglect of females in terms of health care and nutritional status has been taken up by the authors like Mishra et al. (2004), Miller (1987), Harriss (1989), Pande and Malhotra (2006), Behrman (1988) etc. Many authors try to find gender discrimination against girls with respect to other issues as well. For example, gender discrimination in the labour market is very often discussed by the economists. Starting from labour force participation to the wage discrimination. At each level, female face some sort of gender discrimination. Deininger et al. (2013) discussed on wage discrimination in India's informal market. These are important aspects where gender bias is observed in India.

There is one more dimension where gender discrimination exists i.e. consumption and intra-household allocation of resources. It is the most challenging among all the other dimensions because of the unavailability of the data at the individual level, which makes it difficult to analyse gender difference in consumption within the households. Although individual data can be collected for other dimensions it is extremely difficult to get data for say food consumption by every member of a household. This is mainly because food items are cooked within the household and it is not possible to measure the exact quantities when it is served to the members of the household. The information on intra-household allocation of say food is required to find whether there is any gender difference in food consumption at each specific age group in the household. However, it may not be possible to know how much a female member consumes a particular item vis-á-vis a male member even if the respondent cooperates fully with the investigator. Therefore, special types of surveys may be necessary for this purpose in which the daily consumptions are recorded by actual measuring devices (Basu et al., 1986). Basu et al. (1986) collected semi-quantitative data on dietary intakes. The women running the households were asked to recall and report the amounts of various food items served to each household member the previous day in terms of eight containers of the following sizes (in ml): (1) 3,000, (2) 2,000, (3) 1,250, (4) 700, (5) 450, (6) 300, (7) 175, and (8) 100 which were supplied to them. But no such data exists at the national level.

Engel curve provides a method to gain information on individuals level using the household level data. This method is widely used and it is basically a relation between the consumption of any good with respect to income. Engel curve relation can be extended to include demographic variables and therefore serve the purpose of getting any gender difference in the consumption using the household level data. This Engel curve approach is used by many economists (for example, Deaton (1989), Fuwa et al. (2006), Gibson (1997), Gibson and Rozelle (2004), Haddad and Reardon (1993), Himaz (2010), Lee (2008), Subramaniam (1996), Subramanian and Deaton (1991), etc.) but strong evidence of the gender bias is still rare. Those who found some sort of gender difference in consumption expenditure include Fuwa et al. (2006). Fuwa et al. (2006) used the Deaton (1989)'s method of "adult good method"[1] They use data collected by IDE-MVF(Institute of Development Economies – M. Venkatarangaiya Foundation) team in 2005. They found gender bias favouring boys over girls for infants in the intrahousehold allocation of consumption goods. Himaz (2010) try to find the gender difference in educational expenditure in Sri Lanka using Deaton's approach. He found contrary results. He showed education expenditure allocation favours girls over boys in rural Sri Lanka. Similar result is also observed in Fuwa (2014). They also used Deaton's approach and detected pro-girl (5-15 age) bias in intra-household allocation of consumption budget in the rural Philippines.

There is a gap in the literature regarding the measurement procedure of detecting gender difference. In some places, it shows there exists gender bias in favour of male whereas in others the bias favours the female. Also, the Engel curve approach is not always able to detect the gender bias correctly mainly due to incorrect functional form, aggregation

[1]Adult good approach is the estimation of Engel curve for adult goods to analyze gender discrimination. An increase in children will lead to decrease in consumption of adult good given the budget constraint. So the question is how much an adult will forgo its consumption for a female child as compared to male?

problem, less number of data, large number of zero observations, etc. This thesis tries to apply Engel curve in a better functional form using large dataset. This approach is discussed in the next section.

1.3 Objective of the Present Study

The gender difference naturally arises because of the society as the major portion of the human population would rather try to live by what society desire and this directly affects the necessities they require as they try to confirm their lives according to it. The reasons for gender differences in consumption expenditure is of course not restricted to just this reason. Although there should never be gender bias towards one gender over another, the differences between genders are nonetheless undeniable biologically, mentally as well as physically. There might be a lot of other factors that might be responsible for differences in consumption expenditure over different genders. The aim of this study is to simply observe the gender differences that exist in a unique country of India which is known for its diversity in religions, food and clothing habits, society and other features that set it apart from most of the other countries. The consumption patterns of people for this country differ vastly from other country and are unique in its own ways. It is important to study gender difference in consumption expenditure to monitor the differences that are there and to check if some of these difference might prove harmful for the overall sanctity of human life in the Indian population. Of course, this study is just restricted to observe the differences.

From the previous section, it is now clear that many studies have been carried out to find out gender bias/discrimination in India. Some of these have failed while others have somehow tried to establish some kind of gender bias in favour of males mainly in the backward regions of India. The other problem faced by them is the data issue. Data on the individual level for consumption expenditure is not available for the entire country. So they have to work with the aggregated data on the household level and try to infer about the individual consumption behaviour. One method which is widely used is the "Engel curve approach". Another method which has been long-forgotten is the "Regression Decomposition Method". This technique has applications in many situations and does not need many assumptions because it is an identity itself. This method is used here to detect gender difference in food expenditure in the context of India.

Although Engel curve fails many times to detect gender bias one cannot ignore its importance. We try to rectify few of the problems like modifying the Engel function from earlier used form to a semiparametric form, and using data for entire rural India, etc. This approach tries to find gender difference in the consumption of food, milk, and intoxicants.

Lastly, we need to know that if there is gender bias in the consumption expenditure and different people consume differently then how are we going to get the exact size of household in terms of adult equivalent ratios. Adult equivalent is necessary if one is concerned about the individual needs and not the family as a whole. In that sense per capita level does not give the true picture of the resource allocation among the households. We will try to find the "consumer unit scale" for different commodity groups using a modified version of Singh (1972) approach.

As noted earlier, India is a vast country with diverse culture and consumption habits, the data collected at such a large scale over such a diverse population takes into consideration the variety of the effects based on social norms, religious influence, consumption practices and a variety of other influences that govern the food pattern for different age groups. While working with such a diverse data it is only natural to extract an equivalence scale out of it as it is going to be free of such influences. This becomes especially useful when one is to revisit the analysis done in Chapter 3 and 4 or to conduct a similar analysis in any other country with any religion and their food patterns.

In summary, it can be said that chapter 3 looks at the gender bias within households using regression decomposition method for food items as a whole. Chapter 4 uses Deaton's approach and a semiparametric model to examine item-wise gender bias. It may be noted from these two chapters that the item 'milk' seems to stand out in both the chapters in some way or the other. Lastly, Chapter 5 is extracting an adult equivalence scale to exploit the availability of such a vast and diverse data.

1.4 Format of the Thesis

The thesis has altogether five chapters excluding the introduction. The analysis is based on the consumer expenditure data of National Sample Survey Organisation. Mostly the tables are presented at the end of each chapter i.e. within the appendix. The brief reviews and organisation of the chapters are as follows:

Chapter 2: Data: A Discussion in the context of the present study

In this chapter, we primarily discuss the NSS (National Sample Survey) data sets. Mainly the last three quinquennial series for the consumer expenditure data. These data are basically 61^{st}, 66^{th} & 68^{th} rounds which are the 7^{th}, 8^{th} and 9^{th} quinquennial survey of its series. In the next few chapters, these data will be used to analyse the gender bias in the consumption expenditure for India as a whole. The format of this chapter is as follows: Section 2.1 gives the introduction of this chapter. In section 2.2 the survey design and few concepts relating to the data has been explained. Data characteristics are explained in section 2.3. Finally the chapter ends in section 2.4 with some concluding remarks.

Chapter 3: Gender Discrimination in Intra-household Allocation of Food Expenditures: A comparison between 61^{st} and 66^{th} rounds of NSS, India

In this chapter, 'Regression Decomposition Technique' is used to investigate the gender difference in individual-level food consumption of household members given the total food consumption. The idea is based on the identity that the total food expenditure of a household is the sum of food expenditures of individual members in the household. The regressors are the number of members in each age-sex group within the household. The total food expenditure is regressed on these regressors and the regression coefficients produce the average food consumption in the respective group. This procedure has been applied to the Indian household level data of the 61^{st} and 66^{th} round Consumer Expenditure Survey of the National Sample Survey Organization, Govt. of India. Our objective is to see whether there is any gender difference in the food expenditure of different age groups, particularly for the children. And if there is any gender difference in consumption then whether it is in favour of boys or girls. It is well known that in developing countries like India, there is gender bias in favour of boys. We also want to see this gender difference for both the sectors i.e., rural and urban in each round. And lastly, we would like to see if there is any gender difference pattern for different income/total expenditure groups for each round and sector. The format of this chapter is as follows: Section 3.1 gives the introduction of this chapter. In section 3.2, the proposed model is introduced. The data which have been used are introduced in section 3.3. The estimation results are discussed in section 3.4. The chapter ends with some concluding remarks in section 3.5.

Chapter 4: Within Household Gender Bias in Consumption: A Semiparametric Analysis of Engel Curves in Rural India

This chapter uses the long-debated Engel curve approach. However, as discussed earlier, to detect the gender difference in the expenditure pattern in rural India semiparametric forms of Engel curve is used. We try to estimate Engel curves for food, milk, and intoxicants for four different household sizes. In the Engel curve analysis, endogeneity of total expenditure is also considered. We try to solve this endogeneity problem using two different approaches - 2SLS and Control function Approach and the results have been compared.

The chapter is organized as follows. Introduction is given in Section 4.1. Deaton's approach is introduced in section 4.2. The Engel curve approach for detecting gender discrimination are presented in the next section 4.3. Section 4.4 outlines the details of the data. In Section 4.5, the empirical results on estimation of the models are discussed. This chapter ends with some concluding remarks in Section 4.6.

Chapter 5: Estimation of Consumer Unit scale and Economies of Scale

In this chapter, consumer unit scales for different commodity groups have been calculated. Both the 'specific' and 'income' consumer unit scales are estimated. From these scales, adult equivalent scales are calculated for different age-sex groups as well as for different commodity groups. The modified version of Singh (1972) and Singh and Nagar (1973) approach has been used to find these scales. Also, the economies of scales are estimated by an iterative procedure, using a non-parametric form of the Engel relation.

The organization of this chapter is as follows. The introduction to this chapter is presented in section 5.1. Literature review is discussed in section 5.2, where Houthakker and Prais (1971) method and Singh (1972), Singh and Nagar (1973) iterative approaches are discussed in details. In section 5.3, our proposed model has been discussed, which is based on the Singh and Nagar (1973) iterative approach. Section 5.4 discusses the data and its characteristics. Results which have been obtained are discussed in section 5.5. Finally the chapter is concluded in the last section 5.6.

Chapter 6: Summary and Future Ideas

The last chapter of this thesis begins with a brief introduction to the problem studied in this thesis. Major findings of the entire work have been presented in section 6.2. The concluding section contains a few ideas for further work in this area.

2 Data: A Discussion in the context of the present study

2.1 Introduction

Our study of gender difference in consumption expenditure is based on survey data collected in India. The large-scale sample survey for all of India conducted at the country level is collected by the National Sample Survey Office(NSSO), Ministry of Statistics and Programme Implementation, Government of India[1]. NSSO collects primary data on various socioeconomic subjects like- Household Consumer Expenditure, Employment and Unemployment, Annual Survey of Industries, Economic Census, etc. In this study, we have used three rounds of National Sample Survey (NSS) consumer expenditure data: 61^{st}, 66^{th}, and 68^{th} rounds, the corresponding years during which the survey has been conducted are respectively July 2004 to June 2005, July 2009 to June 2010 and July 2011 to June 2012. These surveys are the 7^{th}, 8^{th} and 9^{th} quinquennial surveys of its series. Apart from the quinquennial series, there also exists an "annual series" comprising consumer expenditure surveys conducted in the intervening periods between quinquennial series rounds – starting from the 42nd round (July 1986 - June 1987) and using a

[1]NSSO was set up in 1950 as part of the Indian Statistical Institute for the collection of various kinds of socioeconomic data for the country as a whole through field surveys for national accounting, planning, and other policy purposes. In 1972 it was taken over by the Government of India. It was placed under the Department of Statistics of the Ministry of Planning.

smaller sample. We are considering only the above three quinquennial surveys which are large-scale surveys. These surveys cover the whole of Indian territory except (i) interior villages of Nagaland situated beyond five kilometers of the bus route and (ii) villages in Andaman and Nicobar Islands which remain inaccessible throughout the year.

The household consumer expenditure surveys of NSS collect information from the households on various demographic characteristics as well as consumption of various items in both quantitative and value terms. The unit of analysis is the household. The survey uses the interview method for gathering information from the randomly selected household. In the next three chapters i.e. chapters 3, 4 and 5, these consumption data at the household level are used. In chapter 3 both 61^{st} and 66^{th} rounds data are used. But for the rest of the chapters, we have used the latest available data i.e. 68^{th} round data.

The chapter is organised into the following sections: Section 2.2 provides a brief discussion on the sampling design and a few concepts. Section 2.3 describes some important data characteristics for each round and also the summary statistics of the data. The final section concludes it with some remarks in Section 2.4.

2.2 Sample Design

The NSSO survey design has undergone many changes over the years but the sampling techniques followed in the 61^{st}, 66^{th}, and 68^{th} rounds are the same. The sampling design is a stratified multistage design where the first stage unit (FSU) for the rural sector is the 2001 census villages and for the urban sector, the latest Urban Frame Survey (UFS) serve the purpose. Households is the ultimate stage units (USU) in both the sectors. Every district of a state or a Union Territory (UT), has two basic strata: (i) the rural areas of the district are combined into one stratum called the rural stratum and (ii) all the urban areas are combined to form the urban stratum. Each sample household represents a different number of households in the population. The probability weights are the multipliers. Multipliers are different for different sample households. The weighted sum of the sample observation gives the estimate of the population aggregates.

In NSS 66^{th} and 68^{th} round there are two types of Schedules: Type 1 and Type 2. These two schedules have same item break-ups but they are different in terms of reference periods[2] used for collection of consumption data. Type I schedule collects information on consumption during the 'last 30 days' for frequently purchased items and for infrequently purchased items 'last 365 days' serve as the reference periods. On the other hand, Schedule Type 2 uses three types of reference periods i.e. '7 days', '30 days', and '365 days' for different set of consumption items based on the frequency of their purchase[3]. For the 61^{st} round there is only one schedule i.e. Type – 1, and this data is used along with the 66^{th} round Type – 1 data, for the analysis of gender difference in food consumption in Chapter-3. Table 2.1 provides the item-wise reference period for both the schedules, while dealing with these commodity items we need to convert them into a single reference period, i.e. '30 days' to make the data compatible[4].

Two important concepts i.e. Per Capita Food Expenditure (pcfe) and Monthly Per Capita Consumer Expenditure (MPCE) can be computed from the value of the items given by

[2]The usage of short reference period produces data which reflects fluctuation in consumption at the individual level. For example, the consumption of salt by the household may be more or less the same every week, but for fish or fruits it may not be true and for consumption of clothing and durables it is certainly not. Therefore, a longer reference period such as a year is used for less frequently purchased commodities.

[3]Type – 2 schedules is formed on the recommendation of an Expert Group that had been formed for the purpose of suggesting the most suitable reference period for each item of consumption.

[4]Formula for conversion is =Item Value $\times$ 30/Reference Periods. It gives the consumption value converted in terms of 30 days period.

the household. Income data is not provided by NSS, therefore, one can take MPCE as a proxy for income. The monthly per capita consumer expenditure (MPCE) is defined as the average expenditure per person per month (Household monthly consumer expenditure ÷ household size). It serves as an indicator of the household's level of living. Thereby distributing the population with respect to their MPCE classified economic level.

All the rounds employed two different methods of measurement of Monthly Per Capita Expenditure (MPCE) at the household level – the URP (Uniform Reference Period), which considers all items consumed during 'last 30 days' and MRP (Mixed Reference Period), consists of a mixture of 'last 365 days' for less frequent category and 'last 30 days' for the rest as the reference. There is one another concept i.e. MMRP (Modified Mixed Reference Period), which is contained in the Type 2 Schedule of the NSS dataset and consists of all 3 types of reference period i.e. 'last 7 days','last 30 days' and 'last 365 days' converted in terms of 30 days reference.

The per capita food expenditure is the average monthly food expenditure per person divided by the household size (total food expenditure ÷ household size). For schedule Type 1 the reference period on the food item is '30 days' whereas in Schedule Type 2 there are both '7 days' and '30 days' reference periods for different commodities.

The NSS data provides household consumer information on individuals such as age, gender, marital status, education level etc. In our analysis, we combine information about age and gender to produce 10 age-gender cohorts: (i) males with 0-3 years of age, (ii) females with 0 – 3 years of age, (iii) males with 4-6 years of age, (iv) females with 4-6 years of age, (v) males with 7-12 years of age, (vi) females with 7-12 years of age, (vii) males with 13-18 years of age, (viii) females with 13-18 years of age, (ix) males with age 19 years and above, and finally (x) females with age 19 years and above. The next few chapters of this thesis are devoted to the analysis of the difference in consumption expenditure for these different age-sex cohorts and try to see if there exists any expenditure difference between different gender over same age groups.

In chapter-5, a concept of consumer unit scale is defined which is nothing but weights given to each of these age-sex groups according to the consumption of the different items. These consumer unit scales can later be converted into adult equivalents. The commodity groups considered are (i) cereals and its substitutes, (ii) Pulses and its products, (iii) Milk and its products, (iv) Salt, Sugar, edible oil, and spices, (v) vegetables, (vi) Egg, fish, meat, fruits and dry fruits, (vii) other misc. food expenditure, (viii) clothing, bedding, and footwear, (ix) education and medical expenditure, (x) other misc. non-food expenditure. In the following section, these variables are discussed in more details.

2.3 Data Characteristics and Summary Statistics

For the current discussion, the last three quinquennial rounds of NSS, i.e. 61^{st}, 66^{th}, and 68^{th} rounds have been used. For these rounds, rural and urban sectors are treated differently. In Table 2.2, the summary statistics of the variables used for the estimation in Chapter-3 are presented. It can be seen that $Type-1$ data is used for 66^{th} round to make the result compatible with the 61^{st} round. The average household size decreased for both the rural and urban sector over the years. In both the sectors, the average number of members approximately over all age groups has declined over the years. The average MPCE for the rural area is Rs.517.77 in 2004-05 which increased to Rs.860.31 in 2009-10 similarly it was Rs.911.11 and then it increases to Rs.1480.24 for the urban sector. The per capita food expenditure also increases over the years. In Table 2.3, the summary statistics of variables used in Chapter-4 are shown. Note that only the data from the rural sector has been used for the estimation. Comparing Tables 2.2 and 2.3, one can see the average household size decreases even further to 4.60 in the rural sector. In the total budget, the average share of food is 54%, the share of milk is 15%, and share of intoxicant is only 3% in the rural sector (see Table 2.3). The average share of the number of females in total household size increases with the age but it is lower than the corresponding male counterparts, except for age group 4-6 and above 19 years[5].

In Table 2.4, the variables used for the estimation in Chapter–5 are presented. The average household size is 4.61 and 4.11 for respectively rural and urban sector. From the table, one can see that the average number of males is always higher than female for approximately all the age-groups in all the sectors. The average per capita food expenditure in the rural sector is Rs.785.53 and in the urban sector, it is Rs.1117.08 in 2011-12. The average monthly per capita expenditure (MMRP) is Rs.1379.35 for the rural sector and Rs.2415.35 for the urban sector.

The MPCE not only helps us to find the level of living of the households but it also helps us to divide the households in terms of the level of living. It has already been discussed in Section 2.2 that there can be three different types of MPCE. Now in the following table i.e. Table 2.5, the average MPCE for both rural and urban and average per capita food expenditure (PCFE) for each round is shown. From Table 2.5, it can be seen that the MPCE(URP), MPCE(MRP) and MPCE(MMRP) are different for each round. PCFE as a percentage of MPCE for all the reference period is also shown in the same table i.e. Table 2.5. It is observed that the percentage of PCFE is higher for MPCE (URP) as compared to MPCE(MRP). This is because the value of MPCE(MRP) is higher than the

[5]The share of the no. of male in different age groups is the difference between the share of the total no. of members and the share of the total no. of females members in different age groups.

MPCE(URP). MPCE(MRP) is higher because there are few commodities, the purchase of which cannot be captured with '30 days' reference period, for example, jewellery, clothing, footwear are not purchased every month so MPCE(URP) cannot capture these items correctly. Table 2.5 also gives us the number of sampled households (or observations) for each round and sector. Percentage of the rural sample was 63.62 % in the 61^{st} round which decreased to 58.62 % in the 66^{th} round. In 2004-05, the total population, according to the survey, was 981 million and the rural sector was 74.68% (as a percentage of the total population) whereas, in 2009-10, the population was 1.04 billion and the percentage of the rural sector was 72.95%, and finally in 2011-12, the total population increased to 1.10 billion where 792 million people belonged to the rural sector i.e., 71.43%. From these statistics, we can say that the population in the rural sector is falling. According to the census, the level of urbanisation has increased from 27.81 % in the 2001 Census to 31.16 % in the 2011 Census, while the proportion of rural population declined from 72.19 % to 68.84 %, which is a fall of 3.35 % in the rural sector. So, our result is in line with the census.

In the following chapters, we are going to analyse the expenditure pattern of the household and also the difference between male and female expenditure pattern by considering male and female in some groups according to age. Overall five age groups are considered they are: 0-3, 4-6, 7-12, 13-18 and 19 and above. The sex ratio for these age groups is presented in Table 2.6. Sex ratio is the number of females for every 1000 males. This helps us to understand the current situation of the women population in India. According to the Indian Population Census of 2011, the sex ratio was 940 females per 1000 males. In 2001 census, this sex ratio was 933. Therefore, it can be said that the last decade show an increase in sex ratio. However, the children sex ratio does not shows a rising trend. The children sex ratio falls over the years which is an area of concern for many authors like Jha et al. (2006),Arnold et al. (2002), Miller (1997), Das Gupta and Mari Bhat (1997), Griffiths et al. (2000) etc. Reasons for this decline could be excess female mortality throughout the life course, sex-selective abortion manifested through son-preference and daughter neglect etc.

In Table 2.6, age-wise sex ratio can be seen for each of the three rounds of the NSS. Census sex ratio for both 2001 and 2011 are also presented in the lower section of Table 2.6. Census 2011 shows an improvement in the sex ratio compared to census 2001 for both the sectors as well as for the overall country. But for children between 0-6 years of age, the sex ratio decline according to the census for both the sectors as well as in the overall country. For population aged 7 and above the situation is reversed, i.e. it improves. For this group, the rural and urban data is not available in the report so it is kept blank for both the censuses. Comparing the NSS data, both the 61^{st} round and the 66^{th} round shows a decline of the overall sex ratio i.e. sex ratio falls from 952 to 946. However, for 68^{th} round and 66^{th} round the overall sex ratio remains same i.e. at 946. In case of 0-6 age groups, 66^{th} round shows an improvement compared to the 61^{st} round whereas 68^{th} round again deteriorates for this category. For population aged 7 and above, 66^{th} round shows a deteriorated picture for both the rounds and the country. For the 68^{th} round, rural sector shows an increase in the sex ratio, but in the urban sector, a fall in the sex ratio is observed. For other age groups, 66^{th} round shows a mixed picture, but overall the total sex ratio falls for each age-group (except for 4-6 age group). In the 68^{th} round, the sex ratio falls for few age groups but it also increases for other age groups making the overall effect unchanged compared to the 66^{th} round.

Last but not the least the break-up of the MPCE by 20 broad item groups for the 61^{st}, 66^{th} & 68^{th} rounds are given in Table 2.7, 2.8 & 2.9 respectively. In all these three tables, the first column shows the broad item groups, the second, third and fourth column shows the average value of the different items for rural, urban and the entire population as a whole. Column fifth and sixth shows the percentage of the items in terms of MPCE (URP) for both rural and urban sectors. Lastly, column seventh and eighth shows the same for the MPCE (MRP). From these tables it can be seen that the share of food expenditure decreases over the rounds from 57.74 % in 61^{st} round to 55.80 % in 66^{th} round and eventually to 50.98 % in the 68^{th} round for the rural sector and from 44.13 % in 61^{st} round to 41.96 % in the 66^{th} round and then to 39.87 % in the 68^{th} round for the urban sector when considered for MPCE (URP). On the contrary, the share of non-food expenditure increases over these rounds for both the sectors.

The share of food expenditure as a percentage of the MPCE (MRP) also shows a similar pattern over the rounds. In the rural sector, it was 55.70 % in the 61^{st} round then it comes down to 54.32 % for the 66^{th} round and eventually to 50.66 % in the 68^{th} round. In the urban sector the share of food is 42.05 % in the 61^{st} round, 40.38 % in the 66^{th} round and then it decreases to 38.61 % in the 68^{th} round. This shows the transition of expenditure pattern from food to non-food items which is in line with the development process. It is expected that the share of food expenditure will decline with development and economic prosperity. Also due to the opening of the economy to the global market, demonstration effect may take place [6].

The major decline in the consumption pattern is for the cereals which were 18.08 % in the rural sector and 10.10 % in the urban sector for the 61^{st} round. It falls to 15.71 % and 9.11 % respectively for the rural and urban sector in the 66^{th} round and then to 12.04 % and 7.37 % respectively for the rural and urban sector in the 68^{th} round which is also seen by Mittal (2007).

[6] For more detail discussion refer Mujumdar and Kapila (2006).

2.4 Conclusions

In this chapter, a brief discussion of the data used has been given starting from the survey design to the expenditure pattern for different items of the commodities. These items will be used later to find the difference in the consumption patterns for the genders. This chapter shows that our data (or the sample) is a good representation of the population and it can be used for the analysis as well. These are the only available data set at the national level which gives details on the consumption of the items of commodities by the household.

Also, it was seen that due to urbanisation the rural sector has contracted and the urban sector expanded in terms of the number of households. According to the sex-ratio pattern followed during the last decade, our data shows that it has actually declined over the years. In the rural sector, this ratio has declined but in the urban sector, it has increased in the year 2009-10. For 0-3 years this ratio has declined over the years. Thus, our data somehow resembles the census data in terms of the sex ratio as it is more or less close to the overall figures. However, our objective is to see the overall consumption expenditure within the household. So, this sex-ratio will not affect much of our analysis.

2.5 APPENDIX

Table 2.1: Item-wise reference period for different schedule

Category	Item groups	Reference period for	
		Schedule Type 1	Schedule Type 2
I	Clothing, bedding, footwear, education, medical (institutional), durable goods	'Last 30 days' and 'Last 365 days'	Last 365 days
II	Edible oil; egg, fish & meat; vegetables, fruits, spices, beverages and processed foods; pan, tobacco & intoxicants	Last 30 days	Last 7 days
III	All other food items, fuel, and light, miscellaneous goods and services including non-institutional medical; rents and taxes	Last 30 days	Last 30 days

Source: NSS Supporting Documents

Table 2.2: Summary Statistics of the variables used for different rounds of NSS

		61^{st} round						66^{th} round					
			Rural			Urban			Rural			Urban	
Variable	Description of the Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
fe	Food expenditure (in INR)	76898	1532.00	849.24	43208	1939.88	1010.83	56919	2368.23	1236.86	39375	2954.80	1508.62
m0003	No. of male with 0-3 years age	76898	0.21	0.48	43208	0.15	0.41	56919	0.17	0.43	39375	0.13	0.37
f0003	No. of female with 0-3 years age	76898	0.20	0.47	43208	0.14	0.39	56919	0.16	0.42	39375	0.12	0.37
m0406	No. of male with 4-6 years age	76898	0.20	0.45	43208	0.14	0.39	56919	0.17	0.42	39375	0.13	0.36
f0406	No. of female with 4-6 years age	76898	0.18	0.43	43208	0.12	0.35	56919	0.16	0.40	39375	0.11	0.34
m0712	No. of male with 7-12 years age	76898	0.39	0.66	43208	0.29	0.57	56919	0.36	0.64	39375	0.26	0.53
f0712	No. of female with 7-12 years age	76898	0.34	0.63	43208	0.25	0.54	56919	0.31	0.60	39375	0.23	0.52
m1318	No. of male with 13-18 years age	76898	0.33	0.60	43208	0.31	0.58	56919	0.35	0.63	39375	0.29	0.57
f1318	No. of female with 13-18 years age	76898	0.29	0.57	43208	0.27	0.56	56919	0.29	0.57	39375	0.24	0.53
m19+	No. of male with age 19 years & above	76898	1.36	0.87	43208	1.42	0.89	56919	1.37	0.84	39375	1.40	0.88
f19+	No. of female with age 19 years & above	76898	1.38	0.74	43208	1.33	0.84	56919	1.38	0.72	39375	1.32	0.83
hhsize	household size	76898	4.89	2.40	43208	4.43	2.29	56919	4.71	2.23	39375	4.24	2.19
pcfe	Per capita food expenditure (in INR)	76898	313.44	128.88	43208	438.09	192.81	56919	502.87	201.46	39375	697.55	309.57
MPCE(URP)	Monthly per capita expenditure (in INR) (Uniform reference period)	76898	517.77	264.48	43208	911.11	530.81	56919	860.31	421.23	39375	1480.24	842.10

[a] Type 1 data of 66^{th} round is used to make the comparison with the 61^{st} round.

[b] Outliers are deleted from the sample and also an upper bound for the MPCE(URP) is set for both the rural and urban sector.

Source: Author's calculation based on 2004-05, 2009-10 (Type 1) Consumer Expenditure Household Survey, collected by NSS.

Table 2.3: Summary Statistics of the variables considered for rural 68^{th} round of NSS

Variable	Description	Obs	Mean	Std. Dev.
hhsize	household size	59695	4.60	2.20
lnpcte	log of per capita total expenditure (MRP)	59695	7.02	0.49
shf	share of food expenditure	59695	0.54	0.11
intoxsh	share of intoxicant	59695	0.03	0.04
shb3	share of milk	59695	0.15	0.13
shf0003	share of the no. of female with 0-3 years of age	59695	0.03	0.08
shf0406	share of the no. of female with 4-6 years of age	59695	0.03	0.07
shf0712	share of the no. of female with 7-12 years of age	59695	0.05	0.11
shf1318	share of the no. of female with 13-18 years of age	59695	0.05	0.11
shf19+	share of the no. of female with age 19 years and above	59695	0.33	0.18
sht0003	share of the total no. of members with 0-3 years of age	59695	0.06	0.11
sht0406	share of the total no. of members with 4-6 years of age	59695	0.05	0.10
sht0712	share of the total no. of members with 7-12 years of age	59695	0.12	0.17
sht1318	share of the total no. of members with 13-18 years of age	59695	0.12	0.18
sht19+	share of the total no. of members with age 19 years and above	59695	0.65	0.24

[a] Type 1 data of 68^{th} round is used.

Source: Author's calculation based on 2011-12 (Type 1) Consumer Expenditure Household Survey, collected by NSS.

Table 2.4: Summary Statistics of the variables used from the 68^{th} round of NSS

		Rural			Urban		
Variables	Description	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	58775	0.16	0.42	41124	0.12	0.36
f0003	No. of female with 0-3 years age	58775	0.14	0.40	41124	0.11	0.35
m0406	No. of male with 4-6 years age	58775	0.16	0.40	41124	0.11	0.34
f0406	No. of female with 4-6 years age	58775	0.14	0.38	41124	0.10	0.33
m0712	No. of male with 7-12 years age	58775	0.33	0.61	41124	0.25	0.53
f0712	No. of female with 7-12 years age	58775	0.30	0.59	41124	0.22	0.49
m1318	No. of male with 13-18 years age	58775	0.34	0.62	41124	0.27	0.56
f1318	No. of female with 13-18 years age	58775	0.27	0.56	41124	0.22	0.51
m19+	No. of male with age 19 years and above	58775	1.38	0.86	41124	1.39	0.88
f19+	No. of female with age 19 years and above	58775	1.39	0.74	41124	1.31	0.84
hhsize	household size	58775	4.61	2.18	41124	4.11	2.15
pcfe	percapita food expenditure	58775	785.53	364.67	41124	1117.08	564.62
pc_1	percapita expenditure on c1 (in INR)	58775	153.07	67.04	41124	174.38	73.40
pc_2	percapita expenditure on c2(in INR)	58775	41.12	24.94	41124	53.31	29.27
pc_3	percapita expenditure on c3 (in INR)	58775	113.17	140.04	41124	180.09	157.23
pc_4	percapita expenditure on c4 (in INR)	58775	127.62	61.07	41124	162.76	73.80
pc_5	percapita expenditure on c5 (in INR)	58775	93.71	50.52	41124	120.49	70.70
pc_6	percapita expenditure on c6 (in INR)	58775	105.08	125.33	41124	178.65	186.21
pc_7	percapita expenditure on c7 (in INR)	58775	151.75	173.45	41124	247.41	340.65
pc_8	percapita expenditure on c8 (in INR)	58775	98.85	60.50	41124	159.28	117.57
pc_9	percapita expenditure on c9 (in INR)	58775	135.59	233.41	41124	295.46	441.81
pc_10	percapita expenditure on c10 (in INR)	58775	359.39	294.99	41124	843.52	804.15
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	58775	1379.35	730.28	41124	2415.35	1543.31

[a] Type 2 data of 68^{th} round is used .

[b] c1 - cereals, c2 - pulses, c3 -milk, c4-salt,sugar,oil and spices, c5-vegetables, c6- egg,fish,meat,fruits and dry-fruits, c7-other misc. food exp., c8-clothing,bedding and footwear, c9-education and medical exp.,c10-other misc. nonfood exp.

Source: Author's calculation based on 2011-12 (Type 2) Consumer Expenditure Household Survey, collected by NSS.

Table 2.5: Average of MPCE & PCFE for 61^{st}, 66^{th} & 68^{th} for both Rural and Urban sectors

	61^{st} Round			66^{th} Round			68^{th} Round			
Variables	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total	Description of the variable
MPCE(URP)	558.78	1052.36		927.7	1785.81		1278.94	2399.23		Monthly per capita expenditure (Uniform Reference Period)
MPCE(MRP)	579.17	1104.6		953.05	1856.01		1287.17	2477.02		Monthly per capita expenditure (Mixed Reference Period)
MPCE(MMRP)				1053.64	1984.46		1429.96	2629.65		Monthly per capita expenditure (Modified Mixed Reference Period)
PCFE (URP/MRP)	322.63	464.44		517.68	749.4		652.06	956.47		Per capita Food expenditure (under URP/MRP)
PCFE (MMRP)				631.47	911.28		802.42	1163.18		Per capita Food expenditure (under MMRP)
% of PCFE (URP/MRP) (in MPCE(URP))	57.74	44.13		55.8	41.96		50.98	39.87		% of Per capita Food expenditure (under URP)
% of PCFE (URP/MRP) (in MPCE(MRP))	55.7	42.05		54.32	40.38		50.66	38.61		% of Per capita Food expenditure (under MRP))
% of PCFE (MMRP) (in MPCE(MMRP))				59.93	45.93		56.1	44.2		% of Per capita Food expenditure (under MMRP)
No.of household (URP/MRP)	79298 (63.62%)	45346 (36.38%)	124644	59119 (58.62%)	41736 (41.38%)	100855	59695 (58.72%)	41967 (41.28%)	101662	Number of household size in Type 1 dataset
No.of household(MMRP)				59097 58.63%	41697 41.37%	100794	58775 58.83%	41124 41.17%	99899	Number of household size in Type 2 dataset

[a] In the 61^{st} round Type 2 Schedule was not collected, so MMRP is missing for this round.

[b] PCFE also includes pan, tobacco and intoxicants other than what NSS includes.

Source: NSS Report (crossed checked by the author).

Table 2.6: Sex Ratio for different age groups and different rounds for both Rural and Urban sectors

	61^{st} Round			66^{th} Round			68^{th} Round		
Age groups	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total
0-3	954	910	945	940	942	940	940	930	937
4-6	899	832	885	938	890	928	937	885	925
0-6	927	873	916	939	916	934	939	908	931
7 -12	906	897	904	895	918	900	867	871	868
13- 18	901	934	909	852	881	859	887	852	878
19+	996	961	987	992	962	984	990	972	985
7-19+	963	947	959	950	943	948	951	940	948
Total	956	936	952	948	940	946	949	936	946
	All India			Children with 0-6 years			Population aged 7 and above		
	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total
Census 2001	946	900	933	934	906	927	-	-	934
Census 2011	947	926	940	919	902	914	-	-	944

[a] Lower panel shows the Census Sex Ratio from Census 2001 & 2011 respectively. For more details visit http://www.censusindia.gov.in/.
[b] For the calculation of the above sex ratio individual multiplier has been used provided by the NSS.
[c] Type 1 data is used for each round to make the comparison compatible.
Source: Author's calculation based on 2004-05, 2009-10 & 2011-12 (Type 1) Consumer Expenditure Household Survey, collected by NSS.

Table 2.7: Break-up of MPCE by 20 broad item groups: all-India, for the 61^{st} round

	URP values (MRP)			Percentage URP to total		Percentage MRP to total	
Broad Item groups	Rural	Urban	Total	Rural	Urban	Rural	Urban
Cereals and Cereal substitute	101.04	106.33	102.38	18.08	10.10	17.44	9.63
Pulses	17.91	23.62	19.35	3.20	2.24	3.09	2.14
Milk and milk products	47.31	83.30	56.42	8.47	7.92	8.17	7.54
edible oil	25.72	36.37	28.42	4.60	3.46	4.44	3.29
Meat, fish, egg	18.60	28.47	21.10	3.33	2.71	3.21	2.58
Vegetables	36.23	49.91	39.69	6.48	4.74	6.26	4.52
fruits and dry fruits	10.42	23.64	13.77	1.86	2.25	1.80	2.14
sugar, salt and spices	24.99	30.45	26.37	4.47	2.89	4.32	2.76
beverages, refreshments & proc. food	25.38	65.31	35.49	4.54	6.21	4.38	5.91
pan, tobacco & intoxicants	15.03	17.04	15.54	2.69	1.62	2.60	1.54
food total	322.63	464.45	358.53	57.74	44.13	55.70	42.05
fuel and light	56.84	104.62	68.93	10.17	9.94	9.81	9.47
conveyance	21.03	68.59	33.07	3.76	6.52	3.63	6.21
entertainment	3.46	19.81	7.60	0.62	1.88	0.60	1.79
consumer serv	21.18	74.15	34.59	3.79	7.05	3.66	6.71
misc. goods & other consumer ser	29.11	53.44	35.27	5.21	5.08	5.03	4.84
rent & taxes	3.89	67.95	20.10	0.70	6.46	0.67	6.15
clothing, bedding & footwear	29.56 (44.92)	49.26 (73.22)	34.55 (52.08)	5.29	4.68	7.76	6.63
education	14.90 (18.06)	52.69 (73.70)	24.46 (32.15)	2.67	5.01	3.12	6.67
medical	36.97 (36.34)	54.59 (57.42)	41.43 (41.67)	6.62	5.19	6.27	5.20
durable goods	19.23 (21.74)	42.81 (47.17)	25.20 (28.17)	3.44	4.07	3.75	4.27
non-food total	236.16 (256.56)	587.91 (640.06)	325.21 (353.65)	42.26	55.87	44.30	57.95
All item	558.79 (579.19)	1052.35 (1104.51)	683.74 (712.18)	100.00	100.00	100.00	100.00

[a] Lower panel shows the non-food items whereas the upper panel is for the different food items.
[b] MRP value of the corresponding item are shown in parenthesis.
Source: NSS Report 2004-05.

Table 2.8: Break-up of MPCE by 20 broad item groups: all-India, for the 66th round

	URP values (MRP)			Percentage URP to total		Percentage MRP to total	
Broad Item groups	Rural	Urban	Total	Rural	Urban	Rural	Urban
Cereals and Cereal substitute	145.76	162.69	150.34	15.71	9.11	15.29	8.77
Pulses	35.71	49.77	39.51	3.85	2.79	3.75	2.68
Milk and milk products	80.16	138.71	95.99	8.64	7.77	8.41	7.47
edible oil	34.15	46.10	37.38	3.68	2.58	3.58	2.48
Meat, fish, egg	32.26	48.03	36.53	3.48	2.69	3.38	2.59
Vegetables	57.20	76.66	62.46	6.17	4.29	6.00	4.13
fruits and dry fruits	14.87	37.37	20.96	1.60	2.09	1.56	2.01
sugar, salt and spices	44.96	55.19	47.73	4.85	3.09	4.72	2.97
beverages, refreshments & proc. food	52.03	112.97	68.51	5.61	6.33	5.46	6.09
pan, tobacco & intoxicants	20.60	21.91	20.95	2.22	1.23	2.16	1.18
food total	517.68	749.40	580.36	55.80	41.96	54.32	40.38
fuel and light	87.79	142.76	102.66	9.46	7.99	9.21	7.69
conveyance	37.56	115.21	58.56	4.05	6.45	3.94	6.21
entertainment	8.20	31.50	14.50	0.88	1.76	0.86	1.70
consumer serv	44.47	127.33	66.88	4.79	7.13	4.67	6.86
misc. goods & other consumer ser	46.03	84.08	56.32	4.96	4.71	4.83	4.53
rent & taxes	7.06	133.79	41.33	0.76	7.49	0.74	7.21
clothing, bedding & footwear	54.76 (67.66)	99.95 (118.34)	66.98 (81.37)	5.90	5.60	7.10	6.38
education	26.51 (40.27)	92.91 (162.19)	44.47 (73.25)	2.86	5.20	4.23	8.74
medical	53.24 (57.02)	89.78 (98.79)	63.12 (68.32)	5.74	5.03	5.98	5.32
durable goods	44.42 (39.30)	119.09 (92.61)	64.62 (53.72)	4.79	6.67	4.12	4.99
non-food total	410.02 (435.36)	1036.41 (1106.61)	579.44 (616.92)	44.20	58.04	45.68	59.62
All item	927.70 (953.05)	1785.82 (1856.01)	1159.80 (1197.28)	100.00	100.00	100.00	100.00

[a] Lower panel shows the non-food items whereas the upper panel is for the different food items.
[b] MRP value of the corresponding item are shown in parenthesis.
Source: NSS Report 2009-10.

Table 2.9: Break-up of MPCE by 20 broad item groups: all-India, for the 68^{th} round

	URP values (MRP)			Percentage URP to total		Percentage MRP to total	
Broad Item groups	Rural	Urban	Total	Rural	Urban	Rural	Urban
Cereals and Cereal substitute	154.01	176.79	160.52	12.04	7.37	11.97	7.14
Pulses	41.88	54.48	45.48	3.27	2.27	3.25	2.20
Milk and milk products	116.13	186.47	156.57	9.08	7.77	9.02	7.53
edible oil	48.22	63.62	52.62	3.77	2.65	3.75	2.57
Meat, fish, egg	45.62	66.94	51.71	3.57	2.79	3.54	2.70
Vegetables	61.88	81.40	67.46	4.84	3.39	4.81	3.29
fruits and dry fruits	24.76	55.87	33.65	1.94	2.33	1.92	2.26
sugar, salt and spices	54.88	67.43	58.46	4.29	2.81	4.26	2.72
beverages, refreshments & proc. food	74.59	169.90	101.82	5.83	7.08	5.79	6.86
pan, tobacco & intoxicants	30.10	33.56	31.09	2.35	1.40	2.34	1.35
food total	652.07	956.47	739.04	50.98	39.87	50.66	38.61
fuel and light	118.26	182.54	136.63	9.25	7.61	9.19	7.37
conveyance	60.92	180.98	95.22	4.76	7.54	4.73	7.31
entertainment	14.24	43.77	22.67	1.11	1.82	1.11	1.77
consumer serv	57.57	155.78	85.63	4.50	6.49	4.47	6.29
misc. goods & other consumer ser	63.18	114.20	77.76	4.94	4.76	4.91	4.61
rent & taxes	9.90	189.15	61.12	0.77	7.88	0.77	7.64
clothing, bedding & footwear	97.16 (101.48)	156.53 (171.60)	114.12 (121.52)	7.60	6.52	7.88	6.93
education	39.84 (50.69)	135.73 (193.09)	67.24 (91.38)	3.11	5.66	3.94	7.80
medical	87.67 (94.83)	131.98 (150.45)	100.33 (110.72)	6.85	5.50	7.37	6.07
durable goods	78.14 (64.03)	152.09 (138.99)	99.27 (85.45)	6.11	6.34	4.97	5.61
non-food total	626.88 (635.11)	1442.77 (1520.55)	859.99 (888.09)	49.02	60.13	49.34	61.39
All item	1278.94 (1287.17)	2399.24 (2477.02)	1599.03 (1627.13)	100.00	100.00	100.00	100.00

[a] Lower panel shows the non-food items whereas the upper panel is for the different food items.
[b] MRP value of the corresponding item are shown in parenthesis.
Source: NSS Report 2011-12.

3 Gender Discrimination in Intra-household Allocation of Food Expenditures: A comparison between 61^{st} and 66^{th} rounds of NSS, India

3.1 Introduction

There is a long debate among the scholars regarding the gender discrimination. Also, there is a great deal of evidence that males and females are not treated equally. Not only in India, countries like China, Pakistan, Bangladesh and other parts of Asia has lots of evidence of gender bias against females (e.g. Gupta, 2005; Klasen and Wink, 2003). The literature on gender bias is huge and contains many dimensions like education, health, labour market and intra-household allocation of resources within the household (Chen et al., 1981; Emerson and Souza, 2007; Zimmermann, 2012). Sometimes, measurement of gender bias becomes difficult and cannot be done at the satisfactory level. This is due to a number of reasons. The most challenging among them is the unavailability of the data at the individual level which often makes it harder to analyse gender bias within households. While it may be possible to get data on expenditure on education and even on the health of each member of the household, it is extremely difficult to get data on food consumption for every member in the household. This is mainly because most of the food items are cooked within the household and not possible to measure the exact quantities when it is served to members.

The information on intra-household allocation of food is required to find whether there is any gender difference in food consumption at each specific age group in the household. However, it may not be possible to know how much a female member consumes a particular item vis-á-vis a male member even if the respondent cooperates fully with the investigator. Therefore, special types of surveys may be necessary for this purpose in which the daily consumptions are recorded by actual measuring devices (Basu et al., 1986). Basu et al. (1986) collected one-day semi-quantitative data on dietary intakes. The women running the households were asked to recall and report the amounts of various food items served to each household member the previous day in terms of eight containers of the following sizes (in ml): (1) 3,000, (2) 2,000, (3) 1,250, (4) 700, (5) 450, (6) 300, (7) 175, and (8) 100. But no such data exists at the national level.

Another very famous method, and also one that is widely used, is the indirect Engel curve approach. Engel curve is basically a relation between the consumption of any good with respect to income. Engel curve relation can be extended to include demographic variables into it and therefore serve the purpose to get any gender difference in the consumption of the item using the household level data. This Engel curve approach is used by many economists [e.g. Deaton (1989), Fuwa et al. (2006), Gibson (1997), Gibson and Rozelle (2004), Haddad and Reardon (1993), Himaz (2010), Lee (2008), Subramaniam (1996), Subramanian and Deaton (1991), etc.] but strong evidence of the gender bias is still rare. Those who found some sort of gender difference in consumption expenditure include Fuwa et al. (2006). They applied the Engel curve approach on adult goods and found gender bias favouring boys over girls for the consumption goods in Rural India. Although their findings were not always statistically significant. Himaz (2010) try to find the gender difference in education expenditure in SriLanka. He found a significant difference between girls and boys for school enrolment for the age group 17-19 category favouring girls. Lancaster et al. (2008) uses the Indian data to investigate the gender bias in the intra-household allocation of expenditure. They found significant gender bias in some items especially education expenditure and this bias becomes stronger for the economically backward regions. Although various commodities show the gender difference in consumption expenditure, a study on gender difference in food expenditure at the country level is not yet done in India. In this chapter, we treat different expenditure groups differently to see if there exist any gender difference in food consumption expenditure for all these expenditure groups separately.

There is one more method, introduced by Kemsley (1952), to get the individual level allocation of resources from the family total. The idea is arranged into a simple linear regression model without the constant term. The same approach is used by Case et al. (2002), Mason et al. (1999) and Pal and Bharati (2010). This approach assumes total expenditure or expenditure on a specific commodity as the sum of expenditure by each of the household members, where each person with a given age-sex receives the same amount. Pal and Bharati (2010), however, named it as 'Regression Decomposition Technique'[1]. They decomposed total calorie consumption into the calorie consumptions of individual

[1] Kemsley (1952)'s method is just an application of the Regression Decomposition Identity. The regression technique can be used to estimate the unknown parameters in identities of the form $Y_i = \sum b_i x_i$, where b_i's are unknown. These identities may be termed as Regression Decomposition Identities. This technique has applications in many situations and does not need many assumptions because it is an identity itself. The thesis is a rediscovery of Kemsley (1952)'s method. We do not think that this method has been applied anywhere else. If we take total expenditure and decompose it then the model can be used to find Adult Equivalent Scales. Adult Equivalent Scales tell us the proportion of expenditure that is made to a category of member compared to that of an adult male. Observe that Adult Equivalent Scale loses its meaning if we take other forms such as Engel Curves which take other variables as regressors.

member groups in the households and found that there is no difference in the calorie consumptions of female members relative to that of male members except for higher income groups for both the rural and urban sectors in India.

In this chapter, the issue of gender difference is again addressed with respect to food expenditure, an approach borrowed from Pal and Bharati (2010) and others and is applied on Indian data to get some idea on the gender difference in the food expenditure for different sectors as well as for various expenditure groups. In other words, we want to see whether there is some disparity with respect to food expenditure between genders for all age-groups and expenditure class in case of both the 61^{st} and 66^{th} round of NSS data. We want to see if there is any bias in favour of boys against girls in the households both for rural and urban as well as for each expenditure class. The result of the analysis shows that there are age groups where there is a gender difference in food expenditure. For the age group 0-3 and for the adults, we found a significant gender difference in food expenditure. For each expenditure class, we found few age groups which show the gender difference in food consumptions. For the age group 0-3 and 19 years and above there is some gender difference in food consumption expenditure for many expenditure classes for both the rounds as well as the sectors. Lancaster et al. (2008) found pro-male gender bias in food expenditure for adults in urban Maharastra whereas other similar studies like Subramanian and Deaton (1991) failed to detect gender bias in food consumption in Maharastra.

The chapter is organised as follows: Section 3.2 explains the model and the procedure for estimation of the food expenditure corresponding to the age-sex groups of members of the household. The nature of the data used is discussed in section 3.3. Section 3.4 contains the results and major findings. Finally, the chapter is concluded in Section 3.5.

3.2 The Model

The main assumption of the model is that it is based on the identity that *the total food expenditure of a household is the sum of food expenditures of all the members of the household.*

$$FE_j = f_{j1} + f_{j2} + ... + f_{jQ} \tag{3.1}$$

where, FE_j is the total food expenditure and $f_{j1}, f_{j2}, ..., f_{jQ}$ are the food expenditures made by the first member, second member and so on of the j^{th} household.

The number of members within a household i.e. Q varies from one household to the other. Therefore, we fix Q by taking age in years of all possible members say up to 100 years, assuming that nobody lives beyond the age of 100, and multiply the number of members with the average food expenditure in that age to get the identity as

$$FE_j = b_{j1}x_{j1} + b_{j2}x_{j2} + ... + b_{j100}x_{j100} \tag{3.2}$$

where b_{ji} is the average food expenditure of members with age 'i' and x_{ji} is the number of people having the i^{th} age for the j^{th} household. In most cases x_{ji} will be '0', in some cases it will be '1' and in rare cases, the value of x_{ji} may be greater than 1. Moreover, there are differences in the consumption patterns of male and female members of the households. To account for that we may, however, increase the number of regressors from 100 to 200 taking male and female members separately for all age categories. The number of member in each age-sex category i.e. x_{ji} are known and also the total food expenditure FE_j for each household. Thus, it automatically fits into a regression model with many exogenous variables. Also, because of the special nature of the regressors taking non-zero values only in few cases, we require data from a large number of households to estimate the coefficients.

As a way out we may divide the people within a household into different age-sex groups, say, male with 0-3 years of age, female with 0-3 years of age, male with 4-6 years of age, female with 4-6 years of age and so on. *The total food expenditure of a household is then the sum of food expenditures by different groups of members of the household.* In other words, the model gives us a method to decompose total food expenditure into the food expenditures by different age-sex groups. Now we take K such groups of members for each household and for the j^{th} household the number of members are $x_{j1}, x_{j2}, ..., x_{jK}$ for the age-sex groups 1,2,..., K respectively. The identity for total food expenditure i.e. equation (3.2) becomes[2]

$$FE_j = \beta_{j1}x_{j1} + \beta_{j2}x_{j2} + ... + \beta_{jK}x_{jK} \tag{3.3}$$

where, $\beta_{j1}, \beta_{j2}, ..., \beta_{jK}$ are the average food expenditures of the members for the respective age-sex groups for the j^{th} household. In general $\beta_{j1}, \beta_{j2}, ..., \beta_{jK}$ will vary from household to household and also from one member to the other member within each group[3]. For simplification, it is assumed that the average food expenditure for any group, say k i.e. β_{jk}, has a common mean β_k and a random component u_{jk}, which captures the fluctuations.

$$\beta_{jk} = \beta_k + u_{jk}, for\, all\, k = 1, 2, 3, ..., K \tag{3.4}$$

[2]In this chapter, we are using linear form, while Chapter-4 deals with the semiparametric form of the Engel curve, which is not a translog form. In Chapter-5, we will deal with the Houthakker and Prais (1971) model, which is, in fact, the log-log form. Since the goal was to find the gender difference, one can see easily why these forms were taken. By taking the demographic variables on the right-hand side appropriately, one can compare the coefficients to get the gender difference, if any.

[3]Note that, β_{jk} is the average food expenditure for the k^{th} age-sex group in the j^{th} household, so we are excluding the individual effect within each group.

where, $u_{j1}, u_{j2}, ..., u_{jK}$ are the deviations of the actual food expenditure of the respective group from its mean. Another important thing is to note that the identity (3.3) is valid only when there is no other food expenditure, positive or negative, incurred by the household other than over the household members. But in real life, total household food expenditure includes food expenditure on guests, domestic workers, cattle and other visitors of the house. Therefore, total food expenditure of the household, in this case, will be greater than the sum of food expenditures of the individual members. Also, some members of the household may avail some food from outside the house without paying, for example, domestic workers generally get some food from their employer's house. Thus, the inequality may be the other way round as well. This introduces another error into the model. This error has two components: systematic and erratic. The erratic component will have zero mean, whereas the systematic component is usually non-zero, the sign depends on the expenditures made on food by the household to entertain guests and other individuals after deducting the amount of food (converted in terms of monetary value) taken by the members of the household from outside. The systematic component of the error becomes the intercept term and the erratic component is absorbed in the equation error. In other words, the expected value of this error, if non-zero, will make the intercept term and the deviation of this error from its expected value will be absorbed in the equation error.

Equation (3.3) may then be reformulated as

$$FE_j = \beta_{j0}x_{j0} + \beta_{j1}x_{j1} + \beta_{j2}x_{j2} + ... + \beta_{jK}x_{jK} \tag{3.5}$$

where, β_{j0} is the intercept term[4] and x_{j0} is the intercept variable taking value 1 for all j.

[4] $\beta_{j0} = \beta_0 + u_{j0}$ from equation (3.4) where, β_0 and u_{j0} are respectively the systematic and erratic components of the intercept term.

We can thus rewrite identity (3.3) as

$$\begin{cases} FE_j = (\beta_0 + u_{j0})x_{j0} + (\beta_1 + u_{j1})x_{j1} + ... + (\beta_K + u_{jK})x_{jK} \\ \qquad = \beta_0 x_{j0} + \beta_1 x_{j1} + ... + \beta_K x_{jK} + \Sigma_k u_{jk} x_{jk} \\ \qquad = \beta_0 x_{j0} + \beta_1 x_{j1} + ... + \beta_K x_{jK} + \epsilon_j \\ \qquad = x_j\prime\beta + \epsilon_j \end{cases} \tag{3.6}$$

where,

$$\epsilon_j \quad = u_{j0} + \sum_{k=1}^{K} u_{jk}x_{jk} = \sum_{k=0}^{K} u_{jk}x_{jk} = x_j\prime u_j \;\; assuming \;\; x_{j0} = 1.$$
$$x_j\prime = \begin{bmatrix} x_{j0} & x_{j1} & ... & x_{jK} \end{bmatrix};$$
$$\beta\prime \;\; = \begin{bmatrix} \beta_0 & \beta_1 & ... & \beta_K \end{bmatrix};$$
$$u_j\prime = \begin{bmatrix} u_{j0} & u_{j1} & ... & u_{jK} \end{bmatrix}$$

The assumptions of u_{jk} are as follows:

$$\begin{cases} E(u_{jk}) = 0 \, for \, all \, j = 1, ..., H \, and \, k = 0, 1, ..., K. \\ E(u_{jk}^2) = \varphi_{kk} \, for \, all \, j = 1, .., H. \\ E(u_{jk}u_{j\prime k}) = 0 \, for \, all \, k = 0, 1, ..., K \, where \, j \neq j\prime \\ E(u_{jk}u_{jk\prime}) = \varphi_{kk\prime} \, for \, all \, j = 1, .., H. \\ E(u_{jk}u_{j\prime k\prime}) = 0 \; where \, j \neq j\prime \, and \, k \neq k\prime \end{cases} \tag{3.7}$$

where H is the total number of units (households) taken for regression. These assumptions state that the error terms are independent and identically distributed with an expected value to be zero and variance to be constant. The food expenditure of any household member's may be influenced by another members' food consumption within the very same household ($E(u_{jk}u_{jk\prime}) = \varphi_{kk\prime}$) but it will not be influenced by the consumption of another member of other household whether they belong to same age-group ($E(u_{jk}u_{j\prime k}) = 0$) or different ($E(u_{jk}u_{j\prime k\prime}) = 0$).

The var–cov matrix of u_{jk} for any j will be then,

$$\begin{cases} E(u_j u_j\prime) = E \begin{bmatrix} u_{j0}u_{j0} & u_{j0}u_{j1} & \cdots & u_{j0}u_{jK} \\ u_{j1}u_{j0} & u_{j1}u_{j1} & \cdots & u_{j1}u_{jK} \\ \cdots & \cdots & \cdots & \\ u_{jK}u_{j0} & u_{jK}u_{j1} & \cdots & u_{jK}u_{jK} \end{bmatrix} \\ \qquad = \begin{bmatrix} \varphi 00 & \varphi 01 & \cdots & \varphi 0K \\ \varphi 01 & \varphi 11 & \cdots & \varphi 1K \\ \cdots & \cdots & \cdots & \cdots \\ \varphi 0K & \varphi 1K & \cdots & \varphi KK \end{bmatrix} = \Phi(say) \end{cases} \tag{3.8}$$

where, $u_j\prime = \begin{bmatrix} u_{j0} & u_{j1} & \dots & u_{jK} \end{bmatrix}$.

Note that here, Φ is symmetric and independent of j i.e. it is constant over the household[5].

Before moving ahead one should discuss the interpretation of the estimates of the regression coefficients for expression (3.6). Each element in β, except β_0, is the expected amount of monthly food expenditure for a member of the respective group. This may also be interpreted as the increase in the average amount of total food expenditure due to an increase of a person in the respective category. If the intercept term is significant and positive, i.e. the sum of members' food expenditure is less than the total food expenditure on the average, then it means that there is some extra food expenditure incurred possibly

[5]This model is similar to Random Coefficient Model introduced by Hildreth and Houck (1968). The only difference is that they assumed $\varphi_{kk\prime} = 0$, which is a special case of expression (3.8). Here, it is a more generalized model and based on the assumption that food habit of one person influences that of another within the very same household.

on guests or domestic workers for many of the households which outweighed the food consumption by the members of the households outside the house and if it is negative then the interpretation is the other way round. The variable associated with the intercept term always takes value 1. Thus, it may be interpreted as a 'ghost' member of the household which may consume or produce extra food other than the permanent members.

With this formulation, one can write all the equations in a compact form as

$$FE = X\beta + \epsilon \tag{3.9}$$

where, $FE = (FE_1\ FE_2\ ...\ FE_H)'$, $\epsilon = (\epsilon_1\ \epsilon_2\ ...\ \epsilon_H)'$, and $X = \begin{bmatrix} 1 & x_{11} & ... & x_{1K} \\ 1 & x_{21} & ... & x_{2K} \\ ... & ... & ... & ... \\ 1 & x_{H1} & ... & x_{HK} \end{bmatrix}$

Observe that u_{jk} and ϵ_j are random variables. And the properties of the disturbance ϵ of equation (3.9) can be deduced from equations (3.7) and (3.8) as:

$$\begin{cases} E(\epsilon_j) = 0\, for\ all\ j = 1, ..., H \\ E(\epsilon_j^2) = \sigma_j^2 = E(x_j\prime u_j u_j\prime x_j) = x_j\prime \Phi x_j\ \ for\, all\, j = 1, ..., H \\ where,\ x_j\prime = \begin{bmatrix} 1 & x_{j1} & ... & x_{jK} \end{bmatrix} \\ E(\epsilon_j \epsilon_i) = 0\, for\ all\ i, j = 1, ..., H;\ i \neq j \end{cases} \tag{3.10}$$

The dispersion matrix of ϵ is

$$\Omega = E(\epsilon\epsilon\prime) = diag(\sigma_1^2,\, \sigma_2^2, ..., \sigma_H^2) \tag{3.11}$$

After expanding the expression σ_j^2 from equation (3.10) it is observed that σ_j^2 consists of two parts (say): A and B. The first part i.e. A is the linear function of $x_{j0}, x_{j1}, x_{j2}, ..., x_{jK}$ i.e., $A = \varphi_{00} + 2\varphi_{01}x_{j1} + ... + 2\varphi_{0K}x_{jK}$ and the second part i.e. B is a quadratic function of $x_{j1}, x_{j2}, ..., x_{jK}$. Lets assume the coefficients within A as $a_0 (= \varphi_{00}), a_1 (= 2\varphi_{01}), \ldots, a_K (= 2\varphi_{0K})$ and for B it is $a_{11}, a_{12}, ..., a_{KK}$ for different combinations of $x_{j1}, x_{j2}, ..., x_{jK}$ respectively. Thus,

$$\begin{cases} \sigma_j^2 = x_j{\prime}\Phi x_j = a_0 + a_1 x_{j1} + ... + a_K x_{jK} + a_{11} x_{j1}^2 + a_{12} x_{j1} x_{j2} + ... + a_{1K} x_{j1} x_{jK} \\ \qquad +a_{22} x_{j2}^2 + ... + a_{2K} x_{j2} x_{jK} + ... + a_{K-1,K} x_{j,K-1} x_{jK} + a_{KK} x_{jK}^2 \end{cases} \tag{3.12}$$

From the above equation, one can say that equation (3.9) constitutes a model with heteroscedastic disturbance term. The variance at each sample point is the linear combination of the explanatory variables and its quadratic combination.

In the presence of heteroscedasticity, the ordinary least squares estimator of β i.e.

$$\beta_{OLS} = (X'X)^{-1} X'FE \tag{3.13}$$

is, however, unbiased but generally inefficient. In that case, one should go for the generalised least squares (GLS) estimate of β, only if Ω is known.

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}FE) \tag{3.14}$$

But without the knowledge of Ω, it becomes necessary to somehow estimate Ω and substitute it in equation (3.14) and find out the feasible generalised least squares (FGLS) estimate of β.

Φ is an important component in Ω therefore, it should be estimated first. To make it simple, let us transform the diagonal matrix Ω into a column matrix σ^2 so that the elements in Φ can be estimated. Thus,

$$\sigma^2 = \begin{bmatrix} \sigma_1^2 \\ . \\ . \\ . \\ \sigma_H^2 \end{bmatrix} = Z^* \alpha \tag{3.15}$$

where, Z^* is a (H, (K+1)(K+2)/2) matrix[6] with column vectors as $1, x_1, x_2, ..., x_K$, $x_1^2, x_1x_2, ..., x_1x_K, x_2^2, ..., x_2x_K, ..., x_K^2$ and $\alpha = (a_0\ a_1....a_K\ a_{11}\ a_{12}......a_{1K}\ a_{22}\ a_{23}...$ $a_{2K}...a_{KK})'$. This can be deduced from equation (3.10) and (3.12), i.e. $E(\epsilon_j^2) = \sigma_j^2$ $= x_j\prime\Phi x_j\ for\ all\ j = 1, ..., H$.

To get an estimate of α, one has to obtain an estimate of the left-hand vector i.e. σ^2. We consider the squares of the vector of residuals from the OLS regression of FE on X i.e. equation (3.9), as a substitute for σ^2 in equation (3.15) and the matrix Z^* is known with the help of the original X matrix. Thus it automatically fits into a regression equation and when $\hat{\epsilon}_j^2$ is regressed on Z^* one can found the estimate of $\hat{\alpha}$. This $\hat{\alpha}$, in turn, gives the estimate of σ^2 as

$$\hat{\sigma^2} = \begin{bmatrix} \hat{\sigma}_1^2 \\ . \\ . \\ . \\ \hat{\sigma}_H^2 \end{bmatrix} = Z^* \hat{\alpha} \tag{3.16}$$

[6]Here K=10 has been considered which makes Z^* with dimension H×66 where H is the number of households.

However, to find an estimate of α, there can be four methods based on the assumption of σ^2 function (or indirectly equation 3.8). They are as follows:

- Robust standard error or White standard error.
- Off-diagonal elements of Φ are zero then $Z*$ will be a function of only the square of x_i (same as Johnston (1984)).
- Φ will have all the elements positive then the $Z*$ will be the same as (3.15). However, the problem with this method is that after re-estimating $\hat{\sigma_j^2}$ we may get some of the values of $\hat{\sigma_j^2}$ as negative, which is not desirable. This may be due to some collinearity between the regressors in Z^*.
- Principal Component Model for σ^2.

The central idea of principal component analysis (PCA) is to reduce the dimensionality of a dataset consisting of a large number of inter-related variables while retaining as much as possible of the variation present in the dataset. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables (Jolliffe, 2002, pp.-1)

Therefore, the modified model for σ_j^2 is

$$\sigma_j^2 = \gamma_0 + \gamma_1 P_{1j} + \gamma_2 P_{2j} + ... + \gamma_m P_{mj} + v_j \tag{3.17}$$

In equation (3.17), the set of new variables are $P_{1j}, P_{2j}, ...P_{mj}$ where, P_{ij} is the i^{th} principal component variable for the j^{th} household. These 'm' components, much smaller than the number of regressors[7] in equation (3.12), absorbs and accounts for the maximum pos-

[7]It is seen that the number of principal components is maximum 10 which is much lower than 66 variables considered in equation (3.12).

sible proportion of the total variation in the set of all regressors of Z^*.

The next task is to find the number of PCs to be retained in the analysis. This is known as 'stopping rule'. If we do not include the correct number of PCs then, either relevant information is lost (under-estimation) or noise is included (over-estimated), causing a distortion in underlying patterns of variation/covariation (see Ferré, 1995; Lawrence and Hancock, 1999, for a discussion). A variety of solutions to this problem are given in the literature (Jackson, 2005; Jolliffe, 2002). The most popular among data analysts is Kaiser-Guttman method and it is what has been used here. In this method, only those components for which the eigenvalue is greater than 1 are retained and others are discarded. A satisfactory result is found after incorporating this method. The estimated value of $\hat{\sigma}_j^2$ is non-negative for almost all j.

Using the above methods of finding the value of $\hat{\sigma}_j^2$ in the first step, $\hat{\Omega}$ can be calculated. Substituting $\hat{\Omega}^{-1}$ in equation (3.14) the required $\hat{\beta}$, the feasible generalized regression estimate of β, can be found in the next step.

To see if there is any gender difference in the average food expenditure for different age groups separately we will go for the Wald test. There will be altogether 7 hypotheses to be tested: 1) All $\beta = 0$ 2) All $\beta_{male} = \beta_{female}$ 3) $\beta_{m0003} = \beta_{f0003}$ 4) $\beta_{m0406} = \beta_{f0406}$ 5) $\beta_{m0712} = \beta_{f0712}$ 6) $\beta_{m1318} = \beta_{f1318}$ 7) $\beta_{m19+} = \beta_{f19+}$. The Wald tests yield statistics that are asymptotically distributed as F under the null.

3.3 Data Description

Indian Household Consumer Expenditure Survey data of the National Sample Survey Office (NSSO) under Ministry of Statistics and Programme Implementation has been used for the current study. NSSO conduct various large-scale sample surveys in diverse fields on all India basis. The NSS surveys on consumer expenditure were collected quinquennially from the 27^{th} round (October 1972-September 1973) onward. The last survey was done on July 2011-June 2012, which was the ninth quinquennial survey and it was the 68^{th} round. For the current analysis, the 61^{st} (2004-05) and the 66^{th} round (2009-10) data[8] has been used. These are the 7^{th} and 8^{th} quinquennial survey of its series. Both of these data sets are individual cross-section data. The survey is designed such that it collects detailed data on (1) Demographic characteristics of household members such as relation to the head, sex, age, marital status, education level of each member of the household etc., (2) Household characteristics: household size and type, ownership and type of land, primary source of energy used for cooking and lighting, availability of regular salary earner in the household etc., and (3) Consumption and Expenditure: this includes all the expenditure made on different items separately by the household during a specified period, called the reference period. The goods and services which are not produced or purchased but procured otherwise for consumption for that the imputed values of the goods are considered.

The unit of observation is household as we can only get consumption at the household level unless a special type of data is collected on an individual basis for the consumption expenditure (Basu et al. (1986)). The sample design of the survey follows a stratified multistage design. The total sample size for the 61^{st} round is 1,24,644 households, where the rural households are 79,298 in number, i.e. 63.62 % of the total sample, and the rest

[8]For this analysis, 66^{th} round Type I data has been used to make the results compatible with the 61^{st} round as both have the same reference period(30 days and 365 days) for data collection on consumption.

of the households are urban, i.e. 45346 or 36.38%. Similarly, for 66^{th} round NSS data the total sample of 1,00,855 households were collected out of which 59,119 (or 58.62% of the total) were rural and the rest, i.e. 41736 (41.38%), were urban.

For the outliers detection, the method of 'standard residual technique' is used. The range of which is fixed at ± 3. Also, the upper bound of the monthly per capita consumer expenditure(MPCE) is fixed for both the rural as well as the urban sectors in both the rounds. Therefore, the total loss of data is below 4% in case of both the rounds. Individuals members of the households are considered in 5 age groups, i.e. 0-3, 4-6, 7-12, 13-18, and 19 and above categories. For genders, we separate the variables thereby making a total of 10 age-sex groups for every household. Table 3.1 shows the different monthly per capita consumer expenditure (MPCE) (URP) class according to 10 decile groups from the poorest 10% to the richest 10% of the population. The range of the MPCE is shown for both the rural and urban sectors in both the rounds which is given by NSS. The truncation for the highest expenditure group is also shown. We have done this truncation because they are the richest people of them all and their food expenditure does not increase at the same rate as their income. They might distort the result for the last group. The NSS data provides a multiplier to each household. These are the sampling weights and are calculated from the sampling scheme adopted by NSS. We have used these weights or multipliers for the calculation of the estimates.

The summary statistics of the variables used are presented in Tables 3.2 to 3.5. From the tables, we can see that the average household size decreased for both sector over the years, i.e. in the rural sector it has decreased from 4.89 to 4.71, and for the urban sector, it was 4.43 in 2004-05 which decreased to 4.24 in 2009-10. The average per capita food expenditure was Rs.313.44 for the rural sector and Rs.438.09 for the urban sector in

2004-05, which increased to Rs.502.87 in the rural sector and Rs.697.55 in the urban sector.

The MPCE also increases over the years for both the sectors. The rest of the variables is the number of members in the different age-sex category. The average number of members for approximately all the age-sex groups have seen a downward movement over the years except for the higher age group category.

3.4 Empirical Investigation

The method of estimation proposed in section 3.2 is done for both the rounds of NSS and for both sectors separately. The various methods of estimations[9] are shown in Table 3.6 and 3.7 for the rural and urban sectors respectively of the 61^{st} round. All the four models are compared. The four different assumptions for σ^2 produces four types of results which are shown in each column. The second section of these tables show few statistics. In the third section of these tables, the result of the hypothesis testing is produced, i.e. F-value, while comparing the models we can see that for the case of PCA we have the minimum value of AIC and BIC, therefore, PCA is superior to other methods of estimation. Not only that using PCA we do not get any negative values of estimated σ^2 which is not true for the third model. All the coefficients of the regression model are statistically significant at 1% level for all the models. The result of hypothesis testing shows that there are significant differences in food expenditure over the gender as well as age. For the age groups—0-3 years and 4-6 years, we found a significant gender difference in food expenditure similar result is also seen for the adults using PCA method. In the urban

[9]To estimate σ_j^2 there are four methods. The first method is the famous White Standard Error or, Huber-White standard errors, the second can be similar to Random coefficient model (see Johnston (1984)), where the off-diagonal element of the matrix in equation (3.8) are all zero. The third method is nothing but directly estimating equation (3.12) and lastly the PCA method. Of all the methods, PCA method is the most efficient.

sector, during 2004-05, the result obtained is similar in case of PCA method.

The 66^{th} round results are presented in Table 3.10 and 3.11. Here also PCA method of estimation techniques for σ^2 is a superior method for getting FGLS estimate of β. The result of hypothesis testing shows a significant gender difference in food expenditure for only 13-18 and 19 years and above age categories in case of the rural sector whereas, in case of the urban sector, age groups 7-12 and 19 years and above shows a significant difference in food expenditure. From these tables, it is clear that PCA is a more appropriate method to get an estimate of σ^2. Therefore, for the estimations of the model for various expenditure classes (shown in Table 3.1), we will only report the FGLS estimate of β, where PCA method is used to obtain σ^2.

In Table 3.8, regression result for food expenditure using FGLS method for separate MPCE groups are produced under separate columns. We found the gender difference in food expenditure for mostly higher income groups and it is seen significant in case of the adults. For 0-3 and 4-6 age groups, gender difference exists for higher expenditure groups in case of the rural sector. In case of the urban sector, we found few cases where there is a significant gender difference in food expenditure. However, when we take all expenditure groups together, i.e. the last column, we found significant gender difference for both the children age groups i.e. 0-3 years and 4-6 years.

The regression result for the 66^{th} round are presented in the Tables 3.12 and 3.13 for rural and urban sectors respectively. Here significant gender difference is obtained for last two age groups i.e. 13-18 and 19 years and above categories when all the expenditure groups are taken together. Group 4 and 9 of income class also shows a significant gender difference in food expenditure for the adult group. Last but not the least, in case of the urban sector, we found more traces of gender difference in food expenditure for approximately all age groups. But mostly it is seen in the case of adults. The results are similar to what obtained by Lancaster et al. (2008). They also found a significant gender difference in food expenditure for the adults.

Most of the significant cases, where there is gender bias, are in the age groups 19 years and above. For this age group, the gender bias is in favour of the male for almost every expenditure groups in each round and sectors. An increasing trend in the average food expenditure from lower expenditure group to higher expenditure group is observed for each age-sex group and also for each sector and the rounds. And this average food expenditure also increases with the age for all types of expenditure groups.

3.5 Concluding Remarks

This chapter uses the 'Regression Decomposition Technique' by which it is possible to find out the age-sex composition wise food consumption expenditure from the total food expenditure of the household. However, the model has heteroscedasticity problem. Four different models were assumed to estimate the variance of the error term (or, sigma square). In the next step, the feasible generalised least square (FGLS) estimate of β is estimated using all the four different models of sigma squares. Comparing these models it can be seen that Principal Component Analysis (PCA) gives the most suitable estimate of sigma square, where not a single negative values are found for sigma square, which is desirable.

The parameters are estimated for the NSS 61^{st} and 66^{th} round data. We have found two important results via this analysis. Firstly, PCA method is superior to other models for estimating sigma square. Secondly, evidence of significant gender difference in food expenditure is found for the adult groups. In 2004-05 i.e. using NSS 61^{st} round, we found a significant gender difference in food expenditure for the children age group as well.

3.6 APPENDIX

Table 3.1: MPCE(URP) decile groups for rural and urban sector for 61^{st} and 66^{th} round of NSS

	61st round		66th round	
MPCE	Rural	Urban	Rural	Urban
Gr-1	0-270	0-395	0-450	0-642
Gr-2	270-320	395-485	450-537	642-797
Gr-3	320-365	485-580	537-613	797-945
Gr-4	365-410	580-675	613-685	945-1114
Gr-5	410-455	675-790	685-765	1114-1307
Gr-6	455-510	790-930	765-853	1307-1543
Gr-7	510-580	930-1100	853-974	1543-1843
Gr-8	580-690	1100-1380	974-1144	1843-2303
Gr-9	690-890	1380-1880	1144-1477	2303-3166
Gr-10	890-2000	1880-3000	1477-3000	3166-4500
Gr-11	0-2000	0-3000	0-3000	0-4500

[a] Values are in INR.

[b] Only the last decile groups have an upper bound.

Table 3.2: Summary statistics for the 61^{st} round rural sector.

Variable	Description	Obs	Mean	Std. Dev.
fe	Food expenditure (in INR)	76898	1532.00	849.24
m0003	No. of male with 0-3 years age	76898	0.21	0.48
f0003	No. of female with 0-3 years age	76898	0.20	0.47
m0406	No. of male with 4-6 years age	76898	0.20	0.45
f0406	No. of female with 4-6 years age	76898	0.18	0.43
m0712	No. of male with 7-12 years age	76898	0.39	0.66
f0712	No. of female with 7-12 years age	76898	0.34	0.63
m1318	No. of male with 13-18 years age	76898	0.33	0.60
f1318	No. of female with 13-18 years age	76898	0.29	0.57
m19+	No. of male with age 19 year & above	76898	1.36	0.87
f19+	No. of female with age 19 years & above	76898	1.38	0.74
hhsize	household size	76898	4.89	2.40
pcfe	Per capita food expenditure (in INR)	76898	313.44	128.88
MPCE(URP)	Monthly per capita expenditure (in INR) (Uniform Reference Period)	76898	517.77	264.48

[a] Values are rounded to two decimal place.
[b] Households with MPCE(URP)>2000 has been dropped before the analysis.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.3: Summary statistics for the 61^{st} round urban sector.

Variable	Description	Obs	Mean	Std. Dev.
fe	Food expenditure (in INR)	43208	1939.88	1010.83
m0003	No. of male with 0-3 years age	43208	0.15	0.41
f0003	No. of female with 0-3 years age	43208	0.14	0.39
m0406	No. of male with 4-6 years age	43208	0.14	0.39
f0406	No. of female with 4-6 years age	43208	0.12	0.35
m0712	No. of male with 7-12 years age	43208	0.29	0.57
f0712	No. of female with 7-12 years age	43208	0.25	0.54
m1318	No. of male with 13-18 years age	43208	0.31	0.58
f1318	No. of female with 13-18 years age	43208	0.27	0.56
m19+	No. of male with age 19 years & above	43208	1.42	0.89
f19+	No. of female with age 19 years & above	43208	1.33	0.84
hhsize	household size	43208	4.43	2.29
pcfe	Per capita food expenditure (in INR)	43208	438.09	192.81
MPCE(URP)	Monthly per capita expenditure (in INR) (Uniform reference period)	43208	911.11	530.81

[a] Values are rounded to two decimal place.
[b] Households with MPCE(URP)>3000 has been dropped before the analysis.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.4: Summary statistics for the 66^{th} round rural sector.

Variable	Description	Obs	Mean	Std. Dev.
fe	Food expenditure (in INR)	56919	2368.23	1236.86
m0003	No. of male with 0-3 years age	56919	0.17	0.43
f0003	No. of female with 0-3 years age	56919	0.16	0.42
m0406	No. of male with 4-6 years age	56919	0.17	0.42
f0406	No. of female with 4-6 years age	56919	0.16	0.40
m0712	No. of male with 7-12 years age	56919	0.36	0.64
f0712	No. of female with 7-12 years age	56919	0.31	0.60
m1318	No. of male with 13-18 years age	56919	0.35	0.63
f1318	No. of female with 13-18 years age	56919	0.29	0.57
m19+	No. of male with age 19 years & above	56919	1.37	0.84
f19+	No. of female with age 19 years & above	56919	1.38	0.72
hhsize	household size	56919	4.71	2.23
pcfe	Per capita food expenditure (in INR)	56919	502.87	201.46
MPCE(URP)	Monthly per capita expenditure (in INR) (Uniform reference period)	56919	860.31	421.23

[a] Values are rounded to two decimal place.
[a] Values are rounded to two decimal place.
[b] Households with MPCE(URP)>3000 has been dropped before the analysis.
Source: Author's calculation based on 2009-10 Type-1 Consumer Expenditure Household Survey, collected by NSS.

Table 3.5: Summary statistics for the 66^{th} round urban sector.

Variable	Description	Obs	Mean	Std. Dev.
fe	Food expenditure (in INR)	39375	2954.80	1508.62
m0003	No. of male with 0-3 years age	39375	0.13	0.37
f0003	No. of female with 0-3 years age	39375	0.12	0.37
m0406	No. of male with 4-6 years age	39375	0.13	0.36
f0406	No. of female with 4-6 years age	39375	0.11	0.34
m0712	No. of male with 7-12 years age	39375	0.26	0.53
f0712	No. of female with 7-12 years age	39375	0.23	0.52
m1318	No. of male with 13-18 years age	39375	0.29	0.57
f1318	No. of female with 13-18 years age	39375	0.24	0.53
m19+	No. of male with age 19 years & above	39375	1.40	0.88
f19+	No. of female with age 19 years & above	39375	1.32	0.83
hhsize	household size	39375	4.24	2.19
pcfe	Per capita food expenditure (in INR)	39375	697.55	309.57
MPCE(URP)	Monthly per capita expenditure (in INR) (Uniform reference period)	39375	1480.24	842.10

[a] Values are rounded to two decimal place.

[b] Households with MPCE(URP)>4500 has been dropped before the analysis.

Source: Author's calculation based on 2009-10 Type-1 Consumer Expenditure Household Survey, collected by NSS.

Table 3.6: Food Expenditure models with different assumption for heteroscedasticity in case of 61^{st} round rural sector.

variables	(1) OLS(ROBUST)	(2) with square terms of x	(3) with all x combination	(4) PCA
m0003	112.27***	112.75***	112.87***	111.72***
	(6.98)	(6.16)	(6.53)	(6.15)
f0003	73.76***	88.61***	82.11***	87.64***
	(6.76)	(6.10)	(6.15)	(6.10)
m0406	149.79***	142.92***	135.26***	147.78***
	(7.25)	(6.48)	(6.78)	(6.41)
f0406	130.87***	129.78***	120.59***	131.64***
	(7.45)	(6.81)	(7.24)	(6.78)
m0712	181.86***	179.28***	150.32***	177.60***
	(4.85)	(4.54)	(6.87)	(4.52)
f0712	183.88***	168.57***	150.47***	170.19***
	(5.16)	(4.92)	(5.87)	(4.89)
m1318	242.37***	223.04***	189.80***	225.25***
	(5.49)	(5.27)	(6.28)	(5.23)
f1318	241.68***	217.34***	181.50***	219.63***
	(5.86)	(5.70)	(6.75)	(5.64)
m19+	366.65***	336.26***	289.62***	333.51***
	(4.45)	(4.68)	(6.87)	(4.74)
f19+	329.69***	265.75***	196.59***	267.57***
	(5.20)	(6.32)	(9.30)	(6.28)
_cons	202.86***	0.63***	0.99***	0.62***
	(7.66)	(0.02)	(0.04)	(0.02)
N	76898	76898	76895	76898
adj. R^2	0.522	0.192	0.105	0.191
F	2738.12	1058.34	440.04	1061.05
AIC	1198749.95	216245.64	219268.05	215698.04
BIC	1198851.71	216347.40	219369.80	215799.79
RMSE	587.22	0.99	1.01	0.98
Hypothesis Testing				
All β=0	1750.69***	1161.89***	422.80***	1165.06***
All β_male=β_female	8.13***	23.12***	39.03***	20.91***
β_m0003=β_f0003	16.29 ***	8.24***	12.66***	8.18***
β_m0406=β_f0406	3.51*	2.10	2.61	3.20*
β_m0712=β_f0712	0.08	2.45	0.00	1.19
β_m1318=β_f1318	0.01	0.52	0.90	0.51
β_m19+=β_f19+	21.58***	103.15***	178.01***	92.23***

[a] Hypothesis testing result corresponds to the β value in the upper section of the table.
[b] Households with MPCE(URP)>2000 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] 10 principal components explains 91% variation in case of PCA estimation technique.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.7: Food Expenditure models with different assumption for heteroscedasticity in case of 61^{st} round urban sector.

variables	(1) OLS(ROBUST)	(2) with square terms of x	(3) with all x combination	(4) PCA
m0003	102.48***	106.87***	67.84***	112.09***
	(16.85)	(15.77)	(14.50)	(15.60)
f0003	51.85***	46.96***	57.76***	45.42***
	(17.03)	(16.00)	(15.04)	(15.91)
m0406	154.14***	154.22***	175.82***	152.34***
	(17.76)	(17.33)	(19.18)	(17.14)
f0406	90.57***	106.65***	148.21***	97.89***
	(17.24)	(16.07)	(16.40)	(15.97)
m0712	178.20***	177.08***	168.09***	176.74***
	(11.00)	(10.86)	(11.84)	(10.76)
f0712	177.54***	182.65***	129.92***	180.88***
	(12.29)	(12.72)	(13.36)	(12.52)
m1318	245.32***	231.63***	189.73***	228.76***
	(11.00)	(11.22)	(13.65)	(11.06)
f1318	213.61***	204.88***	178.84***	203.25***
	(11.58)	(11.61)	(11.81)	(11.41)
m19+	411.96***	382.06***	339.06***	378.27***
	(8.29)	(8.68)	(10.81)	(8.74)
f19+	408.08***	354.36***	247.13***	347.59***
	(8.86)	(9.41)	(11.42)	(9.39)
_cons	526.54***	0.91***	1.23***	0.93***
	(13.92)	(0.03)	(0.04)	(0.03)
N	43208	43208	43195	43208
adj. R^2	0.466	0.217	0.153	0.210
F	983.96	507.32	149.19	484.38
AIC	693374.62	122084.54	121792.36	121780.36
BIC	693470.03	122179.95	121887.77	121875.77
RMSE	738.50	0.99	0.99	0.99
Hypothesis Testing				
All β=0	491.77***	370.99***	149.33***	362.15***
All β_male=β_female	2.70**	3.07***	10.99 ***	3.76***
β_m0003=β_f0003	4.22**	6.52***	0.22	8.23***
β_m0406=β_f0406	6.71***	3.92**	1.17	5.22**
β_m0712=β_f0712	0.00	0.09	4.72**	0.05
β_m1318=β_f1318	3.47*	2.40	0.34	2.28
β_m19+=β_f19+	0.08	4.56**	50.57***	5.67**

[a] Hypothesis testing result corresponds to the β value from the upper section of the table.
[b] Households with MPCE(URP)>3000 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] 9 principal components explains 90% variation in case of PCA estimation technique.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.8: Regression result for Food Expenditure using FGLS for different MPCE classes in case of 61^{st} round rural sector.

	(1) Gr-1	(2) Gr-2	(3) Gr-3	(4) Gr-4	(5) Gr-5	(6) Gr-6	(7) Gr-7	(8) Gr-8	(9) Gr-9	(10) Gr-10	(11) Gr-11
m0003	190.74***	202.40***	217.97***	258.98***	275.45***	305.23***	296.31***	318.10***	339.11***	287.42***	103.14***
	(6.18)	(5.27)	(6.31)	(7.07)	(8.43)	(9.40)	(11.40)	(12.67)	(16.60)	(33.36)	(6.30)
f0003	182.69***	195.18***	223.28***	244.86***	263.40***	286.98***	290.24***	324.18***	298.42***	384.65***	79.53***
	(6.13)	(5.06)	(6.16)	(8.02)	(8.14)	(10.32)	(11.33)	(13.05)	(18.84)	(36.04)	(6.21)
m0406	201.92***	215.57***	225.34***	255.88***	279.41***	311.20***	348.32***	364.61***	397.70***	409.17***	140.86***
	(6.04)	(5.47)	(6.55)	(7.01)	(8.70)	(9.71)	(11.63)	(13.91)	(18.40)	(39.37)	(6.50)
f0406	206.82***	208.99***	227.87***	245.05***	278.35***	278.51***	330.79***	322.70***	339.79***	392.16***	128.47***
	(6.72)	(5.95)	(6.99)	(7.50)	(9.99)	(10.70)	(13.76)	(16.53)	(22.57)	(35.15)	(7.72)
m0712	199.42***	208.46***	244.91***	267.20***	280.54***	307.57***	321.69***	346.87***	366.79***	372.36***	170.36***
	(4.56)	(4.37)	(4.37)	(5.27)	(6.63)	(6.78)	(8.13)	(9.38)	(12.09)	(20.87)	(4.60)
f0712	194.49***	218.72***	235.04***	266.74***	277.07***	306.77***	324.46***	356.87***	374.99***	331.99***	164.72***
	(4.85)	(4.03)	(5.47)	(5.75)	(7.07)	(7.77)	(8.87)	(10.69)	(14.36)	(24.69)	(4.98)
m1318	186.37***	217.80***	233.85***	258.50***	279.07***	307.97***	339.40***	353.36***	375.76***	445.40***	219.19***
	(5.45)	(5.21)	(5.22)	(5.43)	(6.25)	(7.10)	(7.25)	(8.14)	(10.76)	(17.11)	(5.55)
f1318	196.98***	208.83***	228.97***	272.82***	289.17***	310.69***	329.50***	356.72***	360.87***	429.72***	212.84***
	(5.72)	(5.14)	(5.79)	(6.64)	(6.67)	(7.01)	(8.31)	(10.11)	(12.76)	(18.65)	(5.71)
m19+	210.06***	215.83***	245.71***	286.75***	302.33***	328.46***	370.07***	401.51***	449.46***	524.63***	327.07***
	(4.39)	(3.95)	(3.69)	(4.72)	(4.55)	(5.29)	(4.86)	(5.88)	(7.94)	(11.15)	(4.98)
f19+	218.84***	220.46***	230.06***	265.84***	287.11***	321.57***	333.63***	376.62***	414.86***	497.21***	261.44***
	(5.17)	(4.61)	(5.26)	(5.72)	(5.14)	(6.64)	(7.42)	(7.43)	(10.20)	(13.25)	(6.40)
_cons	-1.46***	-0.31***	-0.06	-0.21***	-0.01	-0.04	0.07	0.19***	0.36***	0.61***	0.68***
	(0.09)	(0.09)	(0.08)	(0.07)	(0.06)	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.02)
N	4501	4874	5646	6109	6127	7179	8389	9898	11001	12685	78381
adj. R^2	0.670	0.727	0.737	0.724	0.703	0.695	0.699	0.663	0.566	0.465	0.193
Hypothesis testing											
All β=0	584.27***	942.54***	1080.73***	1068.97***	1090.48***	1338.86***	1328.04***	1291.67***	740.99***	555.68***	969.13***
All β_male=β_female	1.16	1.56	1.73	2.98***	1.44	1.52	3.66***	2.32**	2.92***	1.65	17.65***
β_m0003=β_f0003	1.21	1.12	0.41	2.07	1.03	1.69	0.14	0.11	2.41	4.14***	7.69***
β_m0406=β_f0406	0.40	0.79	0.09	1.15	0.01	5.07**	0.98	3.54*	4.02**	0.10	1.61
β_m0712=β_f0712	0.77	3.46*	2.41	0.00	0.14	0.01	0.05	0.46	0.17	1.48	0.68
β_m1318=β_f1318	1.88	1.55	0.40	2.61	1.31	0.07	0.78	0.06	0.81	0.36	0.60
β_m19+=β_f19+	1.66	0.61	5.58**	7.41***	4.65**	0.64	16.54***	6.60***	7.97***	2.47	77.93***

[a] Hypothesis testing result corresponds to the β value in the upper section of the table.
[b] Households with MPCE(URP)>2000 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] PCA has been used to estimate sigma square with almost 10 principal components and the explained variation is approximately 90%.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.9: Regression result for Food Expenditure using FGLS for different MPCE classes in case of 61^{st} round urban sector.

	(1) Gr-1	(2) Gr-2	(3) Gr-3	(4) Gr-4	(5) Gr-5	(6) Gr-6	(7) Gr-7	(8) Gr-8	(9) Gr-9	(10) Gr-10	(11) Gr-11
m0003	224.78***	257.92***	287.41***	331.13***	319.12***	376.67***	492.92***	383.07***	525.73***	457.60***	90.10***
	(15.35)	(14.30)	(15.94)	(19.03)	(27.22)	(25.99)	(30.43)	(40.30)	(53.69)	(94.95)	(15.69)
f0003	199.14***	268.16***	282.51***	315.75***	332.11***	374.22***	396.17***	453.23***	521.64***	629.11***	27.10
	(12.58)	(12.94)	(16.09)	(17.95)	(24.76)	(28.26)	(42.40)	(37.18)	(60.42)	(81.54)	(17.08)
m0406	214.76***	266.52***	271.43***	336.90***	325.00***	410.32***	413.22***	390.90***	482.36***	550.00***	129.36***
	(12.36)	(14.41)	(16.43)	(16.99)	(30.75)	(23.73)	(29.38)	(37.65)	(55.07)	(91.60)	(17.54)
f0406	198.22***	268.09***	304.03***	312.60***	321.67***	365.12***	336.58***	484.28***	467.04***	641.05***	73.54***
	(14.25)	(12.12)	(16.97)	(21.89)	(31.47)	(28.56)	(41.14)	(45.39)	(51.63)	(113.37)	(16.35)
m0712	206.06***	268.80***	296.79***	352.71***	332.54***	395.05***	393.80***	404.46***	448.36***	658.86***	157.21***
	(9.90)	(9.84)	(15.17)	(13.87)	(19.40)	(19.14)	(23.09)	(23.98)	(27.98)	(55.67)	(10.97)
f0712	228.42***	250.23***	314.05***	310.95***	360.10***	347.78***	420.96***	413.93***	565.38***	622.66***	161.60***
	(11.88)	(9.45)	(14.61)	(16.39)	(18.78)	(18.68)	(24.63)	(28.49)	(33.13)	(70.96)	(12.57)
m1318	228.16***	288.58***	328.57***	349.55***	348.49***	411.34***	433.34***	435.41***	485.24***	651.88***	209.59***
	(8.80)	(9.84)	(12.81)	(13.55)	(17.96)	(16.48)	(23.72)	(25.70)	(30.92)	(52.43)	(11.57)
f1318	224.87***	272.60***	297.19***	337.26***	361.60***	378.73***	350.15***	432.65***	492.88***	494.58***	183.46***
	(9.47)	(10.59)	(13.06)	(15.84)	(16.99)	(18.03)	(22.39)	(25.77)	(32.56)	(51.93)	(11.66)
m19+	262.79***	291.09***	343.78***	360.43***	400.10***	415.11***	476.23***	494.01***	563.76***	664.06***	357.22***
	(6.95)	(6.16)	(10.86)	(10.05)	(11.38)	(14.30)	(17.40)	(18.95)	(20.98)	(35.68)	(9.35)
f19+	250.99***	275.50***	300.83***	336.40***	381.28***	387.45***	450.27***	459.27***	475.42***	687.18***	336.57***
	(8.23)	(10.72)	(12.42)	(11.08)	(10.49)	(13.92)	(11.39)	(16.73)	(18.10)	(23.47)	(10.37)
_cons	-0.48***	-0.01	0.14	0.30***	0.39***	0.58***	0.50***	0.89***	1.00***	0.80***	1.03***
	(0.09)	(0.12)	(0.14)	(0.10)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.03)
N	4790	4619	4585	3945	3884	3759	3713	4516	5175	4024	44789
adj. R^2	0.708	0.681	0.566	0.597	0.659	0.675	0.693	0.582	0.569	0.625	0.202
Hypothesis testing											
All β=0	522.61***	458.04***	220.87***	330.71***	327.08***	401.12***	351.75***	291.56***	252.38***	206.62***	401.89***
All β_male=β_female	1.61	1.03	3.25***	1.51	0.55	1.48	3.31***	1.02	3.78***	1.37	2.85***
β_m0003=β_f0003	2.06	0.33	0.05	0.34	0.13	0.00	3.59*	1.62	0.00	2.00	6.72***
β_m0406=β_f0406	0.68	0.01	2.15	0.78	0.01	1.42	2.33	2.27	0.04	0.39	5.24***
β_m0712=β_f0712	1.66	2.31	0.69	3.38*	0.92	2.56	0.62	0.06	7.04***	0.14	0.06
β_m1318=β_f1318	0.07	1.28	2.80*	0.35	0.29	1.79	7.34***	0.01	0.03	4.17**	2.21
β_m19+=β_19+	1.01	1.43	8.79***	2.88*	1.47	2.04	1.70	2.06	12.03***	0.29	2.09

[a] Hypothesis testing result corresponds to the β value in the upper section of the table.
[b] Households with MPCE(URP)>3000 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] PCA has been used to estimate sigma square with no. of principal components between 4-10 and the explained variation is approximately 90%.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.10: Food Expenditure models with different assumption for heteroscedasticity in case of 66^{th} round rural sector.

variables	(1) OLS(ROBUST)	(2) with square terms of x	(3) with all x combination	(4) PCA
m0003	180.71***	184.47***	198.72***	177.46***
	(17.74)	(15.76)	(19.53)	(15.78)
f0003	175.43***	184.83***	262.07***	172.88***
	(17.37)	(15.20)	(30.47)	(15.28)
m0406	258.77***	254.95***	195.28***	252.86***
	(18.06)	(16.86)	(18.98)	(16.61)
f0406	221.80***	247.42***	209.41***	254.44***
	(19.10)	(17.68)	(20.93)	(17.67)
m0712	312.84***	299.52***	270.38***	286.99***
	(12.11)	(11.33)	(15.55)	(11.32)
f0712	309.55***	287.51***	246.86***	276.82***
	(12.93)	(12.29)	(15.15)	(12.28)
m1318	356.57***	337.94***	253.45***	333.65***
	(12.12)	(12.13)	(15.93)	(11.99)
f1318	334.09***	293.47***	319.02***	280.48***
	(13.87)	(13.50)	(30.72)	(13.43)
m19+	502.79***	447.41***	367.59***	436.32***
	(11.05)	(11.40)	(17.59)	(11.60)
f19+	492.88***	375.35***	266.23***	366.51***
	(11.95)	(14.03)	(22.41)	(13.93)
_cons	433.34***	0.81***	1.18***	0.86***
	(18.89)	(0.03)	(0.07)	(0.04)
N	56919	56919	56896	56919
adj. R^2	0.481	0.190	0.136	0.169
F	960.35	410.86	146.83	359.52
AIC	934749.30	160541.71	163776.89	160241.86
BIC	934847.74	160640.16	163875.33	160340.30
RMSE	890.86	0.99	1.02	0.99
Hypothesis Testing				
All β=0	622.65***	438.02***	150.01***	387.94***
All β_male=β_female	0.73	5.01***	9.25***	5.55***
β_m0003=β_f0003	0.05	0.00	3.77**	0.04
β_m0406=β_f0406	2.10	0.10	0.25	0.00
β_m0712=β_f0712	0.03	0.50	1.20	0.37
β_m1318=β_f1318	1.44	6.08***	3.15*	8.95***
β_19+=β_f19+	0.28	19.01***	39.23***	19.00***

[a] Hypothesis testing result corresponds to the β value in the upper section of the table.
[b] Households with MPCE(URP)>3000 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] 10 principal components explains 91% variation in case of PCA estimation technique.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.11: Food Expenditure models with different assumption for heteroscedasticity in case of 66^{th} round urban sector.

variables	(1) OLS(ROBUST)	(2) with square terms of x	(3) with all x combination	(4) PCA
m0003	197.09***	174.72***	181.05***	170.22***
	(30.41)	(29.66)	(38.05)	(29.41)
f0003	87.70***	132.99***	141.50***	122.27***
	(27.74)	(26.72)	(33.01)	(26.60)
m0406	222.90***	231.15***	236.45***	238.73***
	(29.23)	(28.67)	(61.68)	(28.21)
f0406	222.16***	175.05***	155.89***	204.14***
	(32.54)	(31.69)	(38.52)	(31.06)
m0712	330.43***	349.73***	307.01***	342.30***
	(18.03)	(17.88)	(30.80)	(17.69)
f0712	317.30***	285.48***	215.59***	266.82***
	(20.62)	(20.62)	(23.50)	(20.33)
m1318	320.15***	301.35***	226.70***	296.95***
	(18.33)	(18.53)	(19.69)	(18.34)
f1318	278.28***	267.00***	236.02***	263.40***
	(19.71)	(19.24)	(25.64)	(18.97)
m19+	582.36***	526.58***	437.16***	515.18***
	(13.56)	(14.08)	(24.04)	(14.09)
f19+	603.79***	474.32***	227.38***	464.40***
	(14.87)	(16.34)	(24.91)	(16.31)
_cons	931.98***	1.09***	1.56***	1.12***
	(22.60)	(0.03)	(0.07)	(0.03)
N	39375	39375	39346	39375
adj. R^2	0.433	0.188	0.102	0.178
F	774.08	347.94	66.67	325.26
AIC	665804.51	111152.91	111535.87	110886.27
BIC	665898.90	111247.30	111630.25	110980.66
RMSE	1136.36	0.99	1.00	0.99
Hypothesis testing				
All β=0	369.20***	281.36***	73.45***	268.85***
All β_male=β_female	1.98*	2.98***	14.51***	3.19***
β_m0003=β_f0003	7.15***	1.09	0.50	1.47
β_m0406=β_f0406	0.00	1.68	2.49	0.66
β_m0712=β_f0712	0.20	4.94**	7.18***	7.04***
β_m1318=β_f1318	2.15	1.49	0.11	1.47
β_m19+=β_f19+	0.85	5.61**	67.65***	5.41**

[a] Hypothesis testing result corresponds to the β value from the upper section of the table.
[b] Households with MPCE(URP)>4500 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] 9 principal components explains 91% variation in case of PCA estimation technique.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.12: Regression result for Food Expenditure using FGLS for different MPCE classes in case of 66^{th} round rural sector.

	(1) Gr-1	(2) Gr-2	(3) Gr-3	(4) Gr-4	(5) Gr-5	(6) Gr-6	(7) Gr-7	(8) Gr-8	(9) Gr-9	(10) Gr-10	(11) Gr-11
m0003	243.75***	341.65***	379.03***	402.75***	427.87***	484.27***	439.75***	485.50***	511.43***	556.83***	159.47***
	(19.38)	(17.39)	(18.53)	(23.28)	(23.87)	(24.27)	(30.56)	(41.86)	(47.33)	(69.60)	(16.06)
f0003	286.63***	347.56***	354.14***	383.19***	420.86***	453.36***	503.89***	458.31***	558.62***	529.38***	156.46***
	(17.28)	(15.09)	(19.96)	(23.64)	(20.95)	(27.90)	(31.59)	(52.43)	(53.47)	(80.99)	(15.32)
m0406	298.37***	353.96***	399.58***	395.87***	448.49***	499.20***	548.20***	588.74***	657.24***	667.49***	238.06***
	(20.75)	(15.63)	(18.56)	(19.00)	(23.02)	(27.91)	(30.42)	(35.84)	(38.22)	(91.70)	(16.80)
f0406	314.14***	326.86***	355.08***	402.16***	432.49***	442.71***	541.93***	541.41***	605.33***	684.58***	248.41***
	(20.64)	(16.86)	(18.54)	(25.57)	(25.04)	(34.30)	(36.64)	(43.34)	(46.39)	(98.61)	(17.62)
m0712	281.00***	360.99***	411.18***	455.58***	466.21***	507.14***	564.88***	588.18***	539.27***	467.93***	255.86***
	(12.91)	(12.36)	(12.18)	(14.82)	(15.72)	(16.93)	(21.53)	(19.81)	(35.46)	(58.10)	(12.51)
f0712	277.99***	356.46***	406.96***	437.80***	485.04***	499.63***	542.99***	523.39***	581.60***	587.03***	254.54***
	(13.09)	(13.65)	(11.18)	(16.87)	(17.33)	(20.96)	(23.85)	(28.29)	(33.64)	(69.56)	(13.31)
m1318	284.05***	361.98***	394.16***	387.22***	461.21***	452.75***	515.20***	543.27***	555.17***	687.26***	327.01***
	(13.00)	(12.11)	(10.88)	(15.54)	(13.58)	(17.48)	(23.15)	(23.66)	(26.76)	(39.58)	(13.39)
m1318	265.61***	363.61***	349.89***	412.65***	468.49***	459.24***	531.71***	569.00***	570.50***	622.08***	251.32***
	(18.43)	(14.39)	(13.87)	(14.82)	(15.48)	(16.97)	(24.00)	(26.27)	(34.57)	(47.93)	(14.94)
m19+	279.88***	360.93***	389.58***	449.09***	459.67***	512.78***	567.50***	586.59***	681.29***	685.79***	403.13***
	(12.10)	(12.00)	(11.30)	(14.24)	(10.31)	(14.98)	(15.37)	(16.75)	(19.70)	(34.28)	(15.21)
f19+	290.95***	324.21***	365.13***	400.65***	442.78***	480.12***	531.51***	578.89***	596.05***	645.07***	297.66***
	(15.69)	(16.31)	(14.49)	(12.74)	(16.25)	(13.57)	(18.99)	(20.77)	(26.61)	(35.12)	(23.16)
_cons	-0.43***	-0.28*	-0.00	0.04	0.18	0.19*	0.13	0.40***	0.59***	0.93***	1.05***
	(0.13)	(0.15)	(0.12)	(0.11)	(0.12)	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.06)
N	3118	3413	3935	4219	4999	5185	6318	7000	8564	9785	58520
adj. R^2	0.608	0.721	0.760	0.754	0.668	0.677	0.618	0.535	0.485	0.386	0.160
Hypothesis Testing											
All β=0	224.83***	342.17***	559.44***	442.50***	434.19***	393.94***	427.00***	338.46***	235.14***	151.09***	240.75***
All β_male=β_female	0.85	0.75	2.46**	1.96*	0.40	1.13	1.00	0.97	2.10*	0.76	7.14***
β_m0003=β_f0003	2.73*	0.08	1.02	0.38	0.05	0.74	2.19	0.15	0.43	0.07	0.02
β_m0406=β_f0406	0.43	1.67	3.17*	0.05	0.22	1.74	0.02	0.79	0.73	0.02	0.19
β_m0712=β_f0712	0.03	0.10	0.07	0.70	0.66	0.08	0.42	3.03*	0.72	1.43	0.01
β_m1318=β_f1318	0.69	0.01	6.85***	1.44	0.12	0.06	0.27	0.53	0.13	0.89	13.50***
β_m19+=β_f19+	0.29	2.35	1.95	7.05***	0.89	2.62	2.04	0.10	7.96***	1.31	26.03***

[a] Hypothesis testing result corresponds to the β value in the upper section of the table.
[b] Households with MPCE(URP)>3000 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] PCA has been used to estimate sigma square with almost 10 principal components and the explained variation is approximately 90%.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

Table 3.13: Regression result for Food Expenditure using FGLS for different MPCE classes in case of 66^{th} round urban sector.

	(1) Gr-1	(2) Gr-2	(3) Gr-3	(4) Gr-4	(5) Gr-5	(6) Gr-6	(7) Gr-7	(8) Gr-8	(9) Gr-9	(10) Gr-10	(11) Gr-11
m0003	284.88***	423.23***	489.44***	525.42***	531.04***	648.72***	797.85***	753.26***	750.53***	742.86***	109.16***
	(21.59)	(25.36)	(25.57)	(32.82)	(36.32)	(49.81)	(71.51)	(69.56)	(84.42)	(248.27)	(29.59)
f0003	306.84***	451.12***	430.66***	486.63***	440.90***	582.31***	582.66***	699.06***	727.79***	753.89***	74.34**
	(21.44)	(19.58)	(27.40)	(31.95)	(44.13)	(66.02)	(56.93)	(87.01)	(115.96)	(161.13)	(31.79)
m0406	297.24***	449.99***	425.04***	512.10***	484.30***	591.99***	673.39***	820.67***	777.76***	1029.25***	208.21***
	(21.37)	(24.91)	(25.11)	(30.43)	(42.26)	(48.44)	(63.14)	(65.42)	(86.45)	(166.97)	(33.28)
f0406	336.37***	347.89***	488.92***	565.00***	540.36***	462.60***	562.32***	719.83***	920.36***	1003.59***	175.44***
	(20.35)	(30.02)	(30.79)	(30.96)	(51.86)	(54.91)	(80.17)	(79.20)	(100.05)	(187.45)	(32.80)
m0712	350.21***	449.33***	468.49***	542.59***	603.91***	574.10***	584.99***	656.67***	686.73***	1050.71***	299.07***
	(16.63)	(16.09)	(19.18)	(22.78)	(24.72)	(30.83)	(43.50)	(38.94)	(63.55)	(111.54)	(17.97)
f0712	312.58***	455.86***	476.94***	513.46***	515.36***	622.26***	490.77***	645.13***	686.69***	870.62***	239.28***
	(15.93)	(17.80)	(20.30)	(25.21)	(37.03)	(31.83)	(51.77)	(48.37)	(67.50)	(110.70)	(20.82)
m1318	316.85***	426.54***	504.96***	531.68***	529.98***	582.27***	626.23***	686.48***	937.39***	746.95***	257.68***
	(15.35)	(14.38)	(19.00)	(20.25)	(27.15)	(30.92)	(34.98)	(43.96)	(48.15)	(95.74)	(18.80)
f1318	329.79***	444.77***	480.21***	502.26***	521.52***	537.67***	589.76***	673.36***	601.67***	766.99***	231.55***
	(16.46)	(17.51)	(19.92)	(22.78)	(24.97)	(36.71)	(45.10)	(42.02)	(63.35)	(106.65)	(19.29)
m19+	345.26***	465.76***	487.28***	543.77***	581.68***	623.63***	703.06***	756.12***	827.67***	1145.20***	487.40***
	(10.87)	(11.57)	(11.98)	(16.30)	(20.66)	(24.97)	(29.21)	(26.97)	(42.22)	(56.35)	(14.91)
f19+	313.98***	431.13***	450.28***	520.60***	555.53***	574.15***	652.48***	748.74***	704.67***	919.88***	439.95***
	(12.66)	(14.91)	(14.69)	(17.73)	(21.38)	(23.33)	(22.79)	(22.73)	(34.49)	(58.21)	(17.07)
_cons	0.20**	-0.00	0.47***	0.48***	0.73***	0.95***	0.88***	0.92***	1.24***	0.84***	1.23***
	(0.09)	(0.11)	(0.10)	(0.10)	(0.11)	(0.12)	(0.09)	(0.07)	(0.10)	(0.10)	(0.03)
N	4500	3453	3358	3532	3576	3893	3958	4647	5114	3112	41267
adj. R^2	0.527	0.709	0.624	0.532	0.466	0.423	0.562	0.616	0.545	0.555	0.164
Hypothesis testing											
All β=0	273.40***	473.42***	399.53***	296.86***	192.66***	141.86***	234.44***	333.97***	175.62***	110.67***	270.30***
All β_male=β_female	1.62	2.73**	1.95*	0.98	1.64	2.08*	2.05*	0.22	4.86***	2.15*	2.07*
β_m0003=β_f0003	0.58	0.72	2.39	0.64	2.78*	0.67	5.00**	0.22	0.02	0.00	0.61
β_m0406=β_f0406	1.86	7.57***	2.65*	1.55	0.73	3.58*	1.19	0.95	1.12	0.01	0.46
β_m0712=β_f0712	2.47	0.08	0.08	0.71	4.08**	1.27	1.86	0.03	0.00	1.31	4.37**
β_m1318=β_f1318	0.32	0.73	0.82	1.02	0.06	1.05	0.41	0.04	16.37***	0.02	0.86
β_m19+=β_f19+	3.27*	3.87**	3.97**	1.04	0.97	3.41*	2.16	0.05	6.86***	9.08***	3.94**

[a] Hypothesis testing result corresponds to the β value in the upper section of the table.
[b] Households with MPCE(URP)>4500 has been dropped before the analysis.
[c] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[d] Standard errors are in parentheses.
[e] PCA has been used to estimate sigma square with no. of principal components between 4-10 and the explained variation is approximately 90%.
Source: Author's calculation based on 2004-05 Consumer Expenditure Household Survey, collected by NSS.

4 Within Household Gender Bias in Consumption: A Semiparametric Analysis of Engel Curves in Rural India

4.1 Introduction

Households differ from one another with respect to a lot of characteristics, for e.g. size, age composition, education level, expenditure patterns and so on. Expenditure on a commodity for a household depends on prices, income, tastes, and preferences as well as the household characteristics. The relationship between income and quantity consumed gives us the Engel curve for a commodity, *ceteris paribus*. Engel curve also helps to classify commodities into luxury, necessary and inferior. The classical studies of cross-section Engel curves are based on parametric models. But in the parametric form of Engel curve, the quality of the resulting estimator depends heavily on the correctness of the Engel curve specification. Working-Leser form of the Engel curves has long been discarded. Nowadays usually a log-quadratic model is used in the parametric set-up (Banks et al. (1997)). However, nonparametric Engel curves are more general and can capture more complex shapes. So, nonparametric smoothing procedure for estimating regression functions has received attention during the last two decades.

Ernst Engel (1857; 1895) was the first to investigate Engel curves and he found that food expenditures are an increasing function of income and of family size but the food budget shares decrease with income. After that Working (1943) proposed the log-linear budget share specification, which is also known as the Working-Leser model after Leser (1963) found that this functional form fits better than any other alternatives. However, there is some empirical evidence that suggests robust nonlinear relationships between some non-food items and the budget share (Banks et al. (1997) and Hausman et al. (1995)). Motivated by the nonlinearity and specification problem in the parametric form of Engel curve, one of the earlier empirical applications of nonparametric regression methods in econometrics was Kernel estimation of Engel curves[1].

The consumption of any commodity also depends on other characteristics of the household like demographics. Therefore, Engel function can be extended to include demographic variables and can be used to show the gender difference in consumption expenditure for various commodities. Some economists like Deaton (1989), Fuwa et al. (2006),Gibson (1997), Gibson and Rozelle (2004), Haddad and Reardon (1993), Himaz (2010), Lee (2008), Subramaniam (1996), Subramanian and Deaton (1991) etc. used the parametric Engel curve approach to find the gender bias in many aspects like education, food, alcohol, adult goods and children goods within the household. However, Deaton (1989) tried to find the gender bias in a different way. He applied the Engel curve approach for the adult good and named it as 'Adult Good Approach'. His idea is to look into the adult good which is not consumed by children. The budget share spent on such goods will reflect how much of their own consumption the parents are willing to give up for goods that are

[1] see also Delgado and Miles (1997), Gozalo (1997) and Lewbel (1991) for applications of parametric and nonparametric estimation methods of non-linear Engel curves.

shared with their children. So if the household prefers boys then there would be lower expenditure on adult goods as compared to girls.

A lot of attempts have been made in the direction of adult good approach to prove the existence of the gender bias. Strong evidence has been rare. Deaton (1997) (pp.240) notes:, "it is a puzzle that expenditure patterns so consistently fail to show strong gender effects even when measures of outcomes show differences between girls and boys". Case et al. (2002) (pp.11) says "it is not clear whether there really is no discrimination or whether, for some reason that is unclear, the method simply does not work". Failure of this method may be due to small budget shares of adult good. Other possible reason might be that there are a lot of households without any purchases of adult good.

As literature developed further, economists found that the linear function or the parametric form of the Engel curve cannot capture the whole behaviour between income and expenditure share of a commodity[2]. So, they opt for nonparametric forms of Engel curves. The nonparametric smoothing procedure for estimating regression functions received increasing attention in recent years as this method drops any assumptions regarding the functional form and thus any error due to the model specification. In case of the nonparametric form, there exists a dimensionality issue which restricts the Engel curve analysis for further expansion. In the nonparametric model, as the number of parameters increases the requirement for larger datasets also rises to cope up with the dimensionality issue. Most of the cross-sectional studies of gender bias cannot have such large datasets. So, as an easy path, economists go for a semiparametric form of Engel curve, where income has an unknown relation with the expenditure share of any commodity but have a linear relation with the demographic variables. Attfield and Bhalotra (1998), Blundell et al.

[2] Strauss and Thonas (1990) and Subramanian and Deaton (1996) plot nonparametric calorie expenditure curves for Brazil and India respectively.

(1998) and Gong et al. (2005) used this semiparametric form of Engel curve to find the gender bias in the household expenditure data.

Attfield and Bhalotra (1998) estimate the semiparametric Engel curves for rural Pakistan. They found little evidence of gender differences among children, but for the adults, significant bias in favour of the male is observed. They also estimate the quadratic logarithmic specification of the Engel curve for food, adult goods and child goods. On the other hand, Blundell et al. (1998) uses the British Family Expenditure Survey and considers for the endogeneity of total expenditure into the model and to overcome the endogeneity they adapted the Holly and Sargan (1982) augmented regression approach. On comparing semiparametric specification with the Piglog and quadratic logarithmic parametric specifications, they conclude that the Working-Leser or Piglog specification is strongly rejected for some budget shares but the quadratic logarithmic model seemed to provide an acceptable parametric specification. Gong et al. (2005) analyzed expenditure patterns for rural China, focusing on the differences between families with boys and girls. They estimate Engel curves for food and for alcohol, a typical adult good. They found no evidence of gender difference in case of the adult good but for food, significant evidence of gender bias is found only in the age group 13-15 years. The test for no gender difference and equal gender difference across age cannot be rejected for both the commodity.

In this chapter, our objective is to find gender bias in the consumption of food, intoxicant (an adult good) and milk (a child good) for various household sizes, using the semiparametric form of the Engel curve for rural India. The idea of Deaton (1989) is borrowed here to show the gender difference. Endogeneity of the total expenditure is also taken into account and has been tackled by two different approaches: Two Stage Least Square(2SLS) (discussed by Newey and Powell (2003)) and Control Function Approach(CFA) (introduced by Holly and Sargan (1982) as augmented regression approach). Last but not the least, the semiparametric specification is tested against constant, linear and quadratic Engel curves.

Our main findings are as follows. We find significant gender bias for the adults in case of both food and intoxicants for all the household sizes. For milk, evidence of gender bias in favour of the male is seen for the adults only for household sizes 4 and 6 and more.

The chapter is organised as follows: Section 4.2 explains the Deaton's Approach and in Section 4.3 , the semiparametric models for the Engel curves, the estimation and testing techniques are discussed. The nature of the data used is discussed in Section 4.4. Section 4.5 contains the results and major findings. Finally, the chapter ends with some concluding remarks in Section 4.6.

4.2 Deaton's Approach

Consider any age group K, then the total expenditure on say i^{th} good incurred by this group can be written as

$$TE_{b+g}^{(i)} = \beta_b^{(i)} B + \beta_g^{(i)} G \tag{4.1}$$

where, B and G is the number of boys and girls in the K^{th} age group respectively. $TE_{b+g}^{(i)}$ is the total expenditure on the i^{th} commodity incurred by both boy and girl in the K^{th} age group. Let us assume that $p^{(i)} = \beta_b^{(i)} - \beta_g^{(i)}$, then

$$TE_{b+g}^{(i)} = \beta_b^{(i)} B + [\beta_b^{(i)} - p^{(i)}]G = \beta_b^{(i)}(B+G) - p^{(i)}G = \beta_b^{(i)}(T) - p^{(i)}G \tag{4.2}$$

where $T = B + G$ i.e. total member in the K^{th} age group. So, if there is negative coefficient for girls for any age category then this implies positive '$p^{(i)}$' i.e. expenditure on boy is greater than girl for that age group[3].

4.3 Engel Curve Approach for detecting gender discrimination

The Working-Leser form of Engel Curve for any commodity is given by

$$Y_{ij} = \alpha_i + \gamma_i ln(pcte_j) + \sum_{k=1}^{K} \beta_{ki} x_{kj} + \epsilon_{ij} \tag{4.3}$$

where Y_{ij} is the budget share of the i^{th} good for the j^{th} household, $ln(pcte_j)$ is the log of per capita total expenditure for the j^{th} household and x_{kj} denotes the demographic variable for the j^{th} household. The linear relationship between the log of per capita total expenditure and the budget share of any commodity might not always be the true specification. As a way out, one can consider a nonparametric relation between them[4]. The ultimate functional form will then be a semiparametric form of the Engel curve which is

$$Y_{ij} = \alpha_i + F(ln(pcte_j)) + \sum_{k=1}^{K} \beta_{ki} x_{kj} + \epsilon_{ij} \tag{4.4}$$

where everything remains same, as in the previous equation i.e. eq.(4.3), except the linear relation between the log of per capita total expenditure and the budget share of any commodity, which changes to a nonparametric form. The curse of dimensionality is not

[3] Exogenous variables should be defined accordingly–For every age group, two types of variables will be defined one is related to the number of females and the other will be the total of members in the respective age group.

[4] For detail discussion see Atkinson et al. (1990), Attfield and Bhalotra (1998), Blundell et al. (1998), Hårdle and Jerison (1991), Hausman et al. (1991, 1995) and Lewbel (1991).

an issue here, as the nonparametric relation has only one variable i.e. ln(pcte).

Estimation of eq.(4.4) can be possible by the stepwise procedure of Robinson (1988). In this method we take conditional expectations in eq.(4.4) given $ln(pcte_j)$ and then subtracting this conditional equation from eq.(4.4) on both sides of the equation. Thereby giving us

$$Y_{ij} - E(Y_{ij}|ln(pcte_j)) = \sum_{k=1}^{K} \beta_{ki}[x_{kj} - E(x_{kj}|ln(pcte_j))] + \epsilon_{ij} \tag{4.5}$$

The conditional means $E(Y_{ij}|ln(pcte_j))$ and $E(x_{kj}|ln(pcte_j))$ can be estimated nonparametrically using univariate kernel regression in the first step. In the second step, β is estimated by OLS from eq (4.5), after replacing the estimates of the conditional means from the first step. Robinson (1988) showed that the estimate of β is $\sqrt{n}$ - consistent and asymptotically normal [5].

4.3.1 Endogeneity of Total Expenditure

Till now it was assumed that the error term in eq.(4.4) has conditional mean zero given log per capita total expenditures $ln(pcte_j)$ and other variables $x_{.j}$. However, this rarely happens, it is not always equal to zero and hence correlation exists, (see Blundell et al. (1998)). If $E(\epsilon_{ij}|ln(pcte_j)) \neq 0$, then the above estimators will not be consistent. Here, we follow the following two procedures to solve the endogeneity problem.

[5] Yatchew (1997) found another $\sqrt{n}$ - consistent estimator, which give similar estimates but has larger standard errors than Robinson (1988) estimator.

4.3.2 Two stage least squares (2SLS)

In 2SLS approach, the endogenous variable, i.e. log of per capita total expenditure ($ln(pcte_j)$), is a linear function of the demographic variables and other instrumental variables.

$$ln(pcte_j) = \pi' z_j + v_j \tag{4.6}$$

where z consists of x, the demographic variables and at least one other variable which is not in x. Here, parameters of eq.(4.6) are estimated by OLS, which are then used to estimate the value of $ln(pcte_j)$ and are then substituted in eq.(4.4), i.e. the semiparametric form of Engel curve. Thereby solving the endogeneity problem. After that eq.(4.4) can be estimated by the usual Robinson (1988)'s two step procedure.

4.3.3 Control Function Approach (CFA)

Control function approach is same as the 'Augmented Regression Technique' introduced by Holly and Sargan (1982). Blundell et al. (1998) and Gong et al. (2005) used this technique in their article. The procedure takes into account eq.(4.6) with the assumptions

$$E(v|z, x) = 0 \tag{4.7}$$

and

$$E(\epsilon|ln(pcte), z, x) = \rho v \tag{4.8}$$

Then the new transformed model will be

$$Y_{ij} = \alpha_i + F(ln(pcte_j)) + \sum_{k=1}^{K} \beta_{ki} x_{kj} + \rho v_j + \tilde{\epsilon_{ij}} \tag{4.9}$$

with

$$E(\tilde{\epsilon}|ln(pcte), z, x) = 0 \tag{4.10}$$

Using OLS for eq.(4.6) one can estimate $\hat{v}$(the residual), i.e. the consistent estimate of v. The Robinson (1988) technique can then be applied to eq.(4.9), to get the estimates of the parameters.

Now, it is possible to fit the non-linear relation between $ln(pcte)$ and Y by estimating equation (4.11) presented below nonparametrically.

$$Y_{ij} - \sum_{k=1}^{K} \hat{\beta}_{ki} x_{kj} = \alpha_i + F(ln(pcte_j)) + \epsilon_{ij},\ j = 1, \ldots, H \tag{4.11}$$

4.3.4 Hardle and Mammen's (1993) test

Sometimes nonparametric functions may be approximated by some parametric polynomial. To test for the appropriateness of such an approximation, Hardle and Mammen (1993) developed a statistic which compares the nonparametric and parametric regression fits using squared deviations between them. The test-statistic is:

$$T_j = H\sqrt{h} \sum_{j=1}^{H} [\hat{F}(ln(pcte_j)) - \hat{F}(ln(pcte_j), \alpha, \beta)]^2 \pi(.) \tag{4.12}$$

where $\hat{F}(ln(pcte_j))$ is the nonparametric function estimate obtained from eq.(4.11), $\hat{F}(ln(pcte_j), \alpha, \beta)$ is an estimated parametric function, h is the bandwidth used and $\pi(.)$ is a weighting function for the squared deviation between fits. The absence of the rejection of the null (i.e. "accepting" the parametric model) means that polynomial adjustment is at least of the degree that has been tested.

4.4 The Data

The 68^{th} round of the National Sample Survey data[6] is used for the current study. This dataset is produced by the National Sample Survey Office, NSSO –Ministry of Statistics & Programme Implementation, Government of India. It was collected during July 2011–June 2012 and has information on the Indian Household Consumer Expenditure. It is a cross–sectional data and the unit of observation is household. The survey provides the information on the quantity consumed and the expenditure incurred by the households on various consumer goods and services during the reference period. It also contains information on individual's age, number of meals taken at different places during last 30 days, education level etc. The information of other things like the lighting, cooking condition, agriculture land owned and cultivated are also available at the household level. The survey covers the whole of the Indian Union except a few places which remain inaccessible throughout the year. The survey uses a stratified multi-stage sampling design. And the sampling weights are provided as multiplier which can be used to convert data or estimates into population approximation.

The data provide the information of 1,01,662 households from the entire country. Out of which 59,695 number of household belongs to the rural sector and the remaining are the urban household. The total population according to the survey is 1.10 billion. And the rural sector consists of 792 million people which are approximately 71.43% (as a percentage of total population). The survey shows that rural India had an estimated 96 million agricultural households–about 55.76% of the total estimated rural households in the country. In the rural sector economic problems such as poverty, gender bias remains

[6]Type 1 data of 68^{th} round is used. For a brief discussion on the distinction between Type 1 and 2 datasets of NSSO one can refer to Chapter-2.

more acute and intense. Therefore, the current study will focus mainly on rural India. Our objective is to focus on the rural India gender difference in consumption. Household size is one of the factors which might change the spending behaviour of the households. To incorporate such behavioural change households are classified in terms of 4 types based on their sizes–households with size less than or equal to 3, size equal to 4, size equal to 5 and size equal to or greater than 6. The percentage distribution is 30.30%, 22.14%, 19.22% and 28.34% for the respective household sizes in the total rural household population.

Table (4.1), in the Appendix, shows that the per capita value (in terms of rupees) of all the three commodities and the monthly per capita total expenditure for all the four types of households as well as the whole households taken together. All the four variables decrease as the household size increases. The three broad categories of commodities are a) total monthly food expenditure, b) total monthly expenditure on milk (which include milk and other milk products) and c) total monthly expenditure on intoxicant (which consists of the pan, tobacco and other intoxicants). Milk represents as a child good whereas intoxicant represents an 'adult good'. The share of these commodities is taken with respect to monthly total expenditure (under mixed reference period[7].) The demographic variables consists of the number of female and total members (as the share of family size) in each of the 5 age groups: 0–3, 4–6, 7–12, 13–18 and 19 and above. We need to drop the last group (i.e. total members in 19 and above age group) while running the semiparametric model to avoid the multicollinearity issue.

The summary statistics of the variables are presented in separate tables starting from Table (4.4)–(4.7) in the Appendix. Each table represents a particular type of household

[7]MPCE(MRP) is a method of collecting data. Under this method, data on clothing, bedding, footwear, education, institutional medical care, and durable goods are collected over a one-year period, while sticking to the 30-day recall for the rest of the items.

based on its sizes. The average household size, for those households with size less than equal to 3, is 2.30, whereas it is 7.26 for those with size greater than equal to 6. It is kind of obvious that the data has observation where household size is less than 3 as well as greater than 6 in our dataset. The mean value of log per capita total expenditure decreases as household size increases which were also evident from the table (4.1). The mean of the share of food expenditure increases with the household size from 52% to 55%. The economic implication is that as per capita food expenditure and monthly per capita consumer expenditure decreases with the increase in household size, the fall in monthly per capita consumer expenditure is more than the fall in per capita food expenditure, thereby, making the share of food increase with the household size. A similar result is also obtained for the milkshake. In these tables, the share of the female for 19 and above age category also declines as the household size increases supporting the case of "Missing Women"–Gupta (2005). The average number of households having electricity ranges from 67–77% between different household sizes. As household size increases the mean value of having the personal vehicle (D2) also increases ranging from 12%–25%. The per capita monthly expenditure might reduce but the overall income of the household will increase this may happen if the number of earning members increases with the household size. Therefore, the bigger household can afford a personal vehicle. Last but not the least, approximately 4% of the rural population possess PC, laptop, and other peripherals.

As discussed in Section 4.3.1, there can be endogeneity of total expenditure in the semi-parametric model of the Engel function, which means changes in $ln(pcte)$ are associated not only with changes in Y but also with the error term (ϵ). What is needed here is a method to generate only exogenous variation in $ln(pcte)$. To cope up with the endogeneity problem we need to have some instrumental variables or instruments. The instrumental variables (z_j) are those variables which are (1) exogenous i.e. $Cov(z, \epsilon) = 0$. More specifically z should have no "partial" effect on Y and should be uncorrelated with

ϵ (see eq.(4.4)). (2) correlated with the endogenous variable ($ln(pcte)$). We cannot test $Cov(z, \epsilon) = 0$ as this is a population assumption. But we can check the second condition, i.e. the value of the correlation coefficient and rank correlation of the instrumental variables, see Table (4.2). In our model five instruments[8] are chosen: number of meals served to non-household members during last 30 days, total expenditure on jewellery and ornaments during last 365 days, dummy variable for the possessing of electricity (D1), dummy variable for the possession of motorcycle, scooter, jeep and other transport equipment (D2) and dummy variable for the possession of PC, Laptop, other peripherals incl. software or telephone instrument (D3). The correlation coefficient is highest for D2 i.e. above 0.40 for all the categories of household sizes. The second variable is D3 in terms of ranking of correlation coefficient and then D1. The rest of the variable is more or less same correlated with $ln(pcte)$.

The regression of $ln(pcte)$ on the instruments and other exogenous variables are shown in Table (4.3). We can see that all the coefficients of the instruments are significant at 1% level for all the households. The R^2 value is greater than 0.32. The F-statistics for the test of overall fit is also statistically significant at 1%. Therefore, we can conclude that we do not have any weak instruments.

4.5 Results and Empirical Findings

The estimated result of the semiparametric model for food, milk and intoxicants are presented in Tables (4.8), (4.9) and (4.10) respectively in the Appendix. The default kernel regression used for all stages is a Gaussian kernel-weighted local polynomial fit[9]. The optimal bandwidth is used which minimizes the conditional weighted mean integrated squared

[8] The problem with Instrumental Variable estimation method is that it is often very difficult to get appropriate instruments. This may be one of the reasons for taking only a few instruments.

[9] The kernel is of order 2.

error (see Verardi and Debarsy (2012)). For each type of household sizes, three models are presented –Model without endogeneity of ln(pcte), 2SLS and the Control function approach. The first section of the table shows the coefficient of the exogenous variables of the semiparametric model[10]. The coefficients for food and intoxicants shows that there is significant gender bias in favour of the male for the adult. This result is similar to what Attfield and Bhalotra (1998) has obtained in their empirical investigation on rural Pakistan. However, for milk two types of households i.e. with size 4 and 6 and above shows significant gender bias in favour of the male for adults. But for household size less than equal to 3, the gender bias in milk consumption favours female adults. Similarly, for age group 7-12, we found significant and positive gender bias in favour of the female for the food consumption in case of household size 4 and 5. Other significant gender biases in favour of the male are seen in case of milk consumption for household size 5 for all the age groups except the adults. For household size equal to 4, significant gender bias in favour of the male is seen for all the age groups except the children 0-3 years in case of the milk. Here, all the three methods, i.e. without considering for the endogeneity of ln(pcte), 2SLS and CFA, all give approximately similar results.

In the third section of these tables, we see the results of the hypotheses for equal sex difference across age and test for no sex difference. Both of these tests are rejected for food. For milk, only the equal sex difference across age is not rejected for only the household size equal to 5. In case of intoxicants, these tests are significant for all the sizes of household except the last one i.e. household size greater than equal to 6. In the first section of these tables, the coefficient of ρ suggest whether there is endogeneity of ln(pcte) exists or not? If ρ is significant then there is endogeneity within the model. For food and intoxicants, we found significant endogeneity of ln(pcte) for all the category of households.

[10]Note only the female related variable are shown in the tables and following the Deaton (1989) approach–a negative sign means that a family with a female member of any age group will consume less of the commodity than a comparable male member in the same age group.

But for milk, in Table (4.9) the coefficient of ρ is only significant for household size 4 and 5.

Hardle and Mammen (1993) statistic to test whether the nonparametric fit could be approximated by a polynomial fit is shown in the second section of these tables. For the sake of clarity, we rescaled the statistic in such a way that it can be compared with the quantile of a Normal distribution. Note, however, that the test is not normally distributed. The number of bootstrap replicates used to get inference is set to 100. Hardle and Mammen (1993) specification test is done against constant, linear and quadratic specifications. In case of food, the constant model is rejected for all the household sizes. The Linear model is accepted only by the 2SLS approach for households size 4, 5 and 6 and above. Lastly, the quadratic model is accepted for the household size equal to 4 (for all the approach), and for households size 5 and 6 except for the case of 2SLS. In case of milk, Hardle and Mammen (1993) specification test is rejected for all. The only exception is for the household size 4 under CFA approach where the quadratic model is accepted. Result for intoxicant, the only adult good, is presented in Table 4.10. Constant model is accepted only for household size 4 for without endogeneity case. Linearity is accepted for household size 4 (for all the cases) and 5 (only for CFA). The quadratic form is accepted for household size 3 (under without endogeneity case), 4 (under all the cases) and 5 (except for 2SLS).

In Chapter-3, we dealt with only the food expenditure and found significant difference in consumption which favours male for the rural sector in 0-3 and the adult age categories in the 61^{st} round and for the 66^{th} round also the gender difference in food consumption favours the male for 0-3,13-18 and 19 above age groups. Whereas in this Chapter, significant gender bias in favour of male is observed only in the adult age group for food expenditure. For age group 7-12 with household sizes 4 and 5, significant gender bias is found in favour of females- a result similar to Himaz (2010) and Fuwa (2014).

4.6 Conclusion

We analysed expenditure pattern in rural India using NSS 68^{th} round consumer expenditure data collected during July 2011–June 2012. We estimated the semiparametric form of Engel curve considering with and without endogeneity of log per capita total expenditure. To solve the endogeneity issue two methods have been used; 2SLS and CFA. We find evidence of gender differentials in food and intoxicants for the adults, confirming the results of Attfield and Bhalotra (1998). For milk, we found evidence of gender bias for approximately all the age groups in household size 4 and 5. We also found evidence of gender bias in favour of the female in case of milk and food for household size 3, 4 and 5 respectively. To compare semiparametric specification with constant, linear and quadratic parametric specifications we implement the recently developed specification test by Hardle and Mammen (1993) test generated by 100 replications of the wild bootstrap. This test is accepted for food as well as for intoxicant for few cases, which is a contradictory result obtained from Gong et al. (2005). So linear and quadratic Engel curves are suitable for few commodities and households size. The overall test for equal sex difference across age and no gender bias is rejected which imply the existence of gender bias for each household sizes separately.

4.7 APPENDIX

Table 4.1: Average monthly per capita consumption for various household sizes (in Rupees)

Variables	All household	household size$\leqslant$3	household size=5	household size=6	household size$\geqslant$6
Per capita food consumption	652.07	831.11	711.79	641.22	570.87
	(330.39)	(466.53)	(335.11)	(274.44)	(260.55)
Per capita milk consumption	116.13	130.27	119.68	113.64	110.97
	(145.02)	(170.63)	(153.55)	(144.18)	(131.21)
Per capita intoxicant consumption	30.10	45.67	35.73	29.14	22.86
	(62.69)	(92.07)	(87.90)	(48.61)	(36.64)
Monthly per capita total expenditure	1287.17	1737.70	1436.23	1251.70	1087.21
	(962.87)	(1578.38)	(964.78)	(783.18)	(651.27)

[1] Standard deviation are in parentheses.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.2: Correlation coefficient and Rank correlation of log per capita total expenditure and other instrumental variables

		Correlation Coefficient					Rank Correlation				
Variables	Description	All household	hhsize⩽3	hhsize=4	hhsize=5	hhsize⩾6	All household	hhsize⩽	hhsize=4	hhsize=5	hhsize⩾6
mealnonhh	no. of meals served to non-household members during last 30 days	0.23	0.29	0.21	0.24	0.25	0.24	0.31	0.27	0.25	0.26
jewellery	total expenditure on jewellery and ornaments during last 365 days	0.23	0.22	0.26	0.23	0.25	0.07	0.09	0.11	0.12	0.11
D1	dummy, 1 if household have electricity for lighting	0.31	0.32	0.32	0.32	0.33	0.33	0.33	0.33	0.34	0.35
D2	dummy, 1 if household have motor cycle, scooter, jeep and other transport equipment.	0.40	0.43	0.51	0.48	0.47	0.41	0.43	0.51	0.49	0.48
D3	dummy, 1 if household have PC or laptop or other peripherals incl. software or telephone instrument.	0.34	0.37	0.37	0.35	0.34	0.29	0.32	0.31	0.30	0.30

[1] Spearman's rank correlation coefficients is reported here.

Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.3: OLS result of the instruments and other exogenous variables against the endogenous variable (lnpcmrp)

	(1)	(2)	(3)	(4)
lnpcmrp	hhsize⩽3	hhsize=4	hhsize=4	hhsize⩾6
mealnonhh	0.00565***	0.00237***	0.00516***	0.00370***
	(0.000440)	(0.000744)	(0.000906)	(0.000315)
jewellery	0.00000384***	0.00000683***	0.00000435***	0.00000724***
	(0.000000446)	(0.000000997)	(0.000000664)	(0.000000659)
D1	0.289***	0.239***	0.181***	0.196***
	(0.0152)	(0.0139)	(0.0139)	(0.0112)
D2	0.407***	0.456***	0.389***	0.348***
	(0.0183)	(0.0146)	(0.0143)	(0.0115)
D3	0.465***	0.364***	0.348***	0.331***
	(0.0398)	(0.0280)	(0.0304)	(0.0196)
shx2	-0.113	0.112	0.0236	0.0297
	(0.0964)	(0.0968)	(0.110)	(0.0922)
shx11	-0.697***	-0.461***	-0.542***	-0.572***
	(0.0815)	(0.0661)	(0.0826)	(0.0791)
shx4	-0.0362	0.144	-0.0390	0.00945
	(0.154)	(0.0921)	(0.108)	(0.0983)
shx12	-0.640***	-0.385***	-0.191**	-0.372***
	(0.113)	(0.0674)	(0.0855)	(0.0847)
shx6	-0.105	-0.0885	0.0218	-0.0526
	(0.0986)	(0.0659)	(0.0636)	(0.0620)
shx13	-0.372***	-0.0833	-0.202***	-0.236***
	(0.0779)	(0.0517)	(0.0542)	(0.0539)
shx8	-0.210	-0.0752	-0.00621	-0.0121
	(0.129)	(0.0658)	(0.0722)	(0.0675)
shx14	-0.125	0.0306	-0.0656	-0.0850
	(0.105)	(0.0539)	(0.0590)	(0.0606)
shx10	-0.342***	-0.0523	0.100	0.0630
	(0.0466)	(0.0703)	(0.0700)	(0.0659)
_cons	7.223***	6.913***	6.818***	6.738***
	(0.0289)	(0.0350)	(0.0358)	(0.0349)
N	16157	13416	11890	18231
adj. R^2	0.324	0.393	0.379	0.391
F	196.4***	207.0***	147.9***	264.3***

[1] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.

[2] Standard errors are in parentheses.

Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.4: Summary Statistics and definition of the variables used when the household size⩽3

Variable	Obs	Mean	Std. Dev.	Description
hhsize	16157	2.30	0.75	household size
lnpcte	16157	7.29	0.53	log per capita total expenditure (mixed reference period)
shf	16157	0.52	0.12	share of food expenditure
intoxsh	16157	0.03	0.05	share of intoxicant
milksh	16157	0.14	0.13	share of milk
shf0003	16157	0.01	0.07	share of the no. of female with 0-3 age
shf0406	16157	0.01	0.05	share of the no. of female with 4-6 age
shf0712	16157	0.02	0.08	share of the no. of female with 7-12 age
shf1318	16157	0.03	0.11	share of the no. of female with 13-18 age
shf19+	16157	0.44	0.25	share of the no. of female with age 19 years and above
sht0003	16157	0.03	0.10	share of the total no. of members with 0-3 age
sht0406	16157	0.02	0.08	share of the total no. of members with 4-6 age
sht0712	16157	0.05	0.15	share of the total no. of members with 7-12 age
sht1318	16157	0.07	0.18	share of the total no. of members with 13-18 age
mealnonhh	16157	5.57	17.80	no. of meals served to non-household members during last 30 days
jewellery	16157	945.89	18494.60	total expenditure on jewellery and ornaments during last 365 days
D1	16157	0.74	0.44	dummy, 1 if household have electricity for lighting
D2	16157	0.12	0.32	dummy, 1 if household have motor cycle,scooter, jeep and other transport equipment.
D3	16157	0.04	0.20	dummy, 1 if household have PC or laptop or other peripherals incl. software or telephone instrument.

[1] Values are rounded to two decimal places.
[2] Household weights or household multipliers are used for the above summary statistics. Only for lnpcte individual weights are used.
[3] The share of food, intoxicant, and milk are with respect to total expenditure whereas the share of female and total members in various age groups are with respect to household size.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.5: Summary Statistics and definition of the variables used when the household size=4

Variable	Obs	Mean	Std. Dev.	Description
hhsize	13416	4.00	0.00	household size
lnpcte	13416	7.15	0.47	log per capita total expenditure (mixed reference period)
shf	13416	0.53	0.11	share of food expenditure
intoxsh	13416	0.03	0.04	share of intoxicant
milksh	13416	0.15	0.13	share of milk
shf0003	13416	0.03	0.09	share of the no. of female with 0-3 age
shf0406	13416	0.03	0.08	share of the no. of female with 4-6 age
shf0712	13416	0.05	0.11	share of the no. of female with 7-12 age
shf1318	13416	0.05	0.11	share of the no. of female with 13-18 age
shf19+	13416	0.30	0.11	share of the no. of female with age 19 years and above
sht0003	13416	0.07	0.13	share of the total no. of members with 0-3 age
sht0406	13416	0.06	0.12	share of the total no. of members with 4-6 age
sht0712	13416	0.13	0.18	share of the total no. of members with 7-12 age
sht1318	13416	0.13	0.18	share of the total no. of members with 13-18 age
mealnonhh	13416	5.21	21.84	no. of meals served to non-household members during last 30 days
jewellery	13416	992.83	10272.53	total expenditure on jewellery and ornaments during last 365 days
D1	13416	0.77	0.42	dummy, 1 if household have electricity for lighting
D2	13416	0.20	0.40	dummy, 1 if household have motor cycle,scooter, jeep and other transport equipment.
D3	13416	0.04	0.21	dummy, 1 if household have PC or laptop or other peripherals incl. software or telephone instrument.

[1] Values are rounded to two decimal places.
[2] Household weights or household multipliers are used for the above summary statistics. Only for lnpcte individual weights are used.
[3] The share of food, intoxicant, and milk are with respect to total expenditure whereas the share of female and total members in various age groups are with respect to household size.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.6: Summary Statistics and definition of the variables used when the household size=5

Variable	Obs	Mean	Std. Dev.	Description
hhsize	11890	5.00	0.00	household size
lnpcte	11890	7.02	0.45	log per capita total expenditure (mixed reference period)
shf	11890	0.54	0.10	share of food expenditure
intoxsh	11890	0.02	0.03	share of intoxicant
milksh	11890	0.15	0.14	share of milk
shf0003	11890	0.03	0.08	share of the no. of female with 0-3 age
shf0406	11890	0.03	0.08	share of the no. of female with 4-6 age
shf0712	11890	0.07	0.11	share of the no. of female with 7-12 age
shf1318	11890	0.07	0.12	share of the no. of female with 13-18 age
shf19+	11890	0.28	0.12	share of the no. of female with age 19 years and above
sht0003	11890	0.06	0.11	share of the total no. of members with 0-3 age
sht0406	11890	0.06	0.11	share of the total no. of members with 4-6 age
sht0712	11890	0.16	0.18	share of the total no. of members with 7-12 age
sht1318	11890	0.15	0.18	share of the total no. of members with 13-18 age
mealnonhh	11890	5.11	14.12	no. of meals served to non-household members during last 30 days
jewellery	11890	969.99	12884.74	total expenditure on jewellery and ornaments during last 365 days
D1	11890	0.72	0.45	dummy, 1 if household have electricity for lighting
D2	11890	0.21	0.40	dummy, 1 if household have motor cycle,scooter, jeep and other transport equipment.
D3	11890	0.05	0.21	dummy, 1 if household have PC or laptop or other peripherals incl. software or telephone instrument.

[1] Values are rounded to two decimal places.
[2] Household weights or household multipliers are used for the above summary statistics. Only for lnpcte individual weights are used.
[3] The share of food, intoxicant, and milk are with respect to total expenditure whereas the share of female and total members in various age groups are with respect to household size.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.7: Summary Statistics and definition of the variables used when the household size$\geqslant$6

Variable	Obs	Mean	Std. Dev.	Description
hhsize	18231	7.26	1.84	household size
lnpcte	18231	6.88	0.44	log per capita total expenditure (mixed reference period)
shf	18231	0.55	0.11	share of food expenditure
intoxsh	18231	0.02	0.03	share of intoxicant
milksh	18231	0.17	0.14	share of milk
shf0003	18231	0.04	0.07	share of the no. of female with 0-3 age
shf0406	18231	0.04	0.07	share of the no. of female with 4-6 age
shf0712	18231	0.08	0.11	share of the no. of female with 7-12 age
shf1318	18231	0.07	0.10	share of the no. of female with 13-18 age
shf19+	18231	0.27	0.11	share of the no. of female with age 19 years and above
sht0003	18231	0.08	0.10	share of the total no. of members with 0-3 age
sht0406	18231	0.08	0.10	share of the total no. of members with 4-6 age
sht0712	18231	0.16	0.15	share of the total no. of members with 7-12 age
sht1318	18231	0.14	0.15	share of the total no. of members with 13-18 age
mealnonhh	18231	5.90	21.26	no. of meals served to non-household members during last 30 days
jewellery	18231	970.60	8165.01	total expenditure on jewellery and ornaments during last 365 days
D1	18231	0.67	0.47	dummy, 1 if household have electricity for lighting
D2	18231	0.25	0.43	dummy, 1 if household have motor cycle,scooter, jeep and other transport equipment.
D3	18231	0.04	0.20	dummy, 1 if household have PC or laptop or other peripherals incl. software or telephone instrument.

[1] Values are rounded to two decimal places.
[2] Household weights or household multipliers are used for the above summary statistics. Only for lnpcte individual weights are used.
[3] The share of food, intoxicant, and milk are with respect to total expenditure whereas the share of female and total members in various age groups are with respect to household size.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.8: Estimates of the Semi-parametric Model for Food under various household sizes

Variables	hhsize⩽3			hhsize=4			hhsize=5			hhsize⩾6		
	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA
shf0003	-0.0228	-0.0256	-0.0252	-0.0123	-0.0110	-0.0124	0.0113	0.0111	0.0145	0.0117	0.0130	0.0114
shf0406	-0.0053	-0.0117	-0.0060	-0.0004	0.0084	0.0083	-0.0034	-0.0080	-0.0074	-0.0017	0.0043	0.0022
shf0712	0.0148	0.0054	0.0146	0.0203**	0.0168*	0.0185**	0.0208**	0.0254**	0.0245**	0.0109	0.0095	0.0092
shf1318	-0.0133	-0.0174	-0.0187*	-0.0103	-0.0260**	-0.0142	-0.0114	-0.0070	-0.0070	-0.0018	0.0066	0.0040
shf19+	-0.0514***	-0.0605***	-0.0608***	-0.0416***	-0.0499***	-0.0395***	-0.0310***	-0.0196*	-0.0175*	-0.0517***	-0.0361***	-0.0376***
ρ	-	-	0.0499***	-	-	0.0655***	-	-	0.0690***	-	-	0.0672***
no.of hh	16157	16156	16157	13416	13415	13416	11890	11889	11890	18231	18231	18231
Specific tests:												
Constant	58.035***	45.781***	80.096***	53.470***	47.637***	68.574***	67.336***	42.978***	65.141***	51.501***	55.846***	66.588***
Linear	20.173***	3.374***	10.602***	4.308***	1.267	2.745***	8.900***	1.544	3.510***	3.586***	1.634	2.447***
Quadratic	3.868***	4.447***	3.336***	1.590	0.791	1.406	1.747	2.759***	1.471	1.618	3.907***	1.486
Equal sex diff. across age	9.09***	8.63***	12.02***	6.23***	7.03***	6.06***	3.94***	2.63**	2.89**	7.32***	3.88***	4.48***
No sex diff.	37.18***	40.17***	51.57***	6.19***	7.70***	6.01***	3.38***	2.11*	2.31**	6.61***	3.17***	3.75***

[1] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[2] In the second section of the table, specific tests are conducted against semiparametric model using Hardle and Mammen's (1993) test.
[3] The third section of the table shows the result of the F-test with 4 and 5 degrees of freedom respectively.
[4] If ρ is statistically significant then this implies that exogeneity of log per capita total expenditure is rejected.
[5] For the sake of comparison, only the female related variables are reported in the table. However, both the female and the total members in each age-group are taken during the estimation.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.9: Estimates of the Semi-parametric Model for Milk under various household sizes

Variables	hhsize⩽3			hhsize=4			hhsize=5			hhsize⩾6		
	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA
shf0003	-0.0095	-0.0079	-0.0093	-0.0152	-0.0127	-0.0152	-0.0219*	-0.0230*	-0.0213*	0.0012	0.0030	0.0011
shf0406	-0.0123	-0.0166	-0.0122	-0.0184*	-0.0198*	-0.0164*	-0.0226*	-0.0212*	-0.0234**	-0.0209*	-0.0239**	-0.0208*
shf0712	-0.0094	-0.0126	-0.0093	-0.0361***	-0.0356***	-0.0365***	-0.0153**	-0.0148*	-0.0146*	-0.0126*	-0.0103	-0.0127*
shf1318	-0.0092	-0.0053	-0.0089	-0.0369***	-0.0401***	-0.0379***	-0.0274***	-0.0280***	-0.0265***	0.0044	0.0041	0.0046
shf19+	0.0200***	0.0230***	0.0206***	-0.0145**	-0.0182***	-0.0140**	-0.0050	-0.0072	-0.0024	-0.0234***	-0.0242***	-0.0228***
ρ	-	-	-0.0032	-	-	0.0154***	-	-	0.0133***	-	-	0.003
no.of hh	16157	16156	16157	13416	13415	13416	11890	11889	11890	18231	18231	18231
Specific tests:												
Constant	105.571***	39.960***	96.470***	74.312***	32.068***	45.025***	84.358***	30.433***	51.148***	106.129***	61.848***	112.984***
Linear	44.425***	10.689***	39.851***	19.248***	9.802***	21.112***	23.893***	7.030***	20.311***	18.297***	6.747***	18.706***
Quadratic	6.022***	13.766***	7.180***	2.189***	10.733***	1.655	2.321***	5.208***	2.126***	3.193***	4.751***	3.161***
Equal sex diff. across age	8.39***	9.16***	8.61***	2.41**	2.22*	2.73**	1.35	1.11	1.61	2.30*	2.45**	2.26*
No sex diff.	13.95***	15.34***	14.42***	12.49***	12.66***	12.71***	4.53***	4.32***	4.31***	3.25***	3.23***	3.15***

[1] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[2] In the second section of the table, specific tests are conducted against semiparametric model using Hardle and Mammen's (1993) test.
[3] The third section of the table shows the result of the F-test with 4 and 5 degrees of freedom respectively.
[4] If ρ is statistically significant then this implies that exogeneity of log per capita total expenditure is rejected.
[5] For the sake of comparison, only the female related variables are reported in the table. However, both the female and the total members in each age-group are taken during the estimation.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

Table 4.10: Estimates of the Semi-parametric Model for Intoxicants under various household sizes

Variables	hhsize⩽3			hhsize=4			hhsize=5			hhsize⩾6		
	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA	Model 1 without endo	Model 2 2SLS	Model 3 CFA
shf0003	-0.0099	-0.0108	-0.0108	0.0023	0.0021	0.0022	0.0024	0.0032	0.0031	0.0008	0.0009	0.0008
shf0406	-0.0133	-0.0135	-0.0136	-0.0065	-0.0053	-0.0039	-0.0030	-0.0039	-0.0039	0.0016	0.0030	0.0023
shf0712	0.0066	0.0063	0.0065	0.0009	0.0006	0.0004	-0.0038	-0.0038	-0.0030	-0.0019	-0.0024	-0.0022
shf1318	-0.0063	-0.0077*	-0.0082*	0.0018	0.0008	0.0007	0.0000	0.0005	0.0009	-0.0027	-0.0019	-0.0016
shf19+	-0.0423***	-0.0414***	-0.0456***	-0.0236***	-0.0231***	-0.0230***	-0.0194***	-0.0166***	-0.0164***	-0.0083***	-0.0055*	-0.0055*
ρ	-	-	0.0176***	-	-	0.0189***	-	-	0.0156***	-	-	0.0135***
no.of hh	16157	16156	16157	13416	13415	13416	11890	11889	11890	18231	18231	18231
Specific tests:												
Constant	23.866***	59.473***	67.468***	1.291	2.341***	4.697***	7.393***	23.941***	32.747***	4.314***	33.755***	34.858***
Linear	3.710***	6.817***	1.935**	0.853	0.790	0.782	3.226***	4.878***	1.268	2.718***	2.778***	2.664***
Quadratic	1.459	12.419***	1.823**	0.844	0.793	0.817	1.219	5.447***	1.125	2.528***	3.729***	2.650***
Equal sex diff. across age	32.19***	29.13***	36.49***	9.20***	8.64***	8.46***	4.93***	3.90***	3.86***	1.25	0.76	0.68
No sex diff.	132.19***	114.61***	151.70***	10.01***	9.55***	9.38***	6.01***	4.67***	4.40***	1.66	0.91	0.83

[1] ***,**,* indicate significance at the level 1%, 5% and 10%, respectively.
[2] In the second section of the table, specific tests are conducted against semiparametric model using Hardle and Mammen's (1993) test.
[3] The third section of the table shows the result of the F-test with 4 and 5 degrees of freedom respectively.
[4] If ρ is statistically significant then this implies that exogeneity of log per capita total expenditure is rejected.
[5] For the sake of comparison, only the female related variables are reported in the table. However, both the female and the total members in each age-group are taken during the estimation.
Source: Author's calculation based on 2011-2012 Type 1 Consumer Expenditure Household Survey, collected by NSS.

5 Estimation of Consumer Unit scale and Economies of Scale

5.1 Introduction

The measurement of welfare or poverty requires conversion from a household to an individual level, for example, total expenditure converted into per capita level by dividing it by the household size. Per capita does not consider the difference in the resource allocation within the household and thus it fails to consider that everyone is not the same and has different requirements. For this reason, a need arises to find some sort of equivalence scale where an adult male member of a household is taken as a unit and other members are given weights as a fraction of the consumption/requirements of the adult male members. This is known as adult equivalence scale. The New Palgrave Dictionary of Economics defines 'An equivalence scale as a measure of the cost of living of a household for a given size and demographic composition, relative to the cost of living of a reference household (usually a single adult), when both households attain the same level of utility or standard of living' (see Lewbel and Pendakur, 2006, pp.1 for more details). Equivalence scales are difficult to construct because household's/individual's utility is not directly observable. Howsoever, equivalence scales for each individual is assigned a value in proportion to its needs. The factor commonly taken into account to assign values is the age of the members (whether they are adults or children). Several other economic and non-economic factors also influence the consumption pattern of members of a household such as occupation, marital status, income, level of education, tastes and preferences, psychological factors etc. All these factors may not be quantifiable and sometimes these variables may not be equally important and so it does not directly influence the consumption patterns. Therefore, researchers have investigated the effect of those variables which have the above qualities and importance. A wide range of equivalence scales exist, many of which are reviewed in Haque (2006). Not only age but gender also plays an important role in the consumer demand analysis. By considering these factors we can estimate, for example, the cost of maintaining a newly born male child as a fraction of the cost of maintaining an adult male.

The concept of consumer unit scale was introduced by Houthakker and Prais (1971) and considers that every individual has different needs for every item of consumption. For example, an adult male will consume a large quantity of say rice whereas an infant child will not consume any rice at all. The 'consumer unit scale' is defined as the effect of an addition of an extra person to the household on the consumption of the commodities and income whereas 'adult equivalent scale' is the effect of an addition of an extra person (as compared to an adult member) on the household commodities consumption and income. The 'consumer unit scale' defines the scale for every individual consumer in terms of particular commodity. Thus if the commodity is women's clothing the scale will give a unit weight to the number of women but zero to others. For food, it is expected that no coefficient will be zero but for young children, it may be very small. To convert this scale in terms of adult equivalent one should take the adult scale in the numeraire then the scale will turn into an adult equivalent. The concept of 'consumer unit scale' helps to determine the effective number of consumers in a household, which can then be used to compare households in terms of economic well being. The consumer unit scales consists of 'income scale' and 'specific scale'. Houthakker and Prais (1971) used these concepts to determine the consumer unit scale, which was originally introduced by Sydenstricker and King (1921) and later used by Singh and Nagar (1973) and Sarangi and Panda (2011).

According to Haque (2006), the 'income scale' measures the relative income or total expenditure required by the households of different composition to maintain the same level of satisfaction whereas the specific scales are the relative consumption expenditure for the 'specific item' of consumption required by different household types. Another important concept in the household consumption analysis is the economies/diseconomies of scale. A household enjoys income economies/diseconomies of scale if it enjoys a higher/lower standard of living compared to a relatively smaller household with the same level of per unit consumer income (Haque, 2006). Houthakker and Prais (1971) introduced the concepts of 'specific' and 'income' economies of scale in household consumption. According to them, 'specific' economies/diseconomies of scale occur when an equi-proportionate change in income and household size (leaving per capita/per unit income unaltered) results in more/less than proportionate change in per capita/per unit expenditure on any specific item. For example, economies of scale may arise when there is bulk purchasing, storage or preparation of foods. There may be discounts for purchasing large quantities, or it may be that a commodity is only sold in multiples of certain minimum quantities and so the cost per unit reduces. It may happen that all items may not reveal economies of scale in their consumption. It may, indeed, quite possible that at a certain stage diseconomies of scale begin. A large household may incur more expenditures on outings and other pleasure trips. Therefore, there can be economies of scale on few items as well as diseconomies of scale on the other items of consumption in the budget, which may offset, at least partly, the effects of economies of scale in the consumption of other items. In consequence, there would be an 'overall effect' of such economies and diseconomies of scale on household consumption. A household is said to be enjoying income economies/diseconomies of scale, if it is able to enjoy higher/lower standard of living than a relatively smaller household, with the same level of per capita/per unit income.

Most of the works like Houthakker and Prais (1971), Singh (1972), Forsyth (1960), Valenzuela and Rebecca (1996), Woodbury (1944) etc. treated the concepts of consumer unit scale and economies of scale separately[1]. As discussed in Haque (2006), there are four methods to construct such equivalence scales. First, consumer unit scales based on nutritional requirements for different age and sex composition (see Visaria, 1979, for more detailed discussion on this type of scale). Since these scales are based on nutrition so consumer unit scale for non-food items cannot be defined by this method. The second approach is based on survey questionnaires about the needs of the members of the household, which are directly collected from the household. Few works who considered the approach are Kapteyn and Van Praag (1978), Goedhart et al. (1977), Van Praag and Van der Sar (1988) etc. This method is also non-viable to be conducted for the entire country primarily because of the huge expenditure incurred to collect them. The third approach is based on 'public opinion' by asking people about the minimal commodity basket required by each type of person, based on the survival needs. And lastly, equivalence scale is being estimated using the expenditure data of the household. This method was used by many authors like Houthakker and Prais (1971), Nicholson (1949), Forsyth (1960), Singh and Nagar (1973), Bojer (1977), Muellbauer (1975, 1977, 1980), Kakwani (1977), Sarangi and Panda (2011) etc. In this chapter, we will consider the last approach and try to find both types of scales i.e. consumer unit scale as well as economies/diseconomies of scale for the specific commodity groups and the income/total expenditure. Many works are done in this direction and they all have tried to find these scales for different age groups and did not consider gender as a dimension to create the group (except a few). This chapter is an attempt to include gender as an extra dimension for the estimation of these scales.

[1]Many authors like Bojer (1977), Deaton and Muellbauer (1980), Muellbauer (1980), Nelson (1988), Murthi (1994), Phipps (1998), Majumder and Chakrabarty (2010) etc. treated these concepts as one. They compared the requirement of specific commodities and income for any household with that of the reference household. In that case, there are a limited number of observations for each type of household.

So, our objective is to find the scales for different age-sex groups and also the effective household size for these commodity groups and total expenditure.

The theory of consumption behaviour developed so far applies to a single individual. In practice, consumption decisions are made by households i.e. a collection of individuals and data on consumption is available only at the household level until special types of primary surveys are conducted (Basu et al., 1986). So analysis at the individual level cannot be done at a satisfactory level. As discussed in the previous chapter 4, Engel relation provides us with a way to deal with the data situation. Engel curve is the relation between consumed commodity and income/total expenditure. If we divide both sides of the Engel relation by the household size then we will get the Engel relation in per capita form. There are two possible ways of considering the effect of the household composition into the Engel curve analysis: (1) households can be classified into different groups such that each group consists of households of identical composition. Engel curve analysis is then applied to each of these groups and the results are compared. This method is simple and direct but the efficiency of this method is constrained by the non-availability of adequate data size (Majumder and Chakrabarty (2010)); (2) adjust for the differences in household composition in the Engel function itself. This is usually done by incorporating weights associated with the specific commodity item for every age-sex groups. These weights are the same as the one proposed by Houthakker and Prais (1971). After estimating these weights by suitable methods, they are converted into equivalent adult units. This method is more prominent in terms of data and its application is more suited to household survey data. In this chapter, we are dealing with the second approach and modify the Engel function, by incorporating the specific and income scales for different age-sex groups, to get the per unit form of Engel relation. Similarly, after some more modification (which will be discussed later) one can get economies of scale form of the Engel relation.

There are various parametric forms of Engel curves; like linear, double-log, semi-log, log-inverse, hyperbola, quadratic etc. but for each commodity, the exact form of Engel relation is unknown therefore, we will consider the nonparametric form of the Engel curve. There are some assumptions: l) There is no change in the relative prices and the regional price structures remain fixed. 2) Tastes and preferences of the society do not vary. 3) The economy (composition of the market) is fixed. Also, we are using total expenditure instead of income as our data i.e National Sample Survey (NSS) data provides the expenditure data only and not the income. Expenditure data has some benefits over income and they are more reliable than income data because total expenditure is expected to be more stable and hence more directly related to the permanent level of living. If income falls (say), the consumption may not fall due to the level of living and consumption habit.

To find the estimate of consumer unit scale and economies of scale, we use a modified version of Singh and Nagar (1973) & Singh (1972) approach. Our approach is also an iterative procedure but uses a non-parametric Engel relation between expenditure on any commodity and the total expenditure. Thus, here an iterative procedure for estimating 'specific' and 'income' consumer unit scales, as well as 'specific' and 'income' economies of scale, is developed which is essentially a modification of the Houthakker and Prais (1971) method and Singh and Nagar (1973) approach.

We will see that the weighted equivalence scales are highest for the adults and lowest for the children (0-3 years) for most of the groups of the commodities, with a few exception like milk, other misc. food expenditure and medical for approximately all types of households. For lower age groups, the male has higher equivalent income scale but as the age increases female dominate in terms of the scale. Also, economies of scales are seen for approximately every item of consumption. Majumder and Chakrabarty (2010) observed that boys cost less than girls. Lancaster et al. (1999) obtain equivalence scales (averaged over three children groups, viz., 0-4 years, 5-14 years, 15-17 years) to be 0.171 for boys and 0.192 for girls.

This chapter is organised as follows: earlier models are discussed in Section 5.2, where Houthakker and Prais (1971)'s method and Singh (1972), Singh and Nagar (1973) iterative approach are discussed in details. In Section 5.3, our proposed model has been discussed, which is based on the Singh and Nagar (1973) iterative approach. Section 5.4 discusses the data and its characteristics. Result which has been obtained are discussed in section 5.5. Finally the chapter is concluded in the last Section 5.6.

5.2 Earlier Models of Consumer Unit Scales

The importance of estimating consumer equivalence scales was realized from the beginning of the Engel curve analysis. This is because household size and composition are important determinants for consumption of any item other than income. The measurement of effective household size is a big problem and the per capita form of Engel curve could not capture the difference in needs for different age-sex groups. For these reasons, from the very beginning, Engel (1895) himself expressed each household in terms of the total number of children (quet, after the name of a famous Belgian Statistician Quetelet), taking the newly born child=1. Later researchers such as Houthakker and Prais (1971), Singh and Nagar (1973), Sarangi and Panda (2011) have considered adult male as a standard member of the family and tried to attach different weights to individuals of different age-sex groups of the households. The method used by them and their shortcomings are discussed in details hereunder.

5.2.1 The Prais and Houthakker Model

Let us assume that there are K age-sex groups in a household and w_{ik} is the value of the k^{th} (k=1,2,......,K) group on the specific unit scale for the i^{th} (i=1,2,.....I) item of consumption also let w_{ok} is the weight of each member in the k^{th} (k=1,2,......,K) group on the income scale. In fact, w_{ik} is the unit consumer coefficient of a member in the k^{th} group for the item i often expressed as per adult equivalent, since weight of an adult male is taken as 1 and the weight of any other member of the family is expressed as proportional to that of an adult male. Let n_{kj} be the number of persons in the k^{th} (k=1,2,......,K) group in the j^{th} (j=1,2,......,J) household then

$$n_j = \sum_{k=1}^{K} n_{kj} \tag{5.1}$$

is the 'unweighted' size of the j^{th} household. And for any specific i^{th} item,

$$n^*_{ij} = \sum_{k=1}^{K} w_{ik} n_{kj} \tag{5.2}$$

is the 'weighted' size of the same household. In other words, n^*_{ij} is the effective household size with respect to the i^{th} commodity item. Thus, the per unit expenditure within a specified period on the i^{th} item of consumption is given by

$$c^*_{ij} = \frac{C_{ij}}{\sum_{k=1}^{K} w_{ik} n_{kj}} \tag{5.3}$$

where, C_{ij} is the expenditure (within a specified period) on the i^{th} item by the j^{th} household.

On the income side, Houthakker and Prais (1971) measured the different age-sex groups of the household members on the income scale. Since, w_{ok} is the value of the k^{th} group on the income scale, the weighted household size with respect to income/total consumer expenditure would be

$$n_{oj}^{*} = \sum_{k=1}^{K} w_{ok} n_{kj} \tag{5.4}$$

which is the effective household size with respect to income/total consumer expenditure.

The per unit income/total consumer expenditure (within the specified period) is given by

$$x_j^{*} = \frac{X_j}{\sum_{k=1}^{K} w_{ok} n_{kj}} \tag{5.5}$$

where X_j is the income or the total household expenditure (within the specified period) for the j^{th} household.

The Engel function in per unit terms can be stated as

$$c_{ij}^{*} = f_i(x_j^{*}) \tag{5.6}$$

which may assume different functional forms for different items of consumption.

Houthakker and Prais (1971) used the semi-log type of function for equation (5.6) and assumed $w_{ok} = 1$ for all k such that $n_j = \sum_{k=1}^{K} n_{kj} = \sum_{k=1}^{K} w_{ok} n_{kj}$; then x_j^{*} becomes $x_j = \frac{X_j}{n_j}$ then equation (5.6) can be written as

$$c_{ij}^{*} = \frac{C_{ij}}{\sum_{k=1}^{K} w_{ik} n_{kj}} = \alpha_i + \beta_i log x_j = \beta_i(\gamma_i + log x_j) \tag{5.7}$$

where,

$$\gamma_i = \frac{\alpha_i}{\beta_i} \tag{5.8}$$

They fix γ_i a priori and run a regression of z_{ij} on n_{1j},n_{2j},.....,n_{Kj} for the following equation.

$$z_{ij} = \frac{C_{ij}}{\gamma_i + logx_j} = \beta_i(\sum_{k=1}^{K} w_{ik}n_{kj}) \tag{5.9}$$

From the estimates of $\beta_i w_{i1}$, $\beta_i w_{i2}$,........,$\beta_i w_{iK}$ they derive the ratios $\frac{w_{i2}}{w_{i1}}$, $\frac{w_{i3}}{w_{i1}}$,........,$\frac{w_{iK}}{w_{i1}}$.

In this approach, different values of γ_i (fixed a priori) will naturally result in different estimates of these ratios. One should choose that value of γ_i which yields the maximum multiple correlation between c^*_{ij} and $\hat{c}^*_{ij}$ or minimum sum of squared errors $\sum(c^*_{ij} - \hat{c}^*_{ij})^2$. Houthakker and Prais (1971) method for estimating the w_{ik}'s depends initially on the arbitrarily chosen numerical values of γ_i, and it also depends on whether a semi-log or a double-log form has been chosen for equation (5.6). This introduces an element of arbitrariness into the model. The numerical values of the estimates differ with the functional form of the Engel curve used. They also do not have separate values for the w_{ok}'s since they assumed $w_{ok} = 1$ and stick to it throughout the analysis.

Houthakker and Prais (1971) assumed that the 'specific' and 'income' economies of scale enters into the Engel function in an exponential way and similarly modified the Engel function as

$$\frac{C_{ij}}{(n_j)^{\theta_i}} = f_i(\frac{X_j}{(n_j)^{\theta_o}}) \tag{5.10}$$

where θ_i is the coefficient of specific economies of scale in the consumption of the ith (i=1 ... I) item and θ_o is the coefficient of income economies of scale. Rest of the symbols are already defined earlier in this chapter so it can be referred from there. If θ_i, θ_o is equal to one then there is neutrality of scale in both the specific and income economies of scale. However, if θ_i, θ_o<1 then there is economies of scale and if its greater than 1 then diseconomies of scale will persist. The degree of the specific and income economies/diseconomies of scale will depend on the extent of which θ_i and θ_o respectively are less/greater than unity.

They also postulate another specification of the Engel function in which per capita expenditure on a specific item (say i^{th}, i = 1, . . . , I) is explained in terms of per capita income and the size of the household, i.e.

$$\frac{C_{ij}}{n_j} = f_i(\frac{X_j}{n_j}, n_j) \tag{5.11}$$

They assumed these two Engel functions are of the double-log form, and so equation (5.10) and (5.11) can be re-written respectively as

$$log(\frac{C_{ij}}{(n_j)^{\theta_i}}) = \alpha_i + \beta_i\ log(\frac{X_j}{(n_j)^{\theta_o}}) \tag{5.12}$$

$$log(\frac{C_{ij}}{n_j}) = \alpha_i + \beta_i\ log(\frac{X_j}{n_j}) + \delta_i\ log\ n_j \tag{5.13}$$

where α_i, β_i and δ_i are the parameters to be estimated for the i^{th} (i = 1, . . ., I) item. If both these Engel curves explain the consumer behaviour of the same set of households, one would have δ_i as sort of composite parameter which implicitly accounts for the effects of both kinds of (specific and income) economies of scale. Comparing equations (5.12) and (5.13) one would find that

$$\delta_i = \beta_i(1 - \theta_o) - (1 - \theta_i) \tag{5.14}$$

Thus, in order to estimate θ_i from equation (5.14) one requires the estimate of θ_o. For this purpose, Houthakker and Prais (1971) postulate two quality equations in the semi-log form corresponding to (5.12) and (5.13) respectively. These are,

$$P_j = \alpha + \beta \; log(\frac{X_j}{(n_j)^{\theta_o}}) \tag{5.15}$$

$$P_j = \alpha + \beta \; log(\frac{X_j}{n_j}) + \delta \; log \; n_j \tag{5.16}$$

where, P_j is the average price paid by the jth (j = 1,..., J) household for various items of their budget and α, β and δ being the parameters to be estimated. It should be noted that P_j is some sort of weighted average of the prices of various items paid by any particular household. Comparing equation (5.15) and (5.16) one can estimate θ_o as

$$\theta_o = (1 - \delta/\beta) \tag{5.17}$$

From equation (5.13) estimate of β and δ can be obtained and then from (5.17) θ_o can be found. From (5.14), θ_i can be calculated after substituting β, δ & θ_o. This model was refuted by Singh (1972). He thinks Houthakker and Prais (1971) estimate θ_i and θ_o from several relationships and naturally complicates the procedure. Besides, the estimate of θ_o will depend on P_j and on the form of the 'quality equations'. This introduces an element of arbitrariness in the estimation of θ_o, since Houthakker and Prais (1971) assume the quality equations viz., (5.15) and (5.16) to be of the semi-log form and 'P_j', the overall price paid by the consumer, to be a weighted average of prices of different items of the budget.

5.2.2 Singh and Nagar Method

Singh and Nagar (1973) worked out an iterative procedure which yields the estimates of both 'specific' and 'income' scales independently of any such restrictions and assumptions as done in Houthakker and Prais (1971). It enables to use several functional forms without arbitrarily fixing the value of any of the parameters. They scan through various available forms and select for the purpose of the analysis the one which (i) provides initial and final critical levels of demand below and above which the consumer would not have any demand for the item in question however low or high his income may be, (ii) satisfies all Slutsky's conditions and iii) explains the maximum variation in the dependent variable. They call such a form the most 'plausible form' of the Engel function. Given several forms of the demand function satisfying the first two criteria of plausibility, the most plausible form of the Engel function will be obtained by scanning through them and selecting the one which explains maximum variation in the dependent variable. The steps of the iterative procedure they propose may now be outlined as follows:

At first they set w_{ik} and w_{ok} equal to unity for all i = 1, . . . , I and k = 1,... , K [2], so that the Engel function (5.6) reduces to the traditional form in per capita terms. The parameters of this Engel function (in per capita terms) can be estimated by any suitable method. They tried several alternative forms of the function, f_i, and selected the most plausible one as mentioned above. They estimate $\hat{f}_i(x_j)$ from the selected form of the Engel function, for each item, i = 1,. . ., I, of consumption. Then regress

$$\frac{C_{ij}}{\hat{f}_i(x_j)} = \sum_{k=1}^{K} w_{ik} n_{kj} \tag{5.18}$$

[2] One can choose any value for w_{ik} and w_{ok}. They chose unity so that they can begin with the traditional Engel function in per capita terms and end up with Engel function in per unit terms. Finally, the two Engel functions can be compared.

where, the estimates $\hat{w}_{i1}$, $\hat{w}_{i2}$,......$\hat{w}_{iK}$ are the estimated specific consumer unit weights.

They worked out an estimate ($\hat{w}_{ok}$) of w_{ok} as follows:

$$\hat{w}_{ok} = \sum_{i=1}^{I} \lambda_i \hat{w}_{ik}, \qquad \lambda_i = \frac{1}{J} \sum_{j=1}^{J} \frac{\frac{C_{ij}}{n_j}}{\frac{X_j}{n_j}} \tag{5.19}$$

Applying these $\hat{w}_{ik}$ and $\hat{w}_{ok}$ values in equation (5.6) and then again they tried to find a suitable functional form for

$$\frac{C_{ij}}{\sum_{k=1}^{K} \hat{w}_{ik} n_{kj}} = f_i\left(\frac{X_j}{\sum_{k=1}^{K} \hat{w}_{ok} n_{kj}}\right) \tag{5.20}$$

after finding the most plausible form f_i, they obtain

$$\hat{f}_i\left(\frac{X_j}{\sum_{k=1}^{K} \hat{w}_{ok} n_{kj}}\right) \tag{5.21}$$

Then regress

$$\frac{C_{ij}}{\hat{f}_i\left(\frac{X_j}{\sum_{k=1}^{K} \hat{w}_{ok} n_{kj}}\right)} = \sum_{k=1}^{K} \hat{w}_{ik} n_{kj} \tag{5.22}$$

To obtain fresh estimates of $\tilde{w}_{i1}$, $\tilde{w}_{i2}$,....,$\tilde{w}_{iK}$ of the w_{ik}'s. The corresponding estimates of w_{ok}'s may be worked out as follows:

$$\tilde{w}_{ok} = \sum_{i=1}^{I} \lambda_i^* \tilde{w}_{ik}, \qquad \lambda_i^* = \frac{1}{J} \sum_{j=1}^{J} \frac{\frac{C_{ij}}{\sum_{k=1}^{K} \hat{w}_{ik} n_{kj}}}{\frac{X_j}{\sum_{k=1}^{K} \hat{w}_{ok} n_{kj}}} \tag{5.23}$$

After replacing $\hat{w}_{ik}$ and $\hat{w}_{ok}$ by $\tilde{w}_{ik}$ and $\tilde{w}_{ok}$, for all i and k they switch back and forth between (5.20) and (5.23) until the process converges.

Although the specific consumer unit weights are estimated under two constraints: (a) $0 \leq w_{ik} \leq 1$ and b) $\sum_{k=1}^{K} w_{ik} = 1$ (if initial value of w_{ik} is equal to $1/K$ instead of 1), these constraints are liable to create problems. They found negative w_{ik}'s if no constraints are imposed and their magnitudes may sometimes be so large that $\sum_{k=1}^{K} w_{ik} n_{kj}$, the number of equivalent adults in the j^{th} household may become negative which could also make $\frac{C_{ij}}{\sum_{k=1}^{K} w_{ik} n_{kj}}$, per unit expenditure on the i^{th} item, negative–a finding which violates the basic assumption of non-negativity of consumption expenditure in the consumer demand theory.

Singh (1972) also proposed an iterative procedure to estimate the specific and income economies of scale. The idea of incorporating the economies of scale parameter in the Engel function is the same as Houthakker and Prais (1971) i.e. they enter into the Engel function exponentially as proposed in equation (5.10). After the estimation of consumer unit scales and incorporating the economies of scale, the Engel relation will be

$$\frac{C_{ij}}{(\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i}} = f_i\left(\frac{X_j}{(\sum_{k=1}^{K} w_{ok} n_{kj})^{\theta_o}}\right) \tag{5.24}$$

θ_o is defined as a weighted average of θ_i's (i=1,2,.....I) (see Appendix of this chapter for the relation between θ_i & θ_o).

Singh (1972) assumed the basic Engel function is of double-log form and thus defined equation (5.24) as

$$log\ \left(\frac{C_{ij}}{(\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i}}\right) = \alpha_i + \beta_i log\ \left(\frac{X_j}{(\sum_{k=1}^{K} w_{ok} n_{kj})^{\theta_o}}\right) + u_{ij} \tag{5.25}$$

Since w_{ik} and w_{ok} are already known, they gave an alternative formulation of equation (5.25) as

$$log\ \frac{C_{ij}}{\sum_{k=1}^{K} w_{ik}n_{kj}} = \alpha_i + \beta_i\ log\ (\frac{X_j}{(\sum_{k=1}^{K} w_{ok}n_{kj})^{\theta_o}}) - (1-\theta_i) log\ (\sum_{k=1}^{K} w_{ik}n_{kj}) + u_{ij} \quad (5.26)$$

They try to solve this in a step-wise manner.

Step-1: They assume $\theta_o^{[1]} = 1$ and so the expression reduces to

$$log\ \frac{C_{ij}}{\sum_{k=1}^{K} w_{ik}n_{kj}} = \alpha_i + \beta_i\ log\ (\frac{X_j}{(\sum_{k=1}^{K} w_{ok}n_{kj})}) - (1-\theta_i) log\ (\sum_{k=1}^{K} w_{ik}n_{kj}) + u_{ij} \quad (5.27)$$

They estimate the values of the parameters α_i, β_i and $(1 - \theta_i^{[1]})$ using OLS. Note $\theta_i^{[1]}$ is the value of θ_i in the first step.

Step-2: Estimate $\theta_o^{[2]} = \sum_{i=1}^{I} \frac{\bar{w}_i \lambda_i \theta_i}{\sum_{i=1}^{I} \bar{w}_i \lambda_i}$ (see Appendix for the relation between θ_i and θ_o).

Step-3: In equation (5.26) if we put $\theta_o^{[2]}$ then it can be rewritten as

$$log\ \frac{C_{ij}}{\sum_{k=1}^{K} w_{ik}n_{kj}} = \alpha_i + \beta_i\ log\ (\frac{X_j}{(\sum_{k=1}^{K} w_{ok}n_{kj})^{\theta_o^{[2]}}}) - (1-\theta_i^{[2]})\ log\ (\sum_{k=1}^{K} w_{ik}n_{kj}) + u_{ij} \quad (5.28)$$

where again the right-hand side is fully measurable and Step-1 will follow. The process will go on and on until the values of θ_i and θ_o converges. The only point which can be made of this method is that although Singh (1972) was against sticking to any particular form of the Engel function, then how could they do so for estimating the economies of scale.

5.3 Proposed Model

We will consider on the lines of Singh and Nagar (1973) and Singh (1972) approach, but taking non-parametric Engel functions. First it will be assumed that w_{ik} and w_{ok} are equal to unity for all i = 1, . . . , I and k = 1,... , K. The Engel function (5.6) reduces to the traditional form in per capita terms. Then estimate the Engel function non parametrically. Compute $\hat{f}_i(x_j)$ for each item, i = 1,. . . ,I, of consumption.

Then regress equation (5.18) using OLS (Ordinary Least Squares) where the estimates $\hat{w}_{i1}$, $\hat{w}_{i2}$,......$\hat{w}_{iK}$ of the regression coefficient are obtained. The estimate of w_{ok} can be obtained from equation (5.19). Substitute these values of $\hat{w}_{ik}$ and $\hat{w}_{ok}$ in equation (5.6) and apply again the non parametric method to estimate f_i i.e. equation (5.21). Then regress equation (5.22) to obtain fresh estimates of $\tilde{w}_{i1}$, $\tilde{w}_{i2}$,....,$\tilde{w}_{iK}$ of the w_{ik}'s. The corresponding estimates of w_{ok}'s may be worked out from equation (5.23). Replacing $\hat{w}_{ik}$ and $\hat{w}_{ok}$ by $\tilde{w}_{ik}$ and $\tilde{w}_{ok}$, for all i and k and then switch back and forth between (5.20) and (5.23) until the process converges.

For the determination of the economies of scale, again it will be assumed that $\theta_i^{[1]} = \theta_o^{[1]} = 1$. So the Engel function will be reduced into the per unit form i.e. equation (5.20), where w_{ik}'s & w_{ok}'s are known. Estimate this equation non-parametrically and find the value of $\hat{f}_i$, then our model can be rewritten in the form

$$\frac{C_{ij}}{\hat{f}_i(\frac{X_j}{(\sum_{k=1}^{K}\hat{w}_{ok}n_{kj})})} = (\sum_{k=1}^{K} w_{ik}n_{kj})^{\theta_i} \tag{5.29}$$

Taking log on both sides, the above equation becomes

$$log \frac{C_{ij}}{\hat{f}_i(\frac{X_j}{(\sum_{k=1}^{K} \hat{w}_{ok} n_{kj})})} = \theta_i^{[2]} \ log \ (\sum_{k=1}^{K} w_{ik} n_{kj}) + u_{ij} \tag{5.30}$$

Calculate $\theta_o^{[2]}$ using the formula in Step-2 of Singh and Nagar approach. Again substitute these values into the following Engel function

$$\frac{C_{ij}}{(\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i^{[2]}}} = f_i(\frac{X_j}{(\sum_{k=1}^{K} w_{ok} n_{kj})^{\theta_o^{[2]}}} \tag{5.31}$$

Estimate f_i non-parametrically. Follow the same procedure as done above until the values of θ_i and θ_o converges.

5.4 Data and its characteristics

This chapter uses National Sample Survey 68^{th} round Type 2 consumer expenditure data[3]. It was collected during July 2011–June 2012 and it is the latest quinquennial surveys on consumer expenditure done by the National Sample Survey Organization (NSSO). 'Type 2' data has modified mixed recall period i.e. different commodity items have different recall periods. Table 5.1 shows the different commodity groups and their corresponding recall periods for the 68^{th} round Type 2 data[4].

We will consider commodities in certain groups, thereby making 10 broad commodity groups: they are (i) cereals and its substitute (ii) pulses and its products (iii) milk and

[3]The survey covers the whole of the Indian Union except those places which remain inaccessible throughout the years.

[4]During the 1990s, the most suitable reference period for the data collection on the items consumed is being argued and has been decided to produce separate thin samples on the basis of alternative schedules. Under this schedules, data on less-frequently used items are collected over a one-year period, while sticking to the 30-day recall period for more frequently used items and for the most frequently used items a 7-day recall period is considered.

its products (iv) salt, sugar, oil (edible) & spices (v) vegetables (vi) egg, fish, meat, fruits & dry fruits (vii) other misc. food expenditure (which includes pan, tobacco, intoxicants, processed food & beverages) (viii) footwear, clothing & bedding (ix) medical and education (x) other misc. non-food expenditure (which includes- fuel and light, entertainment, minor durable goods, toiletry articles, other household consumables, consumer services, conveyance, rent, tax & durables). These items have been named serially starting from c1 to c10 respectively for our convenience, as shown in Table (5.1)[5]. The data are cross-sectional and the unit of observation is household. The total sample size is 101,651 number of households. The survey provides data on the expenditure incurred by the household as a whole on all the consumer goods and services during the reference period. It provides information on the demographic characteristics of the household at the individual level for each household surveyed. In addition, this survey also provides information on the total expenditure of the household.

The data can be divided into the rural and the urban sector, with rural consists of 71.43% of the population. The average monthly per capita consumption expenditure (MPCE) in 2011-12 was approximately Rs.1430 in rural India and Rs.2630 in urban India. Further, for each sector, three expenditure groups are defined according to the MPCE: i) Lower expenditure group–The poorest 30% of India's population with an average MPCE below Rs.963 for rural sector and Rs.1490 for the urban sector. ii) Middle expenditure group–The MPCE fractile classes between 30% and 70%, with the average MPCE, lies between Rs.963–Rs.1522 for the rural sector and between Rs.1490–Rs.2771 for the urban sector. iii) Higher expenditure group–The MPCE fractile classes above 70% with an upper bound on MPCE as Rs.6000 for rural sector and Rs.10,000 for urban sector. Observation above this level is below 2% level and expenditure patterns of these households also show

[5] Items with recall period '7 days' and '365 days' are converted in terms of '30 days' using the formula= (Item Value $\times$ 30)/Corresponding Recall Period).

some erratic behaviour, so they are excluded from the current analysis. In Table 5.2, the expenditure pattern for all the expenditure groups in case of the rural and the urban sector are presented. From the table, it can be seen that for the items cereals, pulses, salt, sugar, edible oil, spices, vegetables, clothing, bedding, and footwear the allocation in the budget share decreases for higher income groups for both of the sectors. And for the items such as milk, egg, fish, meat, fruits and dry fruits, educational expenditure, medical expenditure, and other misc. food and non-food expenditures the budget share increases for the higher income group for both of the sectors.

The age groups considered are the same as in other earlier chapters i.e., 0-3 years, 4-6 years, 7-12 years, 13-18 years, 19 years and above, and for different genders, separate variables will be used. Therefore, there will be 10 age-sex groups. The summary statistics of the variables used are shown in Tables from 5.3 to 5.10 for all types of households. It can be seen from these tables that the average weighted household size is higher for the rural sector than for the urban sector for all types of expenditure groups. Also, for the low expenditure group this average weighted household size is the highest and for the high expenditure group, this average is the lowest. The per capita expenditure for each commodity group and monthly per capita expenditure (under modified mixed reference period) is higher for the urban sector than the rural counterpart. Similarly, the per capita expenditure of different commodity groups and the monthly per capita expenditure are lower for low expenditure group and it increases with the income groups without any exception, which is quite obvious. Therefore, every commodity groups are normal goods, where the per capita consumption increases with the income. Figure (5.1) & (5.2) shows the Engel curve for the rural and the urban sector respectively. From the figure, it can be seen that for most of the food items the Engel curve bends towards the X-axis, which implies these goods are necessary goods and the consumption will not increase at an increasing rate with the income. For medical, education and other misc.

non-food expenditure the Engel curve bends towards the Y-axis showing the property of luxurious goods. This is true for both the sectors. For almost all age and expenditure groups, the mean of the number of the male is greater than the number of females except in rural sector for adult group (in case of overall expenditure group when taken together) and in rural lower expenditure groups for age 7-12 and adult category.

5.5 Result and Empirical Investigation

The primary interest of the analysis is to estimate the adult equivalent consumer unit scales for different commodities as well as different age-sex groups, which is shown in Table (5.11) - (5.18). In these tables, the estimated values of specific adult equivalent(male) scales for all the 10 commodity groups are shown in the upper section of each table. Also, the income adult equivalent scale is shown in the last column in each of these tables. The values of the economies of scales (θ_i) are shown in the 11^{th} row of these tables. From the tables, one can see that all the equivalent consumer unit scales and the economies of scale are positive, which is quite a desirable result. This is unlike what was obtained by Singh and Nagar(1973). They found many negative values for the w_{ik}'s which was quite strange for them and it was difficult to explain because they estimated the values assuming the constraint $1 \geq w_{ik} \geq 0$. The data that they were using was based on the individual commodity items and not the commodity groups which lead to a high density of zeros in the dependent variable. This might have been one of the reasons why they were getting negative weights. In our case, we are using the commodity groups which reduced the number of zero observation[6]. In the middle section of each of these tables, the relevant test statistic value for the hypothesis testing are presented. The different hypotheses are (i) no difference between the gender equivalence scale for each age groups, (i.e. H_o:$w_{ik} = w_{ik'}$ vs H_1: $w_{ik} \neq w_{ik'}$) where k is for male related groups and k' is for the

[6] If there are a higher percentage of zeros in the dependent variable then one should go for the double-hurdle model or similar types of models. For more details see Demoussis and Mihalopoulos (2001), Yen and Jones (1997) etc.

corresponding female related variable under the same age groups and (ii) the economies of scale is neutral (i.e. H_o: $\theta_i = 1$ vs H_1: $\theta_i \neq 1$). The last section of the tables, shows the effective weighted household size, the per capita consumption and per unit consumption of these variables for each of the commodity groups and the income as a whole.

The result for the rural and the urban sectors when considered for all expenditure together are shown in Tables 5.11 and 5.12 respectively. The iterative procedure for consumer unit scales converges rapidly for both the sectors. The number of iterations is 8 and 9 respectively for the rural and the urban sectors. If we compare the magnitudes of the consumer unit weights then it can be seen that in the rural sector, for $0-3$ years, the weights are higher for male than for females for almost all the categories except cereals and other misc. food expenditure. For $4-6$ years age groups, a similar pattern is noticed but here the commodities where female consumer unit's weights are greater than the corresponding male counterpart increases. In general, it is seen that for lower age groups the consumer unit weights are higher for males than for females in the rural sector. On the other hand, for the urban sector, only the age group 4-6 shows a similar pattern where male consumer unit scale is higher than the female consumer unit scale. For cereals and milk, we found higher consumer unit scales for the male for approximately all the age groups in the urban sector. But overall, we can see that males have lower consumer weights than the female counterparts. The values of the economies of scale (θ_i) are less than 1 for all the commodities and income in case of the rural sector. Also, the hypothesis testing result shows that they are not equal to 1, which implies that there exist economies of scale in the consumption of the commodities and income. For the urban sector, the economies of scale are less than 1 for every commodity except spices, salt, sugar and oil. Further, the hypothesis testing for the economies of scale ($\theta_i = 1$) is also rejected at 1% significant level which means that there are economies of scale in the consumption of these commodities. When comparing the effective household size, i.e. the weighted household size where the weights are the consumer unit weights, with the average household size it is seen that it is always less than the average household size in case of the rural sector (except medical and educational expenditure category), whereas for the urban sector it is lower for all except egg, fish, meat, fruits and medical, educational experience groups. When we look into the per capita and per unit consumption of these commodities, we can see that per capita consumption is lower than the per unit consumption except for the one where effective household size is larger than the average household size.

In the Section 5.4, it has been said that the households can be divided into three categories based on the monthly per capita consumer expenditure (MPCE) i.e. low, middle, and high expenditure groups. The consumer unit scales for the low expenditure group in case of the rural sector are shown in Table 5.13. Equivalence weights favour males for $0-3$ and $4-6$ age group only for the expensive commodities. The income scale favours males for these age groups. For the other age groups, there are one or two where males are favoured but mostly the female's weights are higher in terms of the equivalence scale. Also, almost all of the θ_i(s) are significantly different from 1, the only exception being the egg, fish, and meat category. Each of the θ_i values is less than 1 implying economies of scale in the consumption of these items with the only exception of milk. Effective household size is less than the average household size for many commodities and also for the income. Similarly, the per capita values are higher than the per unit consumption of those commodities where the average household size is less than the effective household size, i.e. for commodities such as vegetables, other misc. food expenditure, clothing bedding and footwear and medical and educational expenditure categories etc. In case of the urban sector (see Table 5.14), most of the weighted consumer unit scale favours females with the only exception in the case of the adult age category i.e. 19 years and above where males have higher weights in the consumer scale for all the commodities as well as income. The θ_i is also significantly different from 1 for all the commodities and all of them shows economies of scale in consumption the only exception is misc. non-food expenditure category. The effective household size and the per capita consumption is lower than the average household size and the per unit consumption respectively for almost all of the commodities and the income except medical and educational expenditure.

For the middle income/expenditure groups, in the rural sector i.e. Table 5.15 shows that the adult equivalence consumer unit scale favours males only for the pulses and milk categories for almost all the age groups. But mostly they favour the female counterparts. The value of the θ_i is significantly different from 1 for all the commodities and the values of θ_i show that there are economies of scale in the consumption of the commodities. The value of the effective household size is higher than the average household size for milk, other misc. food expenditure, clothing, education, and medical expenditure. In case of the urban sector, i.e. Table 5.16, it can be seen that comparatively lower number of cases where males adult equivalence weights are higher than the female counterparts. Here also the values of θ_i show economies of scale for all the commodities. Effective household size is less than the average household size for many commodities but it is higher for eggs, fish, meat, clothing and medical expenditure etc. Also, for the income scale, a higher effective household size is observed.

Lastly, in case of the higher income group, the adult equivalent consumer unit scales shows a similar picture for both of the sectors (see Tables 5.17 and 5.18), i.e. females have higher weights as compared to the male counterparts. θ_i is less than 1 for all the commodities and also statistically significant at 1% level for both the sectors. The effective weighted household size is less than the average household size in case of cereals, spices, vegetables, eggs, and other misc. food expenditures in case of the rural sector. For the urban sector only in case of cereals, pulses, and other miscellaneous food expenditure shows that the effective weighted household size is less than the average household size.

5.6 Conclusion

In this chapter, a modified iterative procedure based on Singh (1972) and Singh and Nagar (1973) approach has been proposed, assuming the Engel relation is nonparametric. This procedure is used to find the adult equivalence scales for 10 commodity groups and income. Also, the economies of scale and the effective weighted household size are estimated after the estimation of consumer unit scales. The data used for the analysis is NSS 68^{th} round type 2 data as it has different recall period which is more appropriate according to the Expert group if estimation is done for different commodity groups.

The results in this chapter have many applications because it gives us the adult equivalence scales for each group of members in the household for every commodity. Firstly, it gives the effective household size of a household for each item. The effective household size may be different for different commodities. Secondly, by looking at the adult equivalent scales, it may be possible to find out whether a group of the member is deprived or not?

The consumer unit scales are also found for different expenditure groups. These expenditure groups are defined on the basis of the decile groups of MPCE. Thus three types of expenditure groups are defined the low, middle and high-income group. The estimation is also done for the overall expenditure group. For approximately all categories of households, it is seen that female consumer unit scale is higher than the male counterpart for most of the commodities. In case of the rural sector and for the lower income group, it is seen that the income scale favours the males in the case of the children category, but as the income rises, the incomes scale favours the female counterpart.

We found positive values for the consumer unit scale and the economies of scale unlike Singh and Nagar (1973) for all the commodities and also for the income. The weighted equivalence scales are highest for the adults and lowest for the children (0-3 years) for most of the commodities with few exceptions like milk, other misc. food expenditure and medical expenses. Also, economies of scales are there for approximately every commodity groups.

5.7 APPENDIX-A

5.7.1 Relation between θ_i & θ_o

Given the budget constraint,

$$\sum_{i=1}^{I} C_{ij} = X_j \tag{5.32}$$

Differentiating the budget constraint with respect to X_j,

$$\sum_{i=1}^{I} \frac{\delta C_{ij}}{\delta X_j} = 1 \tag{5.33}$$

Now, $\frac{\delta C_{ij}}{\delta X_j} = \frac{(\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i}}{(\sum_{k=1}^{K} w_{ok} n_{kj})^{\theta_o}} f_i'$ from equation (5.24)

$$\sum_{i=1}^{I} \frac{\delta C_{ij}}{\delta X_j} = \sum_{i=1}^{I} \frac{(\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i}}{(\sum_{k=1}^{K} w_{ok} n_{kj})^{\theta_o}} f_i' = 1 \tag{5.34}$$

Again, differentiating the budget constraint with respect to n_{kj} will give

$$\sum_{i=1}^{I} \frac{\delta C_{ij}}{\delta n_{kj}} = 0 \tag{5.35}$$

Now $\frac{\delta C_{ij}}{\delta n_{kj}} = \theta_i (\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i - 1} w_{ik} f_i - \frac{(\sum_{k=1}^{K} w_{ik} n_{kj})^{\theta_i} f_i' X_j \theta_o w_{ok}}{(\sum_{k=1}^{K} w_{ok} n_{kj})^{\theta_o + 1}}$ from equation (5.24)

$\sum_{i=1}^{I} \frac{\delta C_{ij}}{\delta n_{kj}} = \sum_{i=1}^{I} \frac{\theta_i w_{ik} C_{ij}}{\sum_{k=1}^{K} w_{ik} n_{kj}} - \frac{X_j \theta_o w_{ok}}{\sum_{k=1}^{K} w_{ok} n_{kj}} = 0$ from equation (5.24) and (5.34)

Solving for θ_o will give

$\theta_o = \sum_{i=1}^{I} (\frac{w_{ik} \sum_{k=1}^{K} w_{ok} n_{kj} C_{ij}}{w_{ok} \sum_{k=1}^{K} w_{ik} n_{kj} X_j}) \theta_i$

If we substitute the value of $w_{ok} = \sum_{i=1}^{I} (\frac{\sum_{k=1}^{K} w_{ok} n_{kj} C_{ij}}{\sum_{k=1}^{K} w_{ik} n_{kj} X_j}) w_{ik}$ in the expression for θ_o the expression will transform to[7]

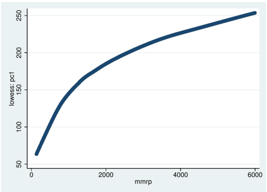
$$\theta_o = \sum_{i=1}^{I} \frac{(w_{ik} C_{ij} \theta_i)/(\sum_{k=1}^{K} w_{ik} n_{kj})}{\sum_{i=1}^{I} w_{ik} C_{ij} / \sum_{k=1}^{K} w_{ik} n_{kj}} \tag{5.36}$$

Replace w_{ik} by $\bar{w}_i = \frac{1}{K}(\sum_{k=1}^{K} w_{ik})$, and $C_{ij}/\sum_{k=1}^{K} w_{ik} n_{kj}$ by say $\lambda_i = \frac{1}{J}\sum_{j=1}^{J}(C_{ij}/\sum_{k=1}^{K} w_{ik} n_{kj})$ will change the expression for θ_o to[8]

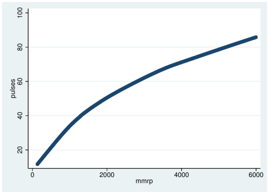
$$\theta_o = \sum_{i=1}^{I} \frac{\bar{w}_i \lambda_i \theta_i}{\sum_{i=1}^{I} \bar{w}_i \lambda_i} \tag{5.37}$$

[7] see Houthakker and Prais (1971)
[8] see Singh (1972) for more details

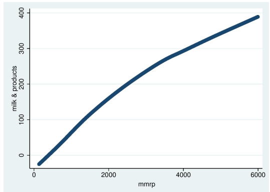
5.7.2 Figures



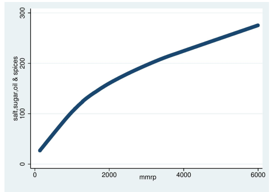
(a) Relation between c1 and mpce



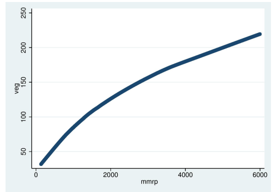
(b) Relation between c2 and mpce



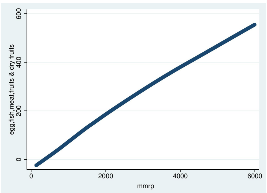
(c) Relation between c3 and mpce



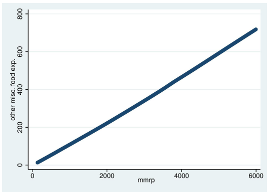
(d) Relation between c4 and mpce



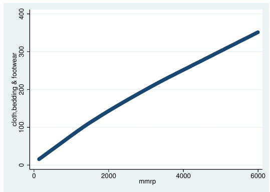
(e) Relation between c5 and mpce



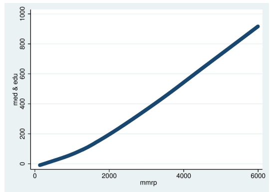
(f) Relation between c6 and mpce



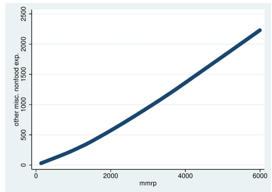
(g) Relation between c7 and mpce



(h) Relation between c8 and mpce

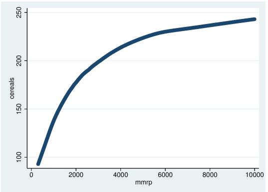


(i) Relation between c9 and mpce

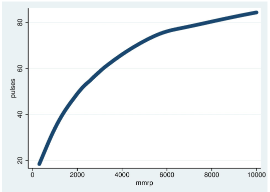


(j) Relation between c10 and mpce

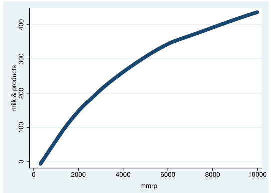
Figure 5.1: Non-parametric relation between per capita consumption of different commodity bundle and mpce for the rural sector



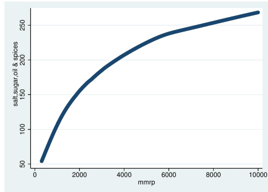
(a) Relation between c1 and mpce



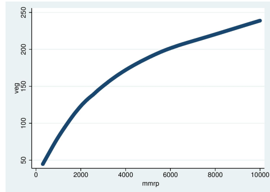
(b) Relation between c2 and mpce



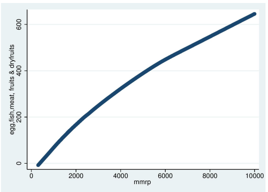
(c) Relation between c3 and mpce



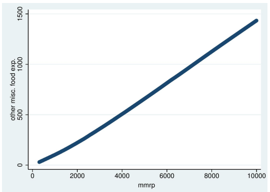
(d) Relation between c4 and mpce



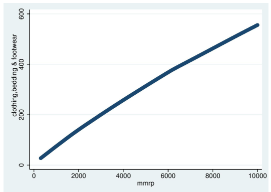
(e) Relation between c5 and mpce



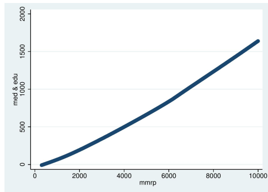
(f) Relation between c6 and mpce



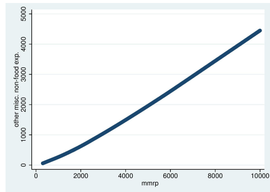
(g) Relation between c7 and mpce



(h) Relation between c8 and mpce



(i) Relation between c9 and mpce



(j) Relation between c10 and mpce

Figure 5.2: Non-parametric relation between per capita consumption of different commodity bundle and mpce for the urban sector

5.7.3 Tables

Table 5.1: 68^{th} round Type 2 data item-wise recall period

Items	variable name	recall period
Cereals and its substitute	c1	30 days
Pulses and its product	c2	30 days
milk and its products	c3	30 days
salt and sugar	c4	30 days
edible oil		7 days
spices		7 days
vegetables	c5	7 days
egg,fish and meat	c6	7 days
fruits		7 days
dry fruits		7 days
beverages	c7	7 days
served processed food		7 days
packed processed food		7 days
pan,tobacco, intoxicant		7 days
clothing	c8	365 days
bedding		365 days
footwear		365 days
education	c9	365 days
medical institutional		365 days
medical non-institutional		30 days
fuel and light	c10	30 days
entertainment		30 days
minor durable goods		30 days
toilet articles		30 days
other household consumables		30 days
consumer services		30 days
conveyance		30 days
rent		30 days
consumer taxes		30 days
durables		365 days

[a] Commodity groups are recalled from NSS 2011-12 Type-2 Consumer Expenditure Household Survey.
[b] Items with recall periods '7 days' and '365 days' are converted into '30 days' using the formula=(item value*30/ corresponding recall period).
[c] c1 –c10 are the different commodity groups defined for the current analysis.

Table 5.2: Expenditure (%) pattern for Rural and Urban different expenditure group on various commodity items.

Commodity Groups	Description	Rural Expenditure Group				Urban Expenditure Group			
		Lower	Middle	Higher	All Groups	Lower	Middle	Higher	All Groups
c1	Cereals and its substitute	16.88	12.55	8.03	11.10	12.70	8.45	4.92	7.22
c2	Pulses and its product	3.83	3.30	2.45	2.98	3.48	2.58	1.61	2.21
c3	milk and its products	5.75	8.12	9.12	8.20	7.48	8.39	6.82	7.46
c4	salt and sugar,edible oil & spices	11.24	10.32	7.78	9.25	10.56	8.02	4.85	6.74
c5	vegetables	9.11	7.50	5.47	6.79	7.40	5.82	3.78	4.99
c6	egg, fish, meat, fruits & dry fruits	5.45	7.64	8.36	7.62	6.97	7.82	7.22	7.40
c7	other misc. food expenditure	10.59	11.08	11.09	11.00	9.91	9.94	10.54	10.24
c8	clothing, bedding & footwear	7.97	7.56	6.59	7.17	7.39	6.92	6.16	6.59
c9	education & medical expenditure	6.06	7.70	12.70	9.83	7.46	10.64	14.59	12.23
c10	Other misc. non-food expenditure	23.12	24.23	28.42	26.05	26.65	31.41	39.51	34.92
MMRP	Monthly per capita expenditure (Modified Mixed reference period)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

[a] Coefficients are rounded to two decimal place.
[b] Commodity groups are redefined for the analysis from NSS 2011-12 Type-2 Consumer Expenditure Household Survey.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.3: Summary Statistics of the variables for rural sector (All Expenditure Group)

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	58775	0.16	0.42
f0003	No. of female with 0-3 years age	58775	0.14	0.40
m0406	No. of male with 4-6 years age	58775	0.16	0.40
f0406	No. of female with 4-6 years age	58775	0.14	0.38
m0712	No. of male with 7-12 years age	58775	0.33	0.61
f0712	No. of female with 7-12 years age	58775	0.30	0.59
m1318	No. of male with 13-18 years age	58775	0.34	0.62
f1318	No. of female with 13-18 years age	58775	0.27	0.56
m19+	No. of male with age 19 years & above	58775	1.38	0.86
f19+	No. of female with age 19 years & above	58775	1.39	0.74
hhsize	household size	58775	4.61	2.18
pcfe	percapita food expenditure	58775	785.53	364.67
pc_1	percapita expenditure on c1 (in INR)	58775	153.07	67.04
pc_2	percapita expenditure on c2(in INR)	58775	41.12	24.94
pc_3	percapita expenditure on c3 (in INR)	58775	113.17	140.04
pc_4	percapita expenditure on c4 (in INR)	58775	127.62	61.07
pc_5	percapita expenditure on c5 (in INR)	58775	93.71	50.52
pc_6	percapita expenditure on c6 (in INR)	58775	105.08	125.33
pc_7	percapita expenditure on c7 (in INR)	58775	151.75	173.45
pc_8	percapita expenditure on c8 (in INR)	58775	98.85	60.50
pc_9	percapita expenditure on c9 (in INR)	58775	135.59	233.41
pc_10	percapita expenditure on c10 (in INR)	58775	359.39	294.99
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	58775	1379.35	730.28

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values below Rs.6000 are considered here. Only 1.5 % observations are above Rs 6000 MPCE.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.4: Summary Statistics of the variables used for urban sector (All Expenditure Groups)

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	41124	0.12	0.36
f0003	No. of female with 0-3 years age	41124	0.11	0.35
m0406	No. of male with 4-6 years age	41124	0.11	0.34
f0406	No. of female with 4-6 years age	41124	0.10	0.33
m0712	No. of male with 7-12 years age	41124	0.25	0.53
f0712	No. of female with 7-12 years age	41124	0.22	0.49
m1318	No. of male with 13-18 years age	41124	0.27	0.56
f1318	No. of female with 13-18 years age	41124	0.22	0.51
m19+	No. of male with age 19 years & above	41124	1.39	0.88
f19+	No. of female with age 19 years & above	41124	1.31	0.84
hhsize	household size	41124	4.11	2.15
pcfe	percapita food expenditure	41124	1117.08	564.62
pc_1	percapita expenditure on c1 (in INR)	41124	174.38	73.40
pc_2	percapita expenditure on c2 (in INR)	41124	53.31	29.27
pc_3	percapita expenditure on c3 (in INR)	41124	180.09	157.23
pc_4	percapita expenditure on c4 (in INR)	41124	162.76	73.80
pc_5	percapita expenditure on c5 (in INR)	41124	120.49	70.70
pc_6	percapita expenditure on c6 (in INR)	41124	178.65	186.21
pc_7	percapita expenditure on c7 (in INR)	41124	247.41	340.65
pc_8	percapita expenditure on c8 (in INR)	41124	159.28	117.57
pc_9	percapita expenditure on c9 (in INR)	41124	295.46	441.81
pc_10	percapita expenditure on c10 (in INR)	41124	843.52	804.15
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed Reference Period)	41124	2415.35	1543.31

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values below Rs.10000 are considered here. Only 2 % observations are above Rs 10000 MPCE.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.5: Summary Statistics of the variables for rural lower expenditure group

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	11119	0.25	0.50
f0003	No. of female with 0-3 years age	11119	0.23	0.49
m0406	No. of male with 4-6 years age	11119	0.24	0.47
f0406	No. of female with 4-6 years age	11119	0.22	0.46
m0712	No. of male with 7-12 years age	11119	0.43	0.69
f0712	No. of female with 7-12 years age	11119	0.45	0.71
m1318	No. of male with 13-18 years age	11119	0.38	0.66
f1318	No. of female with 13-18 years age	11119	0.33	0.62
m19+	No. of male with age 19 years & above	11119	1.41	0.87
f19+	No. of female with age 19 years & above	11119	1.44	0.73
hhsize	household size	11119	5.39	2.19
pcfe	percapita food expenditure	11119	478.04	109.37
pc_1	percapita expenditure on c1 (in INR)	11119	128.38	51.62
pc_2	percapita expenditure on c2(in INR)	11119	29.16	14.65
pc_3	percapita expenditure on c3 (in INR)	11119	43.74	51.15
pc_4	percapita expenditure on c4 (in INR)	11119	85.52	29.58
pc_5	percapita expenditure on c5 (in INR)	11119	69.28	30.50
pc_6	percapita expenditure on c6 (in INR)	11119	41.43	46.84
pc_7	percapita expenditure on c7 (in INR)	11119	80.52	55.39
pc_8	percapita expenditure on c8 (in INR)	11119	60.64	25.36
pc_9	percapita expenditure on c9 (in INR)	11119	46.10	51.92
pc_10	percapita expenditure on c10 (in INR)	11119	175.84	57.97
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	11119	760.63	141.89

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values below Rs.963 are considered here.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.6: Summary Statistics of the variables for urban lower expenditure group

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	11922	0.20	0.46
f0003	No. of female with 0-3 years age	11922	0.19	0.46
m0406	No. of male with 4-6 years age	11922	0.19	0.44
f0406	No. of female with 4-6 years age	11922	0.19	0.44
m0712	No. of male with 7-12 years age	11922	0.39	0.65
f0712	No. of female with 7-12 years age	11922	0.38	0.65
m1318	No. of male with 13-18 years age	11922	0.39	0.68
f1318	No. of female with 13-18 years age	11922	0.35	0.63
m19+	No. of male with age 19 years & above	11922	1.53	1.01
f19+	No. of female with age 19 years & above	11922	1.51	0.85
hhsize	household size	11922	5.33	2.30
pcfe	percapita food expenditure	11922	640.49	161.68
pc_1	percapita expenditure on c1 (in INR)	11922	139.02	54.62
pc_2	percapita expenditure on c2(in INR)	11922	38.05	17.56
pc_3	percapita expenditure on c3 (in INR)	11922	81.91	69.06
pc_4	percapita expenditure on c4 (in INR)	11922	115.65	40.19
pc_5	percapita expenditure on c5 (in INR)	11922	81.07	35.54
pc_6	percapita expenditure on c6 (in INR)	11922	76.31	70.56
pc_7	percapita expenditure on c7 (in INR)	11922	108.48	84.71
pc_8	percapita expenditure on c8 (in INR)	11922	80.92	36.86
pc_9	percapita expenditure on c9 (in INR)	11922	81.73	82.28
pc_10	percapita expenditure on c10 (in INR)	11922	291.85	114.60
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed Reference Period)	11922	1095.00	247.19

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values below Rs.1490 are considered here.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.7: Summary Statistics of the variables for rural middle expenditure group

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	20540	0.16	0.41
f0003	No. of female with 0-3 years age	20540	0.15	0.40
m0406	No. of male with 4-6 years age	20540	0.17	0.41
f0406	No. of female with 4-6 years age	20540	0.14	0.38
m0712	No. of male with 7-12 years age	20540	0.35	0.62
f0712	No. of female with 7-12 years age	20540	0.32	0.60
m1318	No. of male with 13-18 years age	20540	0.36	0.63
f1318	No. of female with 13-18 years age	20540	0.29	0.57
m19+	No. of male with age 19 years & above	20540	1.42	0.87
f19+	No. of female with age 19 years & above	20540	1.41	0.73
hhsize	household size	20540	4.77	2.12
pcfe	percapita food expenditure	20540	733.34	142.16
pc_1	percapita expenditure on c1 (in INR)	20540	152.08	59.26
pc_2	percapita expenditure on c2(in INR)	20540	40.05	19.36
pc_3	percapita expenditure on c3 (in INR)	20540	98.41	87.09
pc_4	percapita expenditure on c4 (in INR)	20540	125.04	42.33
pc_5	percapita expenditure on c5 (in INR)	20540	90.93	39.82
pc_6	percapita expenditure on c6 (in INR)	20540	92.58	81.43
pc_7	percapita expenditure on c7 (in INR)	20540	134.24	94.97
pc_8	percapita expenditure on c8 (in INR)	20540	91.65	36.27
pc_9	percapita expenditure on c9 (in INR)	20540	93.38	96.65
pc_10	percapita expenditure on c10 (in INR)	20540	293.64	98.89
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	20540	1212.01	157.53

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values between Rs.963 and Rs.1522 are considered here.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.8: Summary Statistics of the variables for urban middle expenditure group

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	14707	0.13	0.36
f0003	No. of female with 0-3 years age	14707	0.11	0.35
m0406	No. of male with 4-6 years age	14707	0.11	0.34
f0406	No. of female with 4-6 years age	14707	0.10	0.32
m0712	No. of male with 7-12 years age	14707	0.27	0.55
f0712	No. of female with 7-12 years age	14707	0.21	0.48
m1318	No. of male with 13-18 years age	14707	0.28	0.55
f1318	No. of female with 13-18 years age	14707	0.22	0.51
m19+	No. of male with age 19 years & above	14707	1.44	0.90
f19+	No. of female with age 19 years & above	14707	1.37	0.83
hhsize	household size	14707	4.25	2.03
pcfe	percapita food expenditure	14707	1050.99	243.33
pc_1	percapita expenditure on c1 (in INR)	14707	174.12	62.07
pc_2	percapita expenditure on c2(in INR)	14707	53.15	23.90
pc_3	percapita expenditure on c3 (in INR)	14707	172.77	115.22
pc_4	percapita expenditure on c4 (in INR)	14707	165.13	55.83
pc_5	percapita expenditure on c5 (in INR)	14707	119.89	56.84
pc_6	percapita expenditure on c6 (in INR)	14707	161.15	125.30
pc_7	percapita expenditure on c7 (in INR)	14707	204.78	187.31
pc_8	percapita expenditure on c8 (in INR)	14707	142.65	61.99
pc_9	percapita expenditure on c9 (in INR)	14707	219.18	194.49
pc_10	percapita expenditure on c10 (in INR)	14707	647.13	238.85
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	14707	2059.94	356.16

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE(MMRP) values between Rs.1490 and Rs.2771 are considered here.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.9: Summary Statistics of the variables for rural higher expenditure group

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	27116	0.10	0.33
f0003	No. of female with 0-3 years age	27116	0.08	0.29
m0406	No. of male with 4-6 years age	27116	0.10	0.32
f0406	No. of female with 4-6 years age	27116	0.07	0.27
m0712	No. of male with 7-12 years age	27116	0.21	0.50
f0712	No. of female with 7-12 years age	27116	0.17	0.44
m1318	No. of male with 13-18 years age	27116	0.28	0.56
f1318	No. of female with 13-18 years age	27116	0.21	0.49
m19+	No. of male with age 19 years & above	27116	1.32	0.84
f19+	No. of female with age 19 years & above	27116	1.31	0.74
hhsize	household size	27116	3.86	1.98
pcfe	percapita food expenditure	27116	1170.08	402.09
pc_1	percapita expenditure on c1 (in INR)	27116	179.61	79.80
pc_2	percapita expenditure on c2(in INR)	27116	54.76	32.24
pc_3	percapita expenditure on c3 (in INR)	27116	204.07	199.54
pc_4	percapita expenditure on c4 (in INR)	27116	174.07	72.81
pc_5	percapita expenditure on c5 (in INR)	27116	122.39	63.86
pc_6	percapita expenditure on c6 (in INR)	27116	187.00	175.42
pc_7	percapita expenditure on c7 (in INR)	27116	248.19	267.17
pc_8	percapita expenditure on c8 (in INR)	27116	147.59	77.39
pc_9	percapita expenditure on c9 (in INR)	27116	284.25	368.73
pc_10	percapita expenditure on c10 (in INR)	27116	635.94	400.77
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	27116	2237.86	763.05

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values between Rs.1522-6000 are considered here.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.10: Summary Statistics of the variables for urban higher expenditure group

Variables	Description	Obs	Mean	Std. Dev.
m0003	No. of male with 0-3 years age	14495	0.06	0.25
f0003	No. of female with 0-3 years age	14495	0.06	0.25
m0406	No. of male with 4-6 years age	14495	0.06	0.25
f0406	No. of female with 4-6 years age	14495	0.04	0.21
m0712	No. of male with 7-12 years age	14495	0.14	0.39
f0712	No. of female with 7-12 years age	14495	0.12	0.35
m1318	No. of male with 13-18 years age	14495	0.19	0.45
f1318	No. of female with 13-18 years age	14495	0.14	0.39
m19+	No. of male with age 19 years & above	14495	1.25	0.74
f19+	No. of female with age 19 years & above	14495	1.12	0.79
hhsize	household size	14495	3.18	1.70
pcfe	percapita food expenditure	14495	1714.16	625.85
pc_1	percapita expenditure on c1 (in INR)	14495	212.12	85.73
pc_2	percapita expenditure on c2(in INR)	14495	69.65	36.42
pc_3	percapita expenditure on c3 (in INR)	14495	294.20	196.66
pc_4	percapita expenditure on c4 (in INR)	14495	209.21	90.94
pc_5	percapita expenditure on c5 (in INR)	14495	163.01	89.38
pc_6	percapita expenditure on c6 (in INR)	14495	311.53	252.36
pc_7	percapita expenditure on c7 (in INR)	14495	454.44	528.73
pc_8	percapita expenditure on c8 (in INR)	14495	265.60	151.32
pc_9	percapita expenditure on c9 (in INR)	14495	629.06	669.82
pc_10	percapita expenditure on c10 (in INR)	14495	1703.85	1015.90
MPCE (MMRP)	Monthly per capita expenditure (in INR) (Modified Mixed reference period)	14495	4312.66	1547.87

[a] Coefficients are rounded to two decimal place.
[b] Households with MPCE (MMRP) values between 2771-10000 are considered.
[c] Percapita food expenditure to MPCE are all monthly figures.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.11: 68^{th} round item-wise estimate of the consumer unit scales & the economies of scale for the rural sector.

	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.46	0.89	1.42	0.97	0.83	0.85	1.00	0.82	1.33	0.90	0.92
f0003	0.49	0.75	1.20	0.83	0.78	0.76	1.02	0.77	1.29	0.78	0.84
m0406	0.68	0.85	0.89	0.89	0.84	0.61	1.08	0.80	0.88	0.68	0.79
f0406	0.69	0.77	0.87	0.86	0.86	0.69	1.11	0.76	0.90	0.58	0.77
m0712	0.74	0.75	0.67	0.76	0.82	0.49	1.02	0.77	1.05	0.58	0.73
f0712	0.77	0.77	0.67	0.82	0.84	0.77	1.23	0.81	1.24	0.67	0.83
m1318	0.82	0.83	0.74	0.80	0.87	0.49	0.65	0.88	1.28	0.57	0.73
f1318	0.89	0.82	0.79	0.94	0.99	0.89	0.80	1.07	1.92	0.75	0.93
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.91	1.07	0.84	1.01	0.96	1.01	0.69	1.13	1.65	1.07	1.02
θ_i (economies of scale)	0.97	0.96	0.96	0.99	0.97	0.95	0.92	0.98	0.87	0.99	0.96
F-test: male vs female in different age group											
0-3	1.54	11.74***	8.49***	22.28***	2.83*	1.93	0.21	1.59	0.07	9.98***	
4-6	0.15	3.27*	0.02	1.18	0.33	1.40	0.47	1.03	0.01	6.80***	
7-12	1.44	0.39	0.01	8.79***	0.58	35.64 ***	35.75***	1.67	5.07**	10.73***	
13-18	11.80***	0.18	1.01	43.19***	21.33***	72.15***	17.03***	44.08***	54.32***	45.90***	
19+	24.22 ***	7.55***	11.90***	0.62	2.91*	0.02	90.35***	29.86***	70.16***	8.28***	
F-test: ($\theta_i = 1$)	817.11***	1179.65***	1100.19***	251.59***	916.72***	1200.46***	3590.99***	805.61***	7422.83***	386.07***	
Effective weighted household size	3.98	4.34	4.10	4.34	4.30	4.02	4.11	4.51	6.01	4.10	4.29
Per capita consumption	153.07	41.12	113.17	127.62	93.71	105.08	151.75	98.85	135.59	359.39	1379.35
Per unit consumption	176.89	43.49	127.62	135.04	100.21	119.50	171.16	100.70	103.55	400.28	1477.37

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 8.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.12: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the urban sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.50	0.64	1.30	0.72	0.62	0.76	0.96	0.71	0.89	0.80	0.81
f0003	0.53	0.68	1.24	0.75	0.72	1.14	1.17	0.76	0.70	0.81	0.85
m0406	0.66	0.80	0.98	0.77	0.82	0.71	1.01	0.86	1.00	0.76	0.82
f0406	0.63	0.76	0.73	0.72	0.81	0.85	0.97	0.82	0.87	0.68	0.76
m0712	0.82	0.78	0.66	0.74	0.75	0.72	0.92	0.80	1.14	0.50	0.72
f0712	0.75	0.82	0.65	0.79	0.82	0.83	1.00	0.87	1.38	0.69	0.83
m1318	0.86	0.86	0.56	0.78	0.89	0.71	0.75	0.84	1.41	0.50	0.74
f1318	0.84	0.85	0.70	0.84	0.89	0.86	0.76	0.91	1.55	0.65	0.84
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.93	0.97	0.98	1.02	1.06	1.26	0.79	1.13	1.50	1.09	1.08
θ_i (economies of scale)	0.99	0.98	0.97	1.00	0.99	0.95	0.90	0.98	0.90	0.99	0.97
F-test: male vs female in different age group											
0-3	0.76	0.59	0.68	0.92	5.06**	18.97***	9.89***	1.00	2.09	0.03	
4-6	0.57	0.89	11.11***	1.99	0.05	2.36	0.38	0.84	0.98	3.74*	
7-12	8.24***	1.46	0.01	5.38**	4.04**	3.48*	3.33*	4.49**	6.52**	40.87***	
13-18	0.56	0.10	8.76***	7.95***	0.03	7.09***	0.11	3.66*	2.33	28.15***	
19+	17.29***	1.87	0.36	1.67	5.90**	32.93***	34.58***	22.86***	53.09***	17.88***	
F-test: ($\theta_i = 1$)	88.73***	237.06***	445.28***	6.15**	154.87***	1039.63***	2976.43***	475.88***	3752.15***	78.06***	
Effective weighted household size	3.65	3.78	3.77	3.81	3.91	4.18	3.70	4.04	5.05	3.72	3.92
Per capita consumption	174.38	53.31	180.09	162.76	120.49	178.65	247.41	159.28	295.46	843.52	2415.35
Per unit consumption	195.61	57.76	196.14	174.64	125.82	174.27	273.55	161.04	237.86	919.60	2516.42

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 9.
[d] The value of θ_i is 1 for spices but the hypothesis testing is not insignificant. It can be the case that here θ_i is having diseconomies of scale it would be more prominent if three decimal place is taken.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.13: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the lower income group of the rural sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.55	0.82	1.70	1.13	1.02	1.08	1.10	0.92	2.03	0.86	0.98
f0003	0.57	0.85	1.58	1.01	0.98	0.72	1.15	0.99	2.28	0.87	0.96
m0406	0.80	1.19	1.10	1.19	1.10	0.61	1.30	1.07	2.06	0.93	1.03
f0406	0.77	1.24	1.12	1.03	0.98	0.64	1.56	1.02	1.58	0.84	0.99
m0712	0.85	0.84	0.85	0.85	0.91	0.57	1.40	0.94	1.83	0.79	0.91
f0712	0.80	0.86	0.81	0.86	0.99	0.67	1.64	0.86	1.79	0.69	0.91
m1318	0.90	1.00	0.76	0.92	0.97	0.77	0.82	1.03	2.06	0.83	0.92
f1318	0.91	1.07	0.80	1.18	1.12	0.68	0.94	1.11	3.09	0.87	1.02
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.92	0.99	0.42	0.90	1.10	0.92	0.71	1.15	1.63	0.92	0.92
θ_i (economies of scale)	0.96	0.95	1.03	0.98	0.96	1.01	0.91	0.97	0.80	0.99	0.96
F-test: male vs female in different age group											
0-3	0.08	0.14	0.56	5.29**	0.46	5.54**	0.19	0.98	0.73	0.03	
4-6	0.24	0.39	0.02	7.81***	2.66	0.05	5.71**	0.61	2.57	3.02*	
7-12	1.59	0.17	0.14	0.08	2.46	0.81	9.54***	2.19	0.04	8.8***	
13-18	0.05	1.19	0.08	36.69***	6.71***	0.63	1.81	2.09	20.06***	1.05	
19+	2.53	0.01	19.16***	5.41**	3.37*	0.4	11.04***	6.57***	6.95***	3.94**	
F-test: ($\theta_i = 1$)	244.03***	310.51***	34.05***	124.66***	244.51***	2.38	608.15***	219.67***	1248.29***	37.45***	
Effective weighted household size	4.74	5.29	4.60	5.23	5.54	4.53	5.60	5.57	9.00	4.82	5.16
Per capita consumption	128.38	29.16	43.74	85.52	69.28	41.43	80.52	60.64	46.10	175.84	760.63
Per unit consumption	146.09	29.76	51.31	88.22	67.39	49.35	77.60	58.75	27.59	196.60	794.53

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 5.
[d] The value of θ_i is greater than 1 for milk and also the hypothesis testing is significant. So here we have diseconomies of scale.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.14: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the lower income group of the urban sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.56	0.69	1.54	0.74	0.72	0.57	1.10	0.82	1.00	0.83	0.84
f0003	0.61	0.81	1.38	0.86	0.90	1.09	1.41	0.90	0.86	0.91	0.95
m0406	0.69	0.81	0.70	0.74	0.74	0.48	1.04	0.74	0.93	0.65	0.73
f0406	0.68	0.81	0.57	0.73	0.79	0.82	1.08	0.70	0.96	0.71	0.76
m0712	0.85	0.83	0.55	0.74	0.84	0.58	1.22	0.79	1.13	0.58	0.77
f0712	0.70	0.86	0.66	0.79	0.75	0.67	1.20	0.85	1.31	0.73	0.81
m1318	0.83	0.94	0.59	0.79	0.87	0.56	0.86	0.87	1.30	0.65	0.78
f1318	0.74	0.93	0.87	0.87	0.87	0.54	0.85	0.89	1.55	0.72	0.83
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.79	0.75	0.59	0.81	0.90	0.97	0.68	0.83	0.92	0.84	0.81
θ_i (economies of scale)	0.96	0.95	0.97	0.99	0.97	0.96	0.90	0.96	0.86	1.00	0.96
F-test: male vs female in different age group											
0-3	0.59	2.82*	1.83	5.68**	7.47***	15.12***	7.57***	1.29	0.63	2.49	
4-6	0.07	0.00	1.07	0.02	0.63	6.31***	0.11	0.30	0.03	1.50	
7-12	13.33***	0.31	1.48	2.19	3.49*	0.79	0.09	1.13	2.11	19.10***	
13-18	4.18**	0.09	9.71***	4.29**	0.00	0.03	0.02	0.08	3.85**	4.51**	
19+	34.18***	28.30***	28.45***	34.36***	4.91**	0.11	19.78***	14.18***	0.47	25.78***	
F-test: ($\theta_i = 1$)	226.8***	291.97***	72.34***	63.93***	199.95***	145.68***	767.55***	246.45***	1080.39***	9.76***	
Effective weighted household size	4.40	4.61	4.24	4.55	4.76	4.46	5.02	4.68	5.63	4.41	4.59
Per capita consumption	139.02	38.05	81.91	115.65	81.07	76.31	108.48	80.92	81.73	291.85	1095.00
Per unit consumption	168.43	44.00	103.15	135.42	90.73	91.41	115.44	92.09	77.18	352.45	1270.19

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 6.
[d] The value of θ_i is 1 for other misc. non-food exp. but the hypothesis testing is not insignificant. It can be the case that here θ_i is having diseconomies of scale. It would be more prominent if three decimal place is taken.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.15: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the middle income group of the rural sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.41	0.95	1.87	1.04	0.82	0.60	1.04	0.87	1.92	0.95	0.95
f0003	0.48	0.80	1.64	0.93	0.83	0.84	1.18	0.90	2.00	0.91	0.96
m0406	0.70	0.95	1.32	0.98	0.91	0.87	1.25	0.95	1.67	0.83	0.96
f0406	0.70	0.78	1.10	0.95	0.94	0.95	1.33	0.80	1.44	0.77	0.92
m0712	0.79	0.87	0.99	0.91	0.88	0.70	1.49	0.93	1.91	0.75	0.94
f0712	0.76	0.77	0.86	0.91	0.84	0.95	1.65	0.91	2.03	0.83	0.98
m1318	0.80	0.93	1.17	0.98	0.97	0.66	0.91	1.06	2.46	0.80	0.95
f1318	0.82	0.91	1.01	1.03	0.95	0.99	0.85	1.13	2.87	0.85	1.00
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.90	1.03	0.79	1.03	0.87	1.08	0.76	1.09	2.08	1.03	1.00
θ_i (economies of scale)	0.95	0.93	0.90	0.97	0.95	0.90	0.88	0.96	0.74	0.98	0.93
F-test: male vs female in different age group											
0-3	2.14	5.69**	3.67*	6.18***	0.04	5.08**	2.80*	0.28	0.15	0.85	
4-6	0.01	7.28***	3.17*	0.62	0.44	0.55	0.86	8.26***	1.00	2.24	
7-12	1.04	5.23**	2.56	0.01	1.41	11.15***	6.43***	0.22	0.54	7.21***	
13-18	0.52	0.21	3.88**	2.76*	0.33	18.09***	0.84	2.90*	6.28***	2.32	
19+	11.14***	0.36	7.39***	1.12	14.82***	1.16	17.37***	5.84**	49.39***	1.45	
F-test: ($\theta_i = 1$)	571.63***	731.43***	622.14***	399.45***	619.69***	637.55***	1278.04***	495.95***	2542.92***	178.17***	
Effective weighted household size	4.09	4.57	4.79	4.74	4.39	4.52	4.86	4.83	8.48	4.47	4.69
Per capita consumption	152.08	40.05	98.41	125.04	90.93	92.58	134.24	91.65	93.38	293.64	1212.01
Per unit consumption	177.59	41.82	97.72	125.78	98.77	97.78	131.69	90.52	52.56	313.05	1232.05

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 4.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.16: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the middle income group of the urban sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.56	0.84	1.36	0.83	0.71	1.08	1.16	0.84	1.14	1.04	0.98
f0003	0.48	0.74	1.42	0.79	0.69	1.14	1.29	0.90	1.17	1.03	1.00
m0406	0.69	0.90	1.09	0.82	0.90	0.91	1.10	1.00	1.54	0.93	0.97
f0406	0.63	0.80	1.10	0.82	0.84	1.06	1.12	1.04	1.62	0.82	0.95
m0712	0.82	0.81	0.94	0.81	0.79	0.99	1.10	1.02	2.20	0.74	0.95
f0712	0.81	0.86	0.83	0.82	0.98	1.07	1.12	0.99	2.14	0.83	0.99
m1318	0.90	0.95	0.71	0.86	1.06	1.04	0.91	1.06	2.40	0.69	0.95
f1318	0.90	0.88	0.78	0.87	0.96	1.28	0.97	1.10	2.19	0.78	0.99
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.87	0.97	0.99	1.03	1.04	1.30	0.78	1.16	1.66	1.02	1.05
θ_i (economies of scale)	0.97	0.94	0.88	0.97	0.95	0.86	0.84	0.95	0.81	0.99	0.92
F-test: male vs female in different age group											
0-3	1.96	1.88	0.32	0.77	0.06	0.16	1.44	0.59	0.01	0.02	
4-6	1.13	1.78	0.00	0.00	0.57	1.00	0.03	0.32	0.12	3.43*	
7-12	0.16	1.23	2.10	0.14	11.58***	0.74	0.06	0.25	0.12	5.97**	
13-18	0.00	2.03	0.75	0.14	3.06**	6.62***	0.56	0.76	2.06	6.20***	
19+	22.02***	0.71	0.02	1.38	1.03	17.63***	15.47***	16.14***	33.41***	0.37	
F-test: ($\theta_i = 1$)	190.9***	358.11***	719.72***	141.72***	381.99***	828.65***	992.38***	329.55***	1159.37***	42.81***	
Effective weighted household size	3.75	4.01	4.17	4.05	4.15	4.77	4.05	4.49	6.54	4.02	4.28
Per capita consumption	174.12	53.15	172.77	165.13	119.89	161.15	204.78	142.65	219.18	647.13	2059.94
Per unit consumption	197.78	56.38	176.22	173.11	122.61	143.54	214.78	135.24	141.94	684.10	2047.35

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 4.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.17: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the higher income group of the rural sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.47	0.96	1.25	0.87	0.83	1.02	1.12	0.88	1.52	1.19	1.06
f0003	0.57	0.82	1.13	0.83	0.80	1.10	1.25	0.83	1.62	1.20	1.08
m0406	0.70	0.76	0.98	0.88	0.87	0.83	1.32	0.90	0.93	0.93	0.94
f0406	0.79	0.69	0.94	0.99	1.00	1.03	1.30	1.06	1.64	0.88	1.03
m0712	0.78	0.82	0.99	0.88	1.01	0.80	1.06	1.02	1.48	0.85	0.95
f0712	0.87	0.85	0.74	0.91	0.97	1.10	1.18	1.01	1.61	0.92	1.01
m1318	0.93	0.92	1.01	0.91	1.02	0.79	0.82	1.13	1.84	0.76	0.95
f1318	0.99	0.70	0.76	0.85	1.07	1.06	0.83	1.19	2.16	0.89	1.02
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	0.93	1.17	1.04	1.07	0.99	1.03	0.60	1.19	1.87	1.16	1.09
θ_i (economies of scale)	0.96	0.93	0.85	0.97	0.94	0.86	0.85	0.96	0.78	0.98	0.91
F-test: male vs female in different age group											
0-3	4.17**	3.96**	0.90	0.73	0.14	0.50	2.27	0.50	0.23	0.01	
4-6	3.04*	1.01	0.09	4.00**	4.15**	3.09*	0.04	5.30**	9.60***	0.49	
7-12	7.12***	0.40	8.74***	0.85	0.63	17.01***	3.59*	0.06	0.69	1.85	
13-18	3.10*	23.63***	10.48***	3.26*	1.55	17.38***	0.08	1.63	5.39**	8.63***	
19+	8.38***	21.51***	0.39	7.50***	0.15	0.25	82.98***	29.40***	61.23***	19.86***	
F-test: ($\theta_i = 1$)	518.87***	771.95***	1973.65***	433.66***	856.16***	1346.24***	1803.69***	469.83***	3147.21***	238.71***	
Effective weighted household size	3.55	3.87	3.85	3.81	3.82	3.82	3.37	4.16	5.83	3.95	3.97
Per capita consumption	179.61	54.76	204.07	174.07	122.39	187.00	248.19	147.59	284.25	635.94	2237.86
Per unit consumption	195.47	54.51	204.86	175.73	123.68	188.63	284.56	136.95	188.04	620.57	2176.54

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 5.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

Table 5.18: 68^{th} round item-wise estimate of consumer unit scales & economies of scale for the higher income group of the urban sector.

Variables	Cereals	Pulses	Milk	Spices,salt, sugar & oil	Veg.	Egg,fish, meat,fruits	Other misc. food exp.	Clothing, bedding & footwear	Med & edu.	Other misc. non-food	Income scales
m0003	0.40	0.48	1.26	0.73	0.63	1.10	0.98	0.88	1.04	1.31	1.07
f0003	0.64	0.62	1.26	0.80	0.86	1.82	1.31	1.01	0.55	1.31	1.14
m0406	0.64	0.75	1.47	0.81	0.98	1.12	1.07	1.13	0.95	1.17	1.09
f0406	0.74	0.86	0.99	0.78	1.24	1.25	1.23	1.40	0.82	1.38	1.18
m0712	0.92	0.90	1.08	0.90	0.94	1.24	0.95	1.19	1.48	0.81	1.00
f0712	0.90	0.79	0.60	0.90	1.02	1.39	1.03	1.23	1.56	0.97	1.06
m1318	0.99	0.83	0.94	0.90	0.99	1.06	0.97	1.14	2.13	0.80	1.05
f1318	1.02	0.83	0.71	0.97	1.12	1.33	0.81	1.16	2.16	0.96	1.11
m19+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f19+	1.06	1.10	1.15	1.12	1.07	1.24	0.62	1.11	1.57	1.10	1.11
θ_i (economies of scale)	0.98	0.96	0.91	0.98	0.94	0.86	0.79	0.95	0.80	0.99	0.91
F-test: male vs female in different age group											
0-3	12.27***	2.36	0.00	0.88	5.18**	18.78***	6.86***	1.54	3.48*	0.00	
4-6	1.89	1.42	13.40***	0.24	6.42**	0.57	1.57	6.38**	0.22	5.61**	
7-12	0.06	3.24*	32.00***	0.00	1.60	1.81	0.84	0.34	0.19	7.41***	
13-18	0.93	0.01	9.55***	3.25*	5.06**	7.43***	4.61**	0.07	0.05	10.66***	
19+	5.45**	7.99***	9.15***	20.63***	3.28*	14.55***	66.84***	8.19***	34.35***	10.39***	
F-test: ($\theta_i = 1$)	65.09***	157.41***	449.55***	24.71***	281.95***	668.37***	918.96***	234.09***	1220.82***	29.8***	
Effective weighted household size	3.14	3.13	3.32	3.22	3.25	3.66	2.75	3.43	4.29	3.28	3.36
Per capita consumption	212.12	69.65	294.20	209.21	163.01	311.53	454.44	265.60	629.06	1703.85	4312.66
Per unit consumption	215.30	70.86	281.77	206.57	159.70	271.19	521.41	247.19	462.13	1649.68	4083.92

[a] ***,**,* indicate significance at the level 1%, 5 % and 10%, respectively.
[b] Coefficients are rounded to two decimal place.
[c] The number of iterations used to produce the results in this table is 6.
Source: Author's calculation based on 2011-12 Type-2 Consumer Expenditure Household Survey, collected by NSS.

6 Conclusion

6.1 Introduction

In this thesis, we have studied the gender difference in consumption expenditure in case of various commodities using National Sample Survey data on household consumer expenditure collected throughout India. This National Sample Survey data provides information on the consumer expenditure at the household level. Therefore, to get information at the individual-level, two very important techniques are used. One is 'Regression Decomposition Technique'(Pal and Bharati (2010)) and the other is 'Semi-parametric Engel Curve Approach'(Gong et al. (2005)). Regression Decomposition Technique is used to investigate the gender difference in individual-level food consumption expenditure of the household members given the total food expenditure of the household. The idea is based on the identity that the total food expenditure of a household is the sum of food expenditure of individual members in the household. The regressors are the number of people in various age-sex groups within the household. The total food expenditure is regressed on these variables or regressors to produce the average food consumption expenditure incurred by each member in the respective age-sex group. While dealing with the above method we come across the problem of heteroscedasticity in the model. In the presence of heteroscedasticity, one has to estimate sigma square (variance of the error term) first and then go for Feasible Generalized Least Square estimate of the coefficient in the regression function. There can be various assumptions regarding the functional form of this sigma square. We assumed four types of functions for sigma square and then compared the value of the coefficients for the original regression model. We found that Principal Component Analysis(PCA) is the most appropriate method to deal with the error variance when it is unknown. The above procedure has been applied to the Indian household level data of the 61^{st} and 66^{th} round Consumer Expenditure Survey of National Sample Survey Organization, Government of India to see whether there is any gender difference in food expenditure for the different age-sex group.

The second method which has been used is the semiparametric form of the Engel Curve approach. The Engel curve approach has been widely used in the literature by many authors like Deaton (1989), Fuwa (2014), Fuwa et al. (2006), Gibson (1997), Gibson and Rozelle (2004), Haddad and Reardon (1993), Himaz (2010), Lee (2008), Subramaniam (1996), Subramanian and Deaton (1991) etc. These authors have used the parametric form of the Engel curve. In our analysis, the semiparametric form of the Engel curve is used to detect gender difference in consumption. Working-Leser form of the Engel curves has long been discarded. Nowadays usually a log-quadratic model is used in the parametric set-up (Banks et al. (1997)). However, nonparametric Engel curves are more general and can capture more complex shapes. So, nonparametric smoothing procedure for estimating regression functions has received attention during the last two decades. In case of the nonparametric form, there exists a dimensionality issue which restricts the Engel curve analysis for further expansion. In the nonparametric model, as the number of parameters increases the requirement for larger datasets also rises to cope up with the dimensionality issue. Most of the cross-sectional studies of gender bias cannot have such large datasets. So, as an easy path, economists go for a semiparametric form of Engel curve, where income has an unknown relation with the expenditure share of any commodity but have a linear relation with the demographic variables. Authors who have used the semiparametric form of the Engel curve to detect gender bias in household consumption are Attfield and Bhalotra (1998), Blundell et al. (1998), Gong et al. (2005) etc. We have applied the same approach for food, milk(as a child product) and, intoxicant (a typical adult good) to see whether there exists any gender bias in favour of male using the idea of Deaton's approach.

Lastly, as discussed in Chapter-1 in Section 1.1, we need to know that if there is gender bias in the consumption expenditure and different people consume differently then how are we going to get the exact size of household in terms of adult equivalent ratios. Adult equivalent is necessary if one is concerned about the individual needs and not the family as a whole. In that sense per capita level does not give the true picture of the resource allocation among the households. So we try to find the "consumer unit scale" for different commodity groups using a modified version of Singh (1972) approach. After some modifications in the Engel curve approach as done by Singh (1972) and Singh and Nagar (1973), we have found an iterative technique for the estimation of parameters which enables us to get the 'specific' and 'income' consumer unit scales. Similarly, an iterative method has been used to estimate the economies of scale, where economies of scale parameter enter into the Engel function exponentially as suggested by Houthakker and Prais (1971). Using these techniques, respective adult equivalence consumer unit scales are found for the 10 age-sex groups of the household and also for the 10 different commodities groups. Economies of scale is also estimated for the 10 different commodities.

The last chapter is organised as follows. A summary of the major findings of the entire work is presented in Section 6.2, and the last section i.e., Section 6.3 presents a few ideas for further work in this area.

6.2 Major Findings

In chapter 2, a brief discussion of the data used has been given starting from the survey design to the expenditure pattern for different items of the commodities. These data are the only available data at the national level which gives details on the consumption of the items of commodities by the household.

In this chapter, it is seen that the rural sector has contracted and the urban sector expanded in terms of the number of households over the years. Also, from the 'sex-ratio' pattern followed during the last decade, our data shows that it has actually declined over the years. In the rural sector, this ratio has declined but in the urban sector, it has increased in the year 2009-10. For 0-3 years, this ratio has declined over the years. Thus, showing gender bias in favour of male children. The overall share of food expenditure has declined over the years and a similar result is seen for the cereals.

In chapter 3, which is based on 'Regression Decomposition Technique', by which it is possible to find out the age-sex composition wise food consumption expenditure from the total food expenditure of the household. The parameters are estimated for the NSS 61^{st} and 66^{th} round data. We have found two important results via this analysis. Firstly, PCA method is superior to other models for estimating sigma square of the error term. Secondly, evidence of significant gender difference in food expenditure is found for the adult groups. In 2004-5 i.e. using NSS 61^{st} round, we found significant gender difference in food expenditure for the children age group as well.

In chapter 4, we have analysed expenditure pattern in rural India using NSS 68^{th} round Type 1 consumer expenditure data collected during July 2011–June 2012. We estimated the semiparametric form of Engel curve considering with and without the endogeneity of log per capita total expenditure. To solve the endogeneity issue two methods have been used; 2SLS and CFA. Both the methods give more or less similar result. We found evidence of gender differentials in food and intoxicants for the adults favouring the male, confirming the results of Attfield and Bhalotra (1998). For milk, we found evidence of gender bias for approximately all the age groups in household size 4 and 5. We also find evidence of gender bias in favour of the female in case of milk and food for household size 3, 4 and 5 respectively. To compare semiparametric specification with constant, linear and quadratic parametric specifications we implement the recently developed specification test by Hardle and Mammen (1993) test generated by 100 replications of the wild bootstrap. This test is accepted for food as well as for intoxicant for few cases, which is a contradictory result obtained from Gong et al. (2005). So linear and quadratic Engel curves are suitable for few commodities and households size. The overall test for equal sex difference across age and no gender bias is rejected which imply the existence of gender bias for each household sizes.

In chapter 5, a modified iterative procedure based on Singh (1972) and Singh and Nagar (1973) approach has been proposed, assuming the Engel relation is nonparametric. This procedure is used to find the adult equivalence scales for 10 commodity groups and income. Also, the economies of scale are estimated after the estimation of consumer unit scales. The data used for the analysis is NSS 68^{th} round type 2 data as it has different recall period which is more appropriate according to the Expert group if the analysis is done for different commodity groups. The estimation of consumer unit scale is done for different expenditure groups. These expenditure groups are defined based on the MPCE decile groups. Therefore, three types of expenditure groups are defined the low, middle and high-income group. Also, the analysis is done for the overall expenditure group. In case of the overall expenditure group, the result of the analysis shows that for the rural sector gender bias favours male more than female for most of the commodity groups and for the income scale as well. However, for the urban sector, the reverse is true. Also, mostly weights favouring male is observed for 0-6 years and this bias changes its pattern for the 7 and above age-groups. When we separate out the expenditure groups according to the income i.e. low, middle, and high income/expenditure groups, the result changes. For the low-income group, the gender bias pattern is same as discussed above. But for the middle and higher income group, gender bias favours female more than male. We found positive values for the consumer unit scale and the economies of scale unlike Singh and Nagar (1973) for all the commodities group as well as the income. The weighted equivalence scales are highest for the adults and lowest for the children (0-3 years) for most of the groups with few exceptions like milk, other misc. food expenditure and medical expenses. Also, economies of scales are there for approximately every commodity groups.

6.3 Few Ideas for Further Research

In this thesis, we have studied gender difference in the consumption expenditure for both the rural and the urban sectors of the country and for the country as a whole. The gender difference was also seen for different income/total expenditure groups as well. This analysis can be extended to each state. It will give a clearer picture of the exact pockets where gender biased should be focused on various policies and awareness programs.

Instead of commodity groups, as it is taken in this thesis, similar work may be carried out for each commodity or finer commodity groups. It is likely that in this case, one would face the problem of "zero observation", i.e. for a given item, there may be too many households not consuming the item. Since we have taken only the broad commodity groups, we avoided that problem. We faced only a very few household not consuming the broad commodity groups. Those households were deleted from the data for the Engel curve analysis.

To consider the problem of excess zeros, one can go for the double-hurdle model, originally proposed by Cragg (1971), and do the analysis for each item of consumer expenditure to see if gender bias exist or not for these items. The double-hurdle model assumes two separate hurdles must be passed before a positive level of consumption can be observed. The first hurdle involves the decision of whether to go to market (if it is related to purchase of anything). This is known as participation decision. The second hurdle concerns the level of quantity to purchase (for any item). This is called consumption decision. Now, zero consumption of any item can be observed in two situations. Suppose the individual doesn't want to participate. And the other, if the consumption decision is not positive, for example, it may happen that after going to the market one may not purchase the item, in that case also, the consumption expenditure on the item in question will be zero.

Bibliography

Arnold, F., Choe, M. K., and Roy, T. K. (1998). Son preference, the family-building process and child mortality in India. *Population Studies*, 52(3):301–315.

Arnold, F., Kishor, S., and Roy, T. K. (2002). Sex-selective abortions in India. *Population and Development Review*, 28(4):759–785.

Atkinson, A. B., Gomulka, J., and Stern, N. H. (1990). Spending on alcohol: Evidence from the family expenditure survey 1970-1983. *The Economic Journal*, 100(402):808–827.

Attfield, C. and Bhalotra, S. R. (1998). Intrahousehold resource allocation in rural Pakistan: A semi-parametric analysis. *Journal of Applied Econometrics*, 13(5):463–480.

Banks, J., Blundell, R., and Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics*, 79(4):527–539.

Basu, A., Roy, S. K., Mukhopadhyay, B., Bharati, P., Gupta, R., and Majumder, P. P. (1986). Sex bias in intrahousehold food distribution: Roles of ethnicity and socioeconomic characteristics. *Current Anthropology*, 27(5):536–539.

Behrman, J. R. (1988). Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion? *Oxford Economic Papers*, 40(1):32–54.

Blundell, R., Duncan, A., and Pendakur, K. (1998). Semiparametric estimation and consumer demand. *Journal of Applied Econometrics*, 13(5):435–461.

Bojer, H. (1977). The effect on consumption of household size and composition. *European Economic Review*, 9(2):169–193.

Case, A., Deaton, A., et al. (2002). *Consumption, Health, Gender and Poverty.* World Bank Policy Research Working Paper No. 3020.

Chen, L. C., Huq, E., and D'Souza, S. (1981). Sex bias in the family allocation of food and health care in rural Bangladesh. *Population and Development Review*, 7(1):55–70.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, 39(5):829–844.

Das Gupta, M. and Mari Bhat, P. (1997). Fertility decline and increased manifestation of sex bias in India. *Population Studies*, 51(3):307–315.

Deaton, A. (1989). Looking for boy-girl discrimination in household expenditure data. *The World Bank Economic Review*, 3(1):1–15.

Deaton, A. (1997). *The analysis of household surveys: A microeconometric approach to development policy.* World Bank Publications.

Deaton, A. and Muellbauer, J. (1980). An almost ideal demand system. *The American Economic Review*, 70(3):312–326.

Deininger, K., Jin, S., and Nagarajan, H. (2013). Wage discrimination in India's informal labor markets: Exploring the impact of caste and gender. *Review of Development Economics*, 17(1):130–147.

Delgado, M. A. and Miles, D. (1997). Household characteristics and consumption behaviour: A nonparametric approach. *Empirical Economics*, 22(3):409–429.

Demoussis, M. and Mihalopoulos, V. (2001). Adult equivalent scales revisited. *Journal of Agricultural and Applied Economics*, 33(01):135–146.

Emerson, P. M. and Souza, A. P. (2007). Child labor, school attendance, and intrahousehold gender bias in Brazil. *The World Bank Economic Review*, 21(2):301–316.

Engel, E. (1857). *Die Productions- und Consumptionsverhaeltnisse des Koenigsreichs Sachsen, Die Lebenskosten belgischer Arbeiter-Familien frúher und jetzt.* Number 8 und 9. Reprinted in the Appendix of Engel (1895).

Engel, E. (1895). Die lebenskosten belgischer arbeiter-familien frúher und jetzt. *International Statistical Institute Bulletin*, 9:1–74.

Ferré, L. (1995). Selection of components in principal component analysis: A comparison of methods. *Computational Statistics & Data Analysis*, 19(6):669–682.

Forsyth, F. G. (1960). The relationship between family size and family expenditure. *Journal of the Royal Statistical Society. Series A (General)*, 123(4):367–397.

Fuwa, N. (2014). Pro-girl bias in intra-household allocation in the rural Philippines: Revisiting the "adult goods" approach. *Review of Development Economics*, 18(4):727–740.

Fuwa, N., Ito, S., Kubo, K., Kurosaki, T., and Sawada, Y. (2006). Gender discrimination, intrahousehold resource allocation, and importance of spouses' fathers: Evidence on household expenditure from rural India. *The Developing Economies*, 44(4):398–439.

Gibson, J. (1997). Testing for boy-girl discrimination with household expenditure data: results for Papua New Guinea. *Applied Economics Letters*, 4(10):643–646.

Gibson, J. and Rozelle, S. (2004). Is it better to be a boy? A disaggregated outlay equivalent analysis of gender bias in Papua New Guinea. *Journal of Development Studies*, 40(4):115–136.

Goedhart, T., Halberstadt, V., Kapteyn, A., and Van Praag, B. (1977). The poverty line: Concept and measurement. *Journal of Human Resources*, 12(4):503–520.

Gong, X., Van Soest, A., and Zhang, P. (2005). The effects of the gender of children on expenditure patterns in rural China: A semiparametric analysis. *Journal of Applied Econometrics*, 20(4):509–527.

Gozalo, P. L. (1997). Nonparametric bootstrap analysis with applications to demographic effects in demand functions. *Journal of Econometrics*, 81(2):357–393.

Griffiths, P., Matthews, Z., and Hinde, A. (2000). Understanding the sex ratio in India: A simulation approach. *Demography*, 37(4):477–488.

Gupta, M. D. (1987). Selective discrimination against female children in rural Punjab, India. *Population and Development Review*, 13(1):77–100.

Gupta, M. D. (2005). Explaining Asia's "Missing Women": A new look at the data. *Population and Development Review*, 31(3):529–535.

Haddad, L. and Reardon, T. (1993). Gender bias in the allocation of resources within households in Burkina Faso: A disaggregated outlay equivalent analysis. *The Journal of Development Studies*, 29(2):260–276.

Haque, M. O. (2006). *Income elasticity and economic development: Methods and applications*, volume 42. Springer Science & Business Media.

Hårdle, W. and Jerison, M. (1991). Cross-sectional Engel curves over time. *Recherches Économiques de Louvain/Louvain Economic Review*, 57(4):391–431.

Hardle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, 21(4):1926–1947.

Harriss, B. (1989). Differential female mortality and health care in South Asia. *Journal of Social Studies*, (44):1–123.

Hausman, J. A., Newey, W. K., Ichimura, H., and Powell, J. L. (1991). Identification and estimation of polynomial errors-in-variables models. *Journal of Econometrics*, 50(3):273–295.

Hausman, J. A., Newey, W. K., and Powell, J. L. (1995). Nonlinear errors in variables estimation of some Engel curves. *Journal of Econometrics*, 65(1):205–233.

Hildreth, C. and Houck, J. P. (1968). Some estimators for a linear model with random coefficients. *Journal of the American Statistical Association*, 63(322):584–595.

Himaz, R. (2010). Intrahousehold allocation of education expenditure: The case of Sri Lanka. *Economic Development and Cultural Change*, 58(2):231–258.

Holly, A. and Sargan, J. D. (1982). *Testing for exogeneity in a limited information framework*. Cahiers de Recherches Economiques, No. 8204 Université de Lausanne, Ecole des hautes études commerciales.

Houthakker, H. S. and Prais, S. J. (1971). *The Analysis of Family Budgets*, volume 4. CUP Archive.

Jackson, J. E. (2005). *A user's guide to principal components*, volume 587. John Wiley & Sons.

Jha, P., Kumar, R., Vasa, P., Dhingra, N., Thiruchelvam, D., and Moineddin, R. (2006). Low male-to-female sex ratio of children born in India: national survey of 1·1 million households. *The Lancet*, 367(9506):211–218.

Johnston, J. (1984). *Econometric Methods*, volume 3. McGraw Hill International Editions.

Jolliffe, I. T. (2002). *Principal component analysis*. Springer, New York.

Kakwani, N. C. (1977). On the estimation of consumer unit scales. *The Review of Economics and Statistics*, 59(4):507–510.

Kapteyn, A. and Van Praag, B. (1978). A new approach to the construction of family equivalence scales. *European Economic Review*, 7(4):313–335.

Kemsley, W. (1952). Estimating individual expenditure from family totals. *Applied Statistics*, 1(3):192–201.

Kingdon, G. G. (2002). The gender gap in educational attainment in India: How much can be explained? *Journal of Development Studies*, 39(2):25–53.

Kingdon, G. G. (2005). Where has all the bias gone? Detecting gender bias in the intrahousehold allocation of educational expenditure. *Economic Development and Cultural Change*, 53(2):409–451.

Klasen, S. and Wink, C. (2003). "Missing Women": Revisiting the debate. *Feminist Economics*, 9(2-3):263–299.

Lancaster, G., Maitra, P., and Ray, R. (2008). Household expenditure patterns and gender bias: Evidence from selected Indian states. *Oxford Development Studies*, 36(2):133–157.

Lancaster, G., Ray, R., and Valenzuela, M. R. (1999). A cross-country study of equivalence scales and expenditure inequality on unit record household budget data. *Review of Income and Wealth*, 45(4):455–482.

Lawrence, F. R. and Hancock, G. R. (1999). Conditions affecting integrity of a factor solution under varying degrees of overextraction. *Review of Income and Wealth*, 45(4):455–482.

Lee, Y. F. D. (2008). Do families spend more on boys than on girls? Empirical evidence from rural China. *China Economic Review*, 19(1):80–100.

Leser, C. E. V. (1963). Forms of Engel functions. *Econometrica: Journal of the Econometric Society*, 31(4):694–703.

Lewbel, A. (1991). The rank of demand systems: Theory and nonparametric estimation. *Econometrica: Journal of the Econometric Society*, 59(3):711–730.

Lewbel, A. and Pendakur, K. (2006). *Equivalence scales entry for the New Palgrave Dictionary of Economics*. 2 edition.

Majumder, A. and Chakrabarty, M. (2010). Estimating equivalence scales through Engel curve analysis. *In: Basu B., Chakravarty S.R., Chakrabarti B.K., Gangopadhyay K. (eds) Econophysics and Economics of Games, Social Choices and Quantitative Techniques. New Economic Windows. Springer, Milano:241–251.*

Mason, A. D., Montenegro, C. E., and Khandker, S. R. (1999). *Can we say anything more about gender and poverty using household consumption data*. Poverty Reduction and Economic Management Network, The World Bank, processed.

Miller, B. D. (1987). *Female infanticide and child neglect in rural North India*. In: Scheper-Hughes N. (eds) Child Survival. Culture, Illness and Healing, volume 11. Springer, Dordrecht:95-112.

Miller, B. D. (1997). *The endangered sex: Neglect of female children in rural North India*. Oxford University Press.

Mishra, V., Roy, T. K., and Retherford, R. D. (2004). Sex differentials in childhood feeding, health care, and nutritional status in India. *Population and Development Review*, 30(2):269–295.

Mittal, S. (2007). What affects changes in cereal consumption? *Economic and Political Weekly*, 42(5):444–447.

Moser, C. (2012). *Gender planning and development: Theory, practice and training*. Routledge.

Muellbauer, J. (1975). Identification and consumer unit scales. *Econometrica: Journal of the Econometric Society*, 43(4):807–809.

Muellbauer, J. (1977). Testing the Barten model of household composition effects and the cost of children. *The Economic Journal*, 87(347):460–487.

Muellbauer, J. (1980). The estimation of the Prais-Houthakker model of equivalence scales. *Econometrica: Journal of the Econometric Society*, 48(1):153–176.

Mujumdar, N. A. and Kapila, U. (2006). *Indian agriculture in the new millennium: Changing perceptions and development policy*, volume 2. Academic Foundation.

Murthi, M. (1994). *7 Engel equivalence scales in Sri Lanka: Exactness, specification, measurement error*. In:The Measurement of Household Welfare:164.

Nelson, J. A. (1988). Household economies of scale in consumption: Theory and evidence. *Econometrica: Journal of the Econometric Society*, 56(6):1301–1314.

Newey, W. K. and Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578.

Nicholson, J. L. (1949). Variations in working class family expenditure. *Journal of the Royal Statistical Society. Series A (General)*, 112(4):359–418.

Pal, M. and Bharati, P. (2010). Estimating intra-household gender inequality of consumption: A regression-decomposition analysis. *Paper presented at the 2010 Joint World Conference on Social Work and Social Development: The Agenda, Hong Kong Convention and Exhibition Centre, Hong Kong, China, June 10-14.*

Pande, R. and Malhotra, A. (2006). *Son preference and daughter neglect in India: What happens to living girls?* International Centre for Research on Women.

Phipps, S. A. (1998). What is the income "cost of a child"? Exact equivalence scales for Canadian two-parent families. *Review of Economics and Statistics*, 80(1):157–164.

Robinson, P. M. (1988). Root-N-Consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, 56(4):931–954.

Sarangi, P. and Panda, B. K. (2011). Measurement of unit consumer scales with reference to household size, age, gender composition and occupation. *IUP Journal of Management Research*, 10(2):53.

Singh, B. (1972). On the determination of economies of scale in household consumption. *International Economic Review*, 13(2):257–270.

Singh, B. and Nagar, A. (1973). Determination of consumer unit scales. *Econometrica: Journal of the Econometric Society*, 41(2):347–355.

Strauss, J. and Thonas, D. (1990). *The shape of the calorie-expenditure curve*. Center Discussion paper, number 595.

Subramaniam, R. (1996). Gender-bias in India: The importance of household fixed-effects. *Oxford Economic Papers*, 48(2):280–299.

Subramanian, S. and Deaton, A. (1991). Gender effects in Indian consumption patterns. *Sarvekshana*, 14(4):1–12.

Subramanian, S. and Deaton, A. (1996). The demand for food and calories. *Journal of Political Economy*, 104(1):133–162.

Sydenstricker, E. and King, W. I. (1921). The measurement of the relative economic status of families. *Quarterly Publications of the American Statistical Association*, 17(135):842–857.

Valenzuela, M. and Rebecca, J. (1996). Engel scales for Australia, the Philippines and Thailand: A comparative analysis. *Australian Economic Review*, 29(2):189–198.

Van Praag, B. M. S. and Van der Sar, N. L. (1988). Household cost functions and equivalence scales. *Journal of Human Resources*, 23(2):193–210.

Verardi, V. and Debarsy, N. (2012). Robinson's square root of n consistent semiparametric regression estimator in stata. *Stata Journal*, 12(4):726–735.

Visaria, P. (1979). Demographic factors and the distribution-of income: Some issues. *Economic and Demographic Change: Issues for the 1980's. Liege: IUSSP*, pages 289–320.

Woodbury, R. M. (1944). Economic consumption scales and their uses. *Journal of the American Statistical Association*, 39(228):455–468.

Working, H. (1943). Statistical laws of family expenditure. *Journal of the American Statistical Association*, 38(221):43–56.

Yatchew, A. (1997). An elementary estimator of the partial linear model. *Economics Letters*, 57(2):135–143.

Yen, S. T. and Jones, A. M. (1997). Household consumption of cheese: An inverse hyperbolic sine double-hurdle model with dependent errors. *American Journal of Agricultural Economics*, 79(1):246–251.

Zimmermann, L. (2012). Reconsidering gender bias in intrahousehold allocation in India. *Journal of Development Studies*, 48(1):151–163.