

UNIVERSITY OF BIRMINGHAM

HOUSEHOLD SIZE, ENERGY USE AND THE IMPACT OF CLIMATE ON HOUSEHOLDS

by

YUHANG QU

A thesis submitted to the University of Birmingham for the Degree of

DOCTOR OF PHILOSOPHY

Department of Economics
Birmingham Business School
College of Social Sciences
University of Birmingham

September 2023

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This thesis focuses on studying household equivalence scales, economies of scale in household energy use, as well as the amenity value of climate to households based on household level consumption expenditure data obtained from the nationally representative household surveys in China and India, the two most populous countries in the world. Following a brief Introduction in chapter one, the Quadratic Almost Ideal Demand System model is employed to estimate household equivalence scales in China in chapter two. The results suggest that the household equivalence scales are different between urban and rural households, and that there exist significant economies of scale in Chinese households.

In chapter three, I continue with the investigation of economies of scale but switch attention from China to India and focus on energy consumed by Indian households. In order to do this, two different but not unrelated concepts are introduced for analysis. The empirical results indicate that there exist significant economies of scale in household energy expenditure in India. By estimating the elasticity of per capita energy use with respect to household size for different types of energy, I find that the economies of scale for modern and cleaner energy (i.e., electricity and LPG) are greater than these for traditional and transition energy. In particular, I find that there are diseconomies of scale in households' use of dung cake.

In chapter four, using the same dataset as used in chapter three, I focus on estimating the amenity value of climate to Indian households by incorporating climate variables into a popular system of demand, namely the Almost Ideal Demand System model. The results reveal that the climate plays a significant role in determining households' cost of living in India. I find higher precipitation is an

amenity, while heating and cooling degree days, as well as relative humidity are disamenities to Indian households. Chapter five summarizes the main findings of this thesis and also discuss the limitations as well as possible future work.

Acknowledgements

As my 4-year PhD journey comes to an end, I would like to take this precious opportunity to express my sincere gratitude to all the people who have supported me, without whom I would not have been able to complete my journey.

First and foremost, I would like to express my deep gratitude to my supervisors: Prof. David Maddison, Prof. Robert Elliott and Prof. Eric Strobl. Without their continuous support, insight and patience, this thesis would not have been completed. Special thanks to my lead supervisor, Prof. David Maddison, who has offered invaluable guidance and mentorship throughout my PhD journey. He not only extended academic assistance but also showed genuine care for my well-being in particular during the times of the pandemic. It is from him that I learnt how to deliver a good presentation and the true attitude toward research. I would also like to deeply thank my co-supervisors, Prof. Robert Elliott and Prof. Eric Strobl, for their encouragement and invaluable suggestions on my research. Without the climate data shared by Prof. Eric Strobl, I would not be able to complete my thesis on time.

I am indebted to the University of Birmingham for awarding me the Global Challenges Scholarship, without which my PhD would not have been possible. Meanwhile, I am equally grateful to the Department of Economics, and Birmingham Business School for their funding and support. My journey was further eased by the great support from the PGR administrative staff.

I would like to thank the chair of my Viva, Prof. Anthony Heyes for hosting my PhD Viva, as well as my internal examiner, Prof. Mtthew Cole, and my external examiner, Prof. Zhifu Mi. Their comments and suggestions are of great help to my thesis.

I would also like to thank my friends and PhD classmates, in particular, Dr. Jing Zhang, Dr. Zitai Chen, and Dr. Xiaoyu Qin. I cherish the time we spent together which made my PhD journey fantastic and memorable.

Last, but not least, I reserve my deepest gratitude for my family. To my parents and my sister, thank you for your endless support and encouragement. My love and gratitude for them can hardly be put into words. A special mention to my little niece, who recently celebrated her first birthday. Her presence has brought immeasurable joy to the family.

Table of Contents

Chapter 1 Introduction 1

Chapter 2 Estimating household equivalence scales for China: An application of the QUAIDS model 8

2.1 Introduction 8

2.2 Literature Review 12

2.2.1 Household equivalence scale 13

2.2.2 Expenditure-based approach to estimate household equivalence scales 15

2.2.3 A review of household equivalence scale in China 21

2.3 Methodology 23

2.3.1 The QUAIDS Model 24

2.3.2 Incorporating household demographics into the demand system 26

2.3.3 Price and Expenditure Elasticities 28

2.3.4 Engel's method 29

2.4 Data 30

2.4.1 China Household Finance Survey (CHFS) 30

2.4.2 Pricing data 34

2.4.3 Descriptive Statistics 35

2.5 Results 38

2.5.1 Estimation results for simple QUAIDS model 38

2.5.2 Estimation results for household equivalence scales 41

2.5.3 Estimation results from Engel's method 47

2.6 Discussion 49

2.7 Conclusion 52

Chapter 3 Household size, economies of scale in household energy use in India 54

3.1 Introduction 55

3.2 Literature review 59

3.2.1 Determinants of household energy consumption in India 60

3.2.2 Economies of scale in household energy consumption 66

3.2.3 Empirical applications of Kakwani-Son's method 72

3.2.4 Applications of double-hurdle model in household energy consumption 74

3.2.5 Literature discussion 75

3.3 Methodology 78
3.3.1 Kakwani-Son's method 78
3.3.2 The double-hurdle model 83
3.4 Data 85
3.4.1 Data source 85
3.4.2 Descriptive statistics 88
3.5 Estimation results from Kawani-Son's model 90
3.6 Estimation results from the double-hurdle model 93
3.6.1 Effects of household demographics on household energy expenditure 95
3.6.2 Household size and economies of scales in household energy expenditure 116
3.7 Discussion 118
3.7.1 Comparisons with other studies 118
3.7.2 Implications of declines in household size in India 121
3.8 Conclusion 122
Chapter 4 The amenity value of climate to India households 124
4.1 Introduction 124
4.2 Literature Review 128
4.2.1 Applications of the AIDS model in India 129
4.2.2 Hedonic studies into the amenity value of climate 131
4.2.3 Life satisfaction studies into the amenity value of climate 137
4.2.4 Equivalence scales studies into the amenity value of climate 141
4.2.5 Literature discussion 143
4.3 Methodology 144
4.3.1 The AIDS Model 145
4.3.2 Incorporating climate amenities into the demand system 146
4.3.3 Demand shifters 148
4.3.4 Price and expenditure elasticities 149
4.4 Data and empirical specifications 150
4.4.1 Household survey data 150
4.4.2 District-level climate data 152
4.4.3 Variable definitions 156
4.4.4 Summary statistics 158

4.4.5 Empirical specifications 161
4.5 Regression results 162
4.5.1 Results from the baseline model 163
4.5.2 Sensitivity analysis 166
4.6 Discussion 174
4.7 Conclusions 177
Chapter 5 Conclusions, Limitations and Future Work 180
5.1 Conclusions 180
5.2 Limitations 182
5.3 Possible future work 184
References 186
Appendices 199
Appendix 1 The comparison of all commodity categories across 2013-2017 CHFS survey .. 200
Appendix 2 The estimation results from the Engel's Method 202
Appendix 3 STATA codes for estimating the baseline model in chapter 4 204

List of Tables

Table 2.1 OECD and square root equivalence scales 14
Table 2.2 Variable definitions 35
Table 2.3 Sample means of variables (whole, urban and rural) 36
Table 2.4 Sample means of expenditure shares by household types 37
Table 2.5 Estimation results of the simple QUAIDS model (without household demographics) 39
Table 2.6. Price and expenditure elasticities from simple QUAIDS model 41
Table 2.7 Estimation results for QUAIDS model with scaling function 42
Table 2.8 Estimation results for QUAIDS model with scaling function 44
Table 2.9 Estimation results for QUAIDS model with scaling function 45
Table 2.10 Estimation results for QUAIDS model with scaling function: 46
Table 2.11 Engel's equivalence scales: individuals divided into adults and children 48
Table 2.12 Engel's equivalence scales: children divided into different age cohorts 48
Table 2.13 Engel's equivalence scales: children divided into male and female children .. 49
Table 2.14 Comparisons with previous QUAIDS model applied in different countries . 51
Table 3.1 Percentage change in households by household size during 2001-2011 56
Table 3.2 Household economies of scale in energy consumption from previous studies 71
Table 3.3 Variable definitions 86
Table 3.4 descriptive statistics 89
Table 3.5 Estimated economies of scale indices using Kakwani-Son's method 92
Table 3.6 Estimated marginal effects for traditional fuel: dung cake 98
Table 3.7 Estimated marginal effects for traditional fuel: firewood 100
Table 3.8 Estimated marginal effects for transition fuel: PDS kerosene 105
Table 3.9 Estimated marginal effects for transition fuel: non-PDS kerosene 107
Table 3.10 Estimated marginal effects for modern fuel: electricity 112
Table3.11 Estimated marginal effects for modern fuel: LPG 114
Table 3.12 Estimated semi-elasticities of household size (traditional fuels) 117
Table 3.13 Estimated semi-elasticities of household size (transition fuels) 117
Table 3.14 Estimated semi-elasticities of household size (modern fuels) 118

Table 3.15 Implications of a 0.5 decline in average household size on household energy expenditure 122
Table 4.1 Distribution of households across districts in each month 152
Table 4.2 Variable definitions 156
Table 4.3 Summary statistics for all non-climate variables used in the analysis 159
Table 4.4 Summary statistics for climate variables 160
Table 4.5 Estimation results for the baseline model 163
Table 4.6 Wald test for joint significance of the intercepts and parameters of expenditure 165
Table 4.7 Price and expenditure elasticities using the estimates from the baseline model 166
Table 4.8 Selected results for the extended QUAIDS model 167
Table 4.9 Selected results for the baseline model without homogeneity and symmetry restrictions 168
Table 4.10 Selected results for the baseline model using weighted observations 169
Table4.11 Selected results for the baseline model for urban/rural subsample 170
Table 4.12 Selected results for the baseline model for high-/low-income subsample ... 171
Table 4.13 Selected results for Ray's scaling method 172
Table 4.14 Selected results for the baseline model with different measure of climate ... 173
Table 4.15 Estimated climate equivalence scale for major cities in India 176

List of Figures

Figure 1.1 Changes in average household size over year in China and India.....3

Figure 3.1 India's total carbon emissions from 1990-2020..... 57

Figure 4.1 Number of sample households in each month in 2011-2012 India's household CES survey..... 151

Figure 4.2 District-level annual precipitation in India..... 154

Figure 4.3 District-level annual temperature in India..... 155

Figure 4.4 District-level annual relative humidity in India..... 155

Figure 4.5 Map of ten major cities in India..... 175

Chapter 1 Introduction

Some people say that a growing population is the primary contributor to environmental issues, while others argue that it is the ownership of motor vehicles or the growth in GDP that leads to environment degradation. Beyond these commonly discussed factors however, trends in household size (i.e., the number of members in the household) could also play a significant role in influencing environmental outcomes, including climate change.

In this thesis, I focus on investigating household equivalence scales, economies of scale in household energy use, as well as the amenity value of climate to households based on household level consumption expenditure data obtained from the nationally representative household surveys in China and India, the two most populous countries in the world, accounting for about 35% of world population. Household equivalence scale is the scale to income or total expenditure that makes two different households equally well off. It is defined as the ratio of the income or total expenditure between households with different sizes and/or compositions. Household economies of scale means that households' cost of living does not need to increase proportionately as the number of household members increases, keeping household utility level constant. The amenity value of climate is defined as how much are households willing to pay in order to possess a specified change in climate, holding prices and income unchanged.

Household size, as well as other household socio-economic characteristics, can affect households' economic decisions, such as what and how much to consume (e.g., Deaton and Paxson, 1988), household savings (e.g., Curtis et al., 2017). Knowledge on the patterns

and trends of household size as well as composition is helpful to identify any opportunities or challenges in the achievement of the United Nations sustainable development goals, such as poverty reduction and environment protection (United Nations, 2019). Households with different sizes may have different standards of living, so we cannot compare the welfare levels between households with different sizes. However, household equivalence scales can be used to make the welfare comparisons between households with different sizes and/or compositions.

Household size plays a significant role in affecting household consumption patterns due to economies of scale, which means that total expenditure does not need to be doubled to maintain the same level of living standard as the household size is doubled, because some goods can be shared among household members. it is worthwhile to emphasize that economies of scale do not refer equally to all household commodities. Some commodities are much easier for the household to share than others. One such example of a commodity that is relatively easy to share is energy, considering that households consume energy goods mainly for the purposes such as cooking, lighting, heating and cooling. According to Edgar et al., (2009), 72% of global greenhouse gas (GHG) emissions are directly or indirectly from household consumption on the global level, with the rest coming from government consumption (10%) and investments (18%), following the conventions of national accounts. In particular, 19% of households' total GHG emissions result from household consumption on shelter (i.e., the operation and maintenance of residences), most of which is related to the direct energy consumption. Therefore, the changes in household size can

markedly affect households' energy consumption, and thus affect the environment through different ways, such as GHG emissions, indoor air pollutants, etc.

The average household size in many countries has declined with higher economic development, including China and India. Figure 1.1 shows the downward trends in average household size in China and India in recent decades. Estimating the economies of scale in household energy use is helpful to understand the impact of the decline in household size on energy use, as well as ultimately, the environmental implications of decreasing household size.

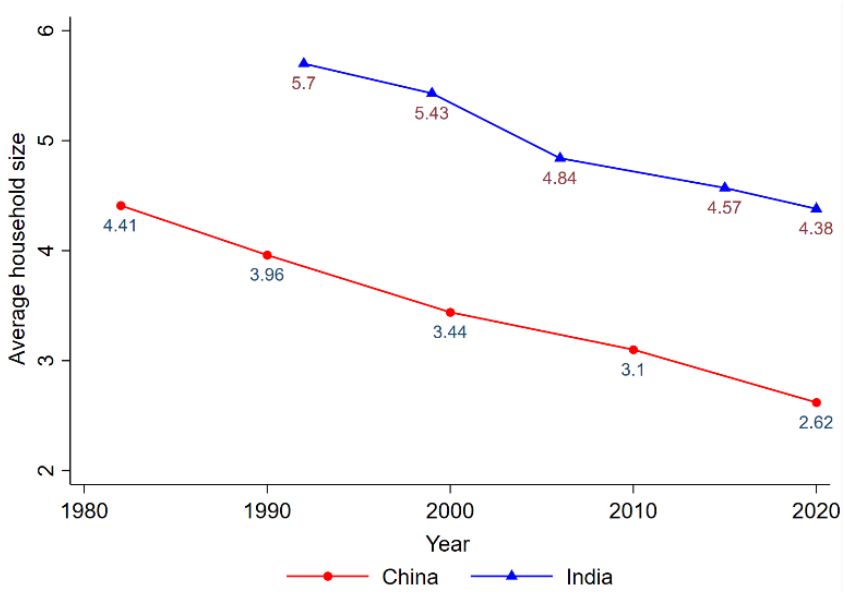


Figure 1.1 Changes in average household size over year in China and India
Data source: China's national population census; Demographic and Health Surveys in India

Similar to household size, households living in different climates may also have different standards of living, keeping all other factors identical. Few people may prefer to live in the freezing tundra or hot desserts an choose instead to reside in places with a moderate

climate. This can be explained by household production theory developed by Becker (1965). More specifically, climate along with purchased commodities are demanded by households as intermediate goods to produce direct service flows (e.g., a warm interior to the house), which enter households' utility function directly. Climate patterns in some locations can facilitate the production of some types of service flows and thus these households can produce the same bundles of service flows at lower costs.

In chapter two, I first investigate household equivalence scales based on the cross-sectional household level data which obtained from the China Household Financial Survey (CHFS) in 2017. This is motivated by the fact that China has replaced its two-child policy by three-child policy very recently. Correspondingly, the government may need other policies to encourage couples to have more children, such as providing subsidies for an additional child. This requires the information on how much income is needed for a household with one more child added, and household equivalence scales can ideally serve for this purpose. The simplest approach to estimate household equivalence scale is Engel's approach, the idea of which is that two households are assumed to be at the same welfare level if their expenditure shares on food are identical. The main problem of this method, however, is that there lacks a theoretical foundation to support this crucial assumption. Therefore, I employ a different method from Engel's method which only considers household food expenditure. More specifically, a demand system model is adopted, namely the Quadratic Almost Ideal Demand System (QUAIDS) model, developed by Banks et al. (1997) for different commodities consumed by household. This model is estimated with household characteristics included using Ray's scaling method (Ray, 1983). Taking the

single-adult household as a reference, the household equivalence scale is estimate by how much extra income does one household need if one more person added into the household in order to obtain the same level of welfare as the reference household. Household members are further disaggregated into different gender-age cohorts to analyse the equivalence scales for households with a different demographic composition.

In chapter three, I estimate the economies of scale in energy consumed by households in India. This analysis focuses on Indian households, which is motivated by the fact that, as mentioned above, India is also experiencing a decline in average household size, and moreover, it is the third largest carbon emitters in the world after China and the United States. Therefore, knowledge on the economies of scale in household energy use can be used to investigate the impact of the decline in household size on GHG emissions, which is the main cause of climate change. This study is based on household level data obtained from Household Consumer Expenditure Survey in 2011-2012, which provides detailed information on households' expenditure on different types of energy. I investigate two different but not unrelated concepts. More specifically, following Kakwani and Son (2005), I attempt to estimate the extent to which energy consumption considered by the household to be a public rather than a private good. Furthermore, I also estimate the per capita elasticity of demand with respect to household size. The former measure is estimated from a Working-Leser demand system model to analyse whether there exists economies of scale in household energy consumption. The latter elasticity is estimated from the double hurdle model due to zero expenditures being observed in sample households and also a suspicion that the factors driving the adoption decision and the decision about how much to

consume, are different. These per capita expenditure elasticities with respect to household size are estimated for different types of energy, including traditional energy (dung cake and firewood) transition energy (kerosene), as well as modern energy (electricity and LPG).

In chapter four, based on the same household survey data used in chapter three, I estimate the amenity value of climate to Indian households by, just as in chapter two, modifying the demand system model, in order to incorporate climate variables into the analysis. I continue to focus on India, one of the countries that are most vulnerable to climate change in the world. Besides, its geographical and topographical characteristics lead to significant variations in climate across the country, which makes it a particularly suitable country for this analysis. Fortuitously, the household survey data also has information on household location at the district level (625 out of 640 districts are covered in the survey). In effect this allows us to measure climate at the level of the household level and thus to have sufficient variations in climate, which is very important to this research. For simplicity and the convenience of analysis, I employ the Almost Ideal Demand System (AIDS) model (Deaton and Muellbauer, 1980) and Barten's price scaling method (Barten, 1964) to include climate into the demand system. I also however go on to consider the QUAIDS model an extension of the AIDS model, and Ray's method of scaling (of which Barten's method is a special case). Climate is measured by annual averages of heating degree days, cooling degree days, precipitation, as well as relative humidity. The value of climate amenities are thus revealed from the role played by climate in determining households' cost of living. I further calculate climatic equivalence scales based on the estimation results, i.e., how much

extra do households living in a particular climate require in order to achieve the same level of welfare as an otherwise identical household residing in some reference climate.

Chapter 2 Estimating household equivalence scales for China: An application of the QUAIDS model

Abstract

Household equivalence scales are used to make the utility of incomes comparable across households with different size and compositions which is of great importance in various fields of economics such as the measurement of inequality and poverty. However, studies estimating household equivalence scales in China are scarce. In this chapter, China Household Finance Survey (CHFS) data is used to estimate household equivalence scales using the Quadratic Almost Ideal Demand System (QUAIDS) model with household demographic characteristics incorporated. Taking a single-adult household as the reference, equivalence scales are estimated for households with children (aged 0-17), additional adults (aged above 18), and with children further split into different age/gender cohorts. The results suggest that there exist significant economies of scale in Chinese households, and that the equivalence scales are different between urban and rural households.

2.1 Introduction

China has the second largest number of people in the world.[1] According to its seventh national census, the total population in 2020 has exceeded 1.4 billion, 63.89% of which resides in urban areas and the total number of households is over 522 million. The average household size is 2.62 in 2020 while that for urban households is 2.49, which is smaller than the corresponding rural household size (2.70). Regarding the composition of Chinese

[1] India has overtaken China recently to be the most populous country in the world, according to the latest estimates of global population from the United Nations.

households, household members aged 0-14, 15-64 and 65 above account for 17.95%, 68.55% and 13.5%, respectively. Notably, the proportion of the children within households has been decreasing while the proportion of elderly people has been increasing since the end of 1970s. This change in the household composition stems from the implementation of the one-child policy.

China's one-child policy was implemented in the late 1970s in order to control the rapid growth of population. Under this policy, the vast majority of couples could only have one child. Since the implementation of the policy, the fertility rate has decreased considerably. However, this policy also caused some problems which appear serious today, including a shortage of labour and gender imbalance. Meanwhile, with the rapid development and the improvement of medical care, the proportion of elderly people have been increasing significantly. To tackle these problems, the Chinese government has progressively loosened the one child policy, and this policy was eventually discontinued in 2015, and was replaced by a two-child policy, which allows a couple to have up to two children. However, the implementation of the new policy does not seem to have brought a significant increase in the fertility rate, and quite recently, the two-child policy has been replaced by a three-child policy in 2021.

There is no doubt that raising children will increase households' costs of living, and thus increase the financial burden on the households. However, the question I wish to address is to what extent will household total expenditure be affected by the presence of additional children? More specifically, how much income or total expenditure is needed for the household with one more child added, in order to keep the welfare level unchanged and

as before? The answers to these questions may be very helpful for policies which aim to encourage families to raise more children, and they can also make the welfare levels or incomes for households with different compositions comparable.

Household equivalence scales are ideally suited to answer those questions mentioned above. An equivalence scale is the scale to income or total expenditure that makes two different households equally well off. It is defined as the ratio of the income or total expenditure between households with different sizes and/or compositions. Taking the so-called modified OECD scale as an example, the household comprising a single adult is treated as the reference, i.e., the scale to this type of household is defined to be one. The modified-OECD scale for a child (aged less than 14) is then 0.3. That means, the income or total expenditure for the household with an adult and a child needs to rise by 30% of the single-adult household in order to attain the same level of welfare. Household equivalence scales have long played a crucial role in various fields of welfare and public economics, including the measurement of inequality and poverty, as well as the design of tax and welfare systems (Biewen and Juhasz, 2017). Apart from the so-called expert scales (such as the OECD scale), various other methods have been proposed to estimate the household equivalence scales, and the results from different methods are not always similar.

Given its policy relevance, studies on the estimation of household equivalence scales in China are surprisingly scarce. Therefore, in this study, I focus on estimating equivalence scales for Chinese households using nationally representative household survey data. In order to do this, the Quadratic Almost Ideal Demand System (QUAIDS) model is

employed which is extended to incorporate household composition variables and other demographic characteristics. The results from the QUAIDS model suggest that taking the single-adult household as the reference, the cost of living of the reference household would increase by about 19% with one more adult added, while having additional one child would increase the cost by about 11%. Furthermore, children are divided into different age/gender cohorts, and the equivalence scales of an additional child aged 0-1, 3-6, 7-12, 13-17 are 0.17, 0.13, 0.06 and 0.14, respectively. There are no significant differences between the equivalence scales of male and female children. I also find that there exist significant economies of scale in households' consumption expenditure. In addition, the Engel's method, which has been widely used in existing literature to estimate household equivalence scales is also estimated in this study. However, this method does not perform well because the Engel's results show that the cost of living for larger households sometimes could even be lower than that for households with less members.

This study contributes to the literature in the following ways. First, to my knowledge, this is the first study to employ the QUAIDS model to estimate household equivalence scales in China. Second, compared to the only other study also applying a demand system model to estimate household equivalence scale in China (Chen, 2006), this study is based on more recent and nationally representative household data. More specifically, the previous study uses household level data obtained in 1993 and 1997, which only covers urban households located in 7 provinces in China, while the data used for the analysis is more recent (2017) and covers both urban and rural households from 29 provinces in China. Third, by comparing the estimated household equivalence scales from two different

models, i.e., the QUAIDS model as well as the Engel's method, a simple and widely used approach in existing literature, it is found that the QUAIDS model works better than the Engel's method.

The remainder of this chapter is organized as follows. In section 2.2, the related literature, in particular, those studies related to China is reviewed. Section 2.3 introduces the QUAIDS model, how I incorporate household demographic characteristics into the demand system, as well as the Engel's method. Section 2.4 describes the data used in this study and provides descriptive statistics. The estimation results are presented in section 2.5, and section 2.6 discusses the estimation results and compares them to those in previous related studies. Section 2.7 concludes.

2.2 Literature Review

In this section, the definition of household equivalence scale is first introduced, followed by three different approaches that have been used to estimate household equivalence scales, which are: (1) the assessments of 'experts' or more or less arbitrary rules-of-thumb; (2) the subjective approach which is based on the data on individuals' subjective evaluations of utility obtained from; (3) expenditure-based approach. In this study, household equivalence scales in China are investigated by estimating the QUAIDS model based on household consumption expenditure data, which belongs to the expenditure-based approach. Therefore, I mainly focus on the review of literature that uses this approach, and in particular, those that identify household equivalence scales from the QUAIDS model.

2.2.1 Household equivalence scale

The household equivalence scale S(.) is defined as the ratio of the expenditure functions between households with different demographic characteristics:

$$\mathrm{S}(u, \boldsymbol{p}, \boldsymbol{z}_h, \boldsymbol{z}_r) = \frac{E_h(u, \boldsymbol{p}, \boldsymbol{z}_h)}{E_r(u, \boldsymbol{p}, \boldsymbol{z}_r)}$$

where E(.) is the expenditure function, u is the utility level, **p** is the price vector, and **z** is the vector of household demographic characteristics. The subscripts r and h denote two different households, with r as the reference household. The reference household is usually set as a household with a single adult, although other household type such as the household with a couple of adults could also be chosen in some studies.

However, equivalence scales cannot be estimated if utility is assumed to be ordinal, because it requires the comparisons of utility, something that is not possible under the assumption of ordinal utility (Pollak and Wales, 1979; Dudel et al., 2020). Therefore, any approach that attempts to identify the household equivalence scales needs to deal with this issue.

There have been three main approaches used to estimate household equivalence scales in the literature. The first approach is just based on the assessments of 'experts' or more or less arbitrary rules-of-thumb, which is less evidence-based than other approaches. Examples include the so-called OECD equivalence scale (including both the old OECD scale and OECD-modified scale), as well as the square root equivalence scale (see Table 2.1).

Table 2.1 OECD and square root equivalence scales

Household composition	Equivalence scale		
	Old OECD scale	OECD-modified scale	Square root scale
1 adult (reference)	1	1	1
2 adults	1.7	1.5	1.4
2 adults, 1 child	2.2	1.8	1.7
2 adults, 2 children	2.7	2.1	2.0
2 adults, 3 children	3.2	2.4	2.2

Source: OECD (2005)

The second approach to estimating household equivalence scales is based on the data on individuals' subjective evaluations of utility obtained from income (Koulovatianos et al., 2005; Biewen and Juhasz, 2017; Borah et al., 2019). This subjective approach is pioneered by Van Praag (1986). The idea of this approach is that households with different compositions are asked to state which income level would be considered as "good", "sufficient", "bad", etc. These survey data can be used to estimate household cost functions, from which we can derive the implicit household equivalence scales. The equivalence scales estimated from this approach are typically quite flat, which means a small increase in costs and large economies of scale for additional members included in the household.

The third approach is to estimate equivalence scales from the data on household consumption and expenditure. As mentioned above, the assumption of ordinal utility makes the equivalence scale impossible to be identified. Thus, an additional assumption is required to overcome this identification problem, which is the so-called "Independence of Base" (IB) assumption (Blundell and Lewbel, 1991). The IB assumption means that equivalence scales do not depend on the base level of utility at which the expenditure

functions are compared. In practice, this assumption is often equivalent to stating that equivalence scales are independent of households' income or expenditure level.[2]

2.2.2 Expenditure-based approach to estimate household equivalence scales

Using household expenditure data, equivalence scales can be estimated by using the share of expenditure on a particular commodity category as a proxy of household welfare. Prominent examples of this approach include the Engel's method as well as its variant, namely the Rothbarth model (Rothbarth, 1943). Recently, Dudel et al. (2020) review various expenditure-based approaches to estimate the equivalence scale and employed German household expenditure data to estimate equivalence scales using those methods. Their results suggest that the estimated scales vary to the method applied.

Engel's method is the simplest and perhaps the best-known approach to estimate household equivalence scales using household expenditure data. The idea of this method is that two households are assumed to be at the same welfare level if their expenditure shares on food are identical. It is grounded in two empirical regularities: one is Engel's law, which refers to the share of food expenditure decreasing with an increase in household income or total expenditure (Engel, 1895); the other regularity is that for households with the same income or total expenditure, the food share component increases with the number of children (Deaton and Muellbauer, 1986). Therefore, equivalence scales can be

[2] The IB assumption is a very strong restriction. Although some approaches have been proposed that relax the IB assumption, and some studies by using subjective-wellbeing approach show that household equivalence scales decrease with household income, the expenditure-based approach based on the IB assumption remains the standard approach to estimate household equivalence scales.

directly calculated by comparing the total expenditure between households with different compositions, but which have the same expenditure shares on food.

Deaton and Muellbauer (1986) analyse the theoretical basis for measuring the cost of children to a household and used Engel's method to empirically estimate household equivalence scales for children in Sri Lanka and Indonesia. The Working-Leser form of Engel equation is used to estimate Engel curve for food, in which the food share is a function of the logarithm of per capita expenditure, as well as household demographics (the exact specification of this form of Engel equation is shown in section 3). Taking a childless couple as a reference, their estimation results suggest that in Sri Lanka, household expenditure needs to rise by 41% with one additional child included, while a couple with two children needs 77% more income to be equally well off. For Indonesian households, the equivalence scales for small children (aged less than 5 years) are similar to those in Sri Lanka: two adults plus one additional child (two children) needs 45% (86%) more total expenditure to be as well off as a childless couple.

Engel's method has been widely used in the estimation of household equivalence scales in the literature due to the simplicity of the estimation procedure and the modest data requirements (such as Tsakloglou, 1991; Bosch-Domènech, 1991; Deaton, 1997; Phipps and Garner, 1994; Jayasinghe et al., 2016). However, this method also has its limitations. Fundamentally, there is no theoretical foundation for the crucial assumption that the food share can correctly reflect the welfare level. Besides, it has been found to be problematic when comparing the food shares between two households with different compositions. Nicholson (1976) argues that the equivalence scale may be overestimated by the Engel's

method. More specifically, the consumption patterns of children are usually different from those of adults. The share of the expenditure for children, or at least for very young children spent on food is likely to be greater than the proportion of the expenditure for adults spent on food. Therefore, the food share of the new income "allocated" to the household to keep the welfare level unchanged due to the arrival of the children is larger than that of the original total income. This restoration of the food share after the inclusion of the children implies that the welfare level of the couple plus an additional child is underestimated by their food share, and thus a further increase in income or total expenditure is needed in order to be as well off as the childless couple. That is, the equivalence scale of children would be overestimated by the Engel's method.

There is a widely known variant of the Engel method suggested by Rothbarth (1943), which is also used to estimate household equivalence scales in the literature (Deaton and Muellbauer, 1986; Nelson, 1992; Deaton, 1997; Dudel et al., 2020). Under the Rothbarth method, the expenditure on goods only consumed by adults are used to proxy the welfare level of adults. A couple with children and a childless couple are assumed to be equally well off if their expenditures on adult goods are identical, which implies that the arrival of children is assumed to affect the expenditure on adult goods only through income-like effects. Therefore, the equivalence scale for children can be estimated from the ratio of income or total expenditure between the household with-children and childless household. Rothbarth (1943) uses a very broad way to define the adult goods, whilst subsequent researchers use a narrower definition of adult goods, which usually include alcohol and tobacco, as well as adult clothing.

Deaton and Muellbauer (1986) also adopt the Rothbarth method to estimate the equivalence scale for children in Sri Lanka and Indonesia. The Rothbarth scales are smaller than the Engel scales: the arrival of a child now costs 12% more for a previously childless couple and income or total expenditure needs to rise by 21% with the arrival of two children in Sri Lanka. For Indonesian households, a small child costs 10% more while the cost of the arrival of two small children is only 16% more for a previously childless couple.

There are, however, some problems with the Rothbarth model as well. First, similar to Engel's method, it lacks a convincing basis for the assumption on the identification of adults' welfare level. Second, it is always hard to find goods that are only consumed by adults. Third, the presence of children may affect the expenditure on adult goods through substitution effects. In particular, the household compensated for the cost associated with the arrival of children will spend more expenditure on adult goods, and thus the Rothbarth method tends to underestimate household equivalence scales for children (Deaton and Muellbauer, 1986; Deaton, 1997).

Household equivalence scales can also be identified from the estimation of a demand system, which considers not just the expenditure on a single commodity (such as food in the Engel's method), but the expenditures on all commodities. Since Stone (1954) firstly developed the Linear Expenditure System model (LES), which is explicitly derived from consumer theory, different specifications of the demand system have been proposed and used to analyse household demands in the literature, including the Rotterdam model (Barten, 1964), the Translog model (Christensen et al., 1975), as well as the Almost Ideal Demand System(AIDS) model (Deaton and Muellbauer, 1980) and its quadratic extension,

QUAIDS model (Banks et al; 1997). In particular, the QUAIDS model, which essentially assumes that the expenditure share of each commodity is quadratic in the logarithm of total expenditure, has become the standard approach to model household demand in the literature (section 2.3 provides more details about the QUAIDS model).

Michelini (2001) uses cross-sectional data from the Household Economic Survey in New Zealand to estimate the QUAIDS model for the purpose of estimation of household equivalence scale. Household expenditures are aggregated into six commodity categories: food, housing, apparel, transport, health and education, as well as recreation. Household demographic characteristics are introduced into the demand system by Ray's (1983) price scaling technique, for the purpose of estimation of equivalence scales. To be brief, Ray's method directly comes from the definition of the general equivalence scale, with the scaling function defined as the ratio of the total income or expenditure of a household with different size and/or compositions necessary to obtain the same level of welfare as the reference household (section 2.3 also provides more details about Ray's method). Michelini specifies a simple form for the scaling function, which is a linear combination of household composition variables, including the number of adults, infant children (aged less than 3), young children (aged 3-7), teenage children (aged 8-14), as well as a dummy variable which takes one if there are more than three children in the household. The results show that taking a single-adult household as the reference, the equivalence scale for an additional adult, infant child, young child and teenage child are 0.68, 0.72, 0.58 and 0.49, respectively.

Blacklow et al (2010) also use the QUAIDS model with household demographics incorporated into the model by Ray's method in order to estimate the equivalence scales

for Australian households, but they construct a different form of the scaling function which is nonlinear in household demographic variables. Four waves of Australian Household Expenditure Survey data are used for the estimation and analysis. Households' expenditures are divided into nine categories, and prices used for each commodity are measured by the quarterly consumer pricing index, which is re-weighted to match the commodity categories.[3] Household members are classified as adults and children, but adults are not defined like in other studies. Instead, in this study an adult is defined as an individual aged over 23, or aged 15-23 but neither a full-time student nor a dependent. The equivalence scales for different types of households are estimated. With a single-adult household as the reference, the equivalence scale is 1.36 for household with 2 adults, 1.526 for 2 adults plus an infant, 1.685 for 2 adults plus an infant and a child aged less than 15, and 1.953 for 2 adults plus an infant, two children aged less than 15 and one child aged above 15.

Vashchilko (2014) uses the same specification of the scaling function as employed by Blacklow et al. (2010) to estimate household equivalence scales for Belarus. The QUAIDS model is estimated using the Belarusian Household Survey on Incomes and Expenditures in 2006 and 2010. Equivalence scales for different types of households are estimated. Compared to a single-adult household, the equivalence scale for the household with two adults is 1.37; for the household with one adult and 1 child aged less than 12, the

[3] The nine commodity categories include: (1) accommodation; (2) Electricity and fuel; (3) food and non-alcoholic beverages; (4) alcohol and tobacco; (5) clothing and footwear; (6) health and personal care; (7) transport; (8) recreation; and (9) miscellaneous items.

equivalence scale is 1.16; for the household with one adult and 1 child aged 13-17, the equivalence scale is 1.27.

Dudel et al. (2020) estimate an extended QUAIDS model for 12 commodities where the parameters on total expenditure are allowed to vary according to household demographic characteristics, and the simple linear form of the scaling function is used. Different from other studies in which household compositions are measured by the number of different group of members (e.g., adults and children), a set of dummy variables for five different household types are used in this study: a single adult (A, the reference), two adults (AA), two adults and a child (AAC), two adults and two children (AACC), as well as two adults and three children (AACCC). Using German household survey data, they estimate the household equivalence scales for different types of households.

2.2.3 A review of household equivalence scale in China

There are very few studies in the existing literature that have attempted to estimate household equivalence scales in China, two of which are based on the Engel's and Rothbarth methods. Tian et al. (2018) use Engel's method to estimate equivalence scales for both male and female children in order to study whether there exists gender discrimination in China. The rationale is that if the arrival of a boy increases the food share by more than the arrival of a girl, it implies that parents are willing to spend more on boys than girls, which suggests gender bias in intra-household resource allocation. Their results indicate that gender bias still exists in China, in particular in rural areas. Liu et al. (2021) employ both the Engel's and Rothbarth model to estimate the costs of children to Chinese households. They find that, taking a childless couple as reference household, the presence

of one child would increase household total expenditure by 63.91%-67.62%, while in the presence of two children, total expenditure would increase by 139.28%-149.99%.

To my knowledge, Chen (2006) is the first and only application of the demand system modelling approach to estimating household equivalence scales in China. Using the data taken from Chinese Urban Household Survey, Chen (2006) estimates the AIDS model to compute equivalence scales for urban households. The household demographic variables are introduced into the demand system by a different method proposed by Lewbel (1985), which is based on the technique of direct modification of the expenditure function in order to allow total expenditure to be affected by household demographics. Thus, the equivalence scale can be estimated by the ratio of the modified expenditure functions between the households with different compositions and the reference household. Only five household types are chosen for estimation, while other types of households are dropped from the sample.[4] A set of dummy variables for children are used to capture household compositions, which are K_1 (kid aged 0-2), K_2 (kid aged 3-6), and K_3 (kid aged 7-18). Since the number of adults in the sample households is at most two, the number of adults is measured by a dummy which equals to 0 for single adult households. In addition, regional dummies are also introduced into the demand system. The equivalence scales of different household types are calculated with the single-adult household as the reference, based on the estimation results of the AIDS model, which are further applied to calculate the poverty lines for households with different compositions.

[4] The five household types chosen by Chen (2006) are: single person without child, single parent with one child, single parent with two children, childless couple, and couple with one child.

Compared to Chen (2006), I advance the estimation of household equivalence scales in China in several ways. First, a different (but related) demand system model (the QUAIDS model) is employed. The QUAIDS model can better capture the non-linear Engel's curve compared to the AIDS model. Second, this study employs a different method to incorporate household demographic characteristics into the demand system, which is more convenient to interpret household equivalence scales. Third, the household expenditure data used in Chen (2006) is out of date, and only covers the urban households from 10 provinces. This study however, is based on the data taken from a more recent and nationally representative household survey, which includes both urban and rural households from 29 provinces in China. Besides the QUAIDS model, Engel's method is also applied in this study, the results from which are then used to make comparisons with those from the QUAIDS model and the work of Tian et al. (2018).

2.3 Methodology

In this study, I employ the QUAIDS model to estimate household equivalence scales in China, which is a quadratic extension of the AIDS model. The AIDS model gives an arbitrary first-order approximation to any demand system which satisfies all standard axioms of consumer theory and aggregates perfectly over consumers (Deaton and Muellbauer, 1980). Furthermore, Banks et al. (1997) find that the Engel curves for some commodities (which capture the relation between the expenditure share on some commodity and the logarithm of total expenditure) are non-linear. Therefore, they extend the original AIDS model by incorporating a non-linear (quadratic) term of the logarithm

of total expenditure, i.e., the QUAIDS model. The advantage of the QUAIDS model is that it can better capture the effect of quadratic terms of log expenditure on consumption shares than the traditional AIDS model. The AIDS model and its quadratic extension have been widely used for modelling household demand. This section introduces the specifications of the QUAIDS model and Ray's price scaling method (Ray, 1983), which is used to incorporate household demographic characteristics into the demand system. Latterly, Engel's method is also introduced in this section.

2.3.1 The QUAIDS Model

According to Banks et al. (1997), The QUAIDS model can be derived from an indirect utility function taking the form:

$$ln\, V = \left\{\left[\frac{ln\, y - ln\, a(\boldsymbol{p})}{b(\boldsymbol{p})}\right]^{-1} + \lambda(\boldsymbol{p})\right\}^{-1} \tag{2.1}$$

where y is total expenditure, $\boldsymbol{p}$ is the price vector, the term $\left[\frac{\ln y - \ln a(\boldsymbol{p})}{b(\boldsymbol{p})}\right]^{-1}$ is the indirect utility function of the PIGLOG demand system, $\ln a(\boldsymbol{p})$, $b(\boldsymbol{p})$ and $\lambda(\boldsymbol{p})$ are functions of prices.[5] In particular, $\ln a(\boldsymbol{p})$ takes the translog form:

$$ln\, a(\boldsymbol{p}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i\, ln\, p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}\, ln\, p_i\, ln\, p_j \tag{2.2}$$

where p_i is the price of good i, ($i = 1, 2, \dots, n$).

$b(\boldsymbol{p})$ is the Cobb-Douglas price aggregator:

$$b(\boldsymbol{p}) = \prod_{i=1}^{n} p_i^{\beta_i} \tag{2.3}$$

$\lambda(\boldsymbol{p})$ is a homogenous of degree zero function of log prices:

[5] Price-Independent Generalized Logarithmic (PIGLOG) demand system refers to that expenditure shares are linear in log total expenditure alone.

$$\lambda(\boldsymbol{p}) = \sum_{i=1}^{n} \lambda_i \ ln\, p_i \tag{2.4}$$

The commodity shares of the QUAIDS model can be obtained by applying Roy's identity to the indirect utility function (Eq. 2.1), which are expressed as:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \, ln\, p_j + \beta_i \, ln\left[\frac{y}{a(\boldsymbol{p})}\right] + \frac{\lambda_i}{b(\boldsymbol{p})}\left\{ln\left[\frac{y}{a(\boldsymbol{p})}\right]\right\}^2 \tag{2.5}$$

where w_i is expenditure share of good i, which equals to the expenditure on good i divided by total expenditure. α, β, γ and λ are parameters to be estimated. Note that the AIDS model is nested within the QUAIDS if λ_i is set to be zero for all i (Banks et al., 1997).

Furthermore, in order to comply with the demand theory and to reduce the number of parameters to be estimated, three constraints are further imposed on the model, which are the adding-up restriction, homogeneity and Slutsky symmetry.

(1) The adding up restriction means that the sum of expenditure shares must be one, which implies that:

$$\sum_{i=1}^{n} \alpha_i = 1 \,; \sum_{i=1}^{n} \gamma_{ij} = 0, \forall j; \sum_{i=1}^{n} \beta_i = 0, and \sum_{i=1}^{n} \lambda_i = 0 \tag{2.6}$$

(2) Homogeneity means that the demand equation should be homogenous of degree zero in prices and total expenditure, which implies that the sum of the parameters of log-prices for each equation must be zero:

$$\sum_{j=1}^{n} \gamma_{ij} = 0, \forall i \tag{2.7}$$

(3) Slutsky symmetry deals with the substitution effects between goods, which implies that the cross-partial price derivatives are equal:

$$\gamma_{ij} = \gamma_{ji}, \forall i \neq j \tag{2.8}$$

2.3.2 Incorporating household demographics into the demand system

Different techniques can be used to extend the demand system to include demographic variables, such as demographic scaling, demographic translating (Pollak and Wales, 1978), as well as Ray's price scaling method (Ray, 1983). With demographic scaling, both prices and quantities of consumption goods are scaled by demographic variables. The scaling parameters can be either the same or different for each commodity. Suppose the original demand system is represented as $q^i = q^i(\boldsymbol{p}, y)$, $i = 1,2,3, \dots, n$, where q^i denotes the quantity of good i, $\boldsymbol{p}$ is the vector of prices, and y is the total expenditure. The vector of demographic characteristics is denoted by $\mathbf{z} = (z_1, z_2, z_3, \dots, z_k)$. Using this demographic scaling procedure, the original demand system is replaced by: $q_i = t_i q^i(p_1 t_1, p_2 t_2, p_3 t_3, \dots, p_n t_n, y)$, where the t's are the demographic scaling parameters, which are the functions of $\boldsymbol{z}$: $t_i = t_i(\boldsymbol{z})$. It can be seen that by using demographic scaling method, the changes in household demographics would induce a change in the effective prices of the marketed goods proportionally. The demographic scaling can be interpreted as reflecting the number of "equivalent adults" on the scale appropriate to good i.

Using the demographic translating method by contrast, the original system of demand equations is replaced by: $q_i = d_i + q^i(p, y - \sum_{i=1}^{n} p_i d_i)$, where d's are translating parameters, $d_i = d_i(\mathrm{z})$. The demographic translating can be interpreted as allowing the "necessary" or "subsistence" parameters of the demand system to depend on the demographic variables.

In this study however, I employ Ray's scaling approach to incorporate household demographic variables into the QUAIDS model. Different from demographic scaling and translating method which directly modify the original demand equations, Ray's scaling technique stems explicitly from the definition of the general equivalence scale, which is for each household with different demographic characteristics, the ratio between the expenditure of obtaining the same level of utility at given prices and the expenditure of the reference household. More specifically, for each household, the expenditure function can be expressed as follows:

$$e(\boldsymbol{p}, \boldsymbol{z}, u) = m_0(\boldsymbol{p}, \boldsymbol{z}, u) \times e^R(\boldsymbol{p}, u) \tag{2.9}$$

where $e^R(\boldsymbol{p}, u)$ is the expenditure function of the reference household; $m_0(\boldsymbol{p}, \boldsymbol{z}, u)$ is the scaling function which scales the expenditure function to account for the impact of demographic characteristics on expenditure; u represents utility, and $\boldsymbol{z}$ is the $k \times 1$ vector of demographic variables.

Furthermore, the scaling function is decomposed into two parts:

$$m_0(\boldsymbol{p}, \boldsymbol{z}, u) = \bar{m}_0(\boldsymbol{z}) \times \phi(\boldsymbol{p}, \boldsymbol{z}, u) \tag{2.10}$$

where the first term, $\bar{m}_0(\boldsymbol{z})$ measures the changes in a household's expenditure which only depend on $\boldsymbol{z}$, ignoring any changes in consumption patterns, while the second term, $\phi(\boldsymbol{p}, \boldsymbol{z}, u)$ measures the changes in relative prices and the actual goods consumed by the household. $\bar{m}_0(\boldsymbol{z})$ takes following form:

$$\bar{m}_0(\boldsymbol{z}) = 1 + \boldsymbol{\rho}'\boldsymbol{z} \tag{2.11}$$

where $\boldsymbol{\rho}$ is the vector of parameters to be estimated. The functional form of $\phi(\boldsymbol{p}, \boldsymbol{z}, u)$, however, is not unique. Following Poi (2012), $\phi(\boldsymbol{p}, \boldsymbol{z}, u)$ is defined as:

$$ln\,\phi(\boldsymbol{p}, \boldsymbol{z}, \boldsymbol{u}) = \frac{\prod_{i=1}^{n} p_i^{\beta_i} \left(\prod_{i=1}^{n} p_j^{\boldsymbol{\eta}_i' \boldsymbol{z}} - 1\right)}{\frac{1}{u} - \sum_{i=1}^{n} \lambda_i \, ln\, p_i} \tag{2.12}$$

The adding-up restriction ($\sum_{i=1}^{n} w_i = 1$) further requires that $\sum_{j=1}^{n} \eta_{sj} = 0$, ($\mathrm{s} = 1, 2, \ldots, \mathrm{k}$). It is easy to prove that once the prices for all commodities are set to be equal, $ln\,\phi(\boldsymbol{p}, \boldsymbol{z}, u)$ will equal to 0, and thus the scaling function will just equal to $\bar{m}_0(\boldsymbol{z})$, i.e.

$$m_0(\boldsymbol{p}, \boldsymbol{z}, u) = \bar{m}_0(\boldsymbol{z}) = 1 + \boldsymbol{\rho}' \mathbf{z} \tag{2.13}$$

The advantage of this functional form of $ln\,\phi(\boldsymbol{p}, \boldsymbol{z}, u)$ is that the resulting expenditure share equations are closely paralleled with those without demographics. The extended demand system with demographic variables takes the following form:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \, ln\, p_j + (\beta_i + \boldsymbol{\eta}_i' \boldsymbol{z}) \, ln \left[\frac{y}{\bar{m}_0(\boldsymbol{z}) a(\boldsymbol{p})}\right] + \frac{\lambda_i}{b(\boldsymbol{p}) c(\boldsymbol{p}, \boldsymbol{z})} \left\{ln \left[\frac{y}{\bar{m}_0(\boldsymbol{z}) a(\boldsymbol{p})}\right]\right\}^2 \tag{2.14}$$

where

$$\mathrm{c}(\boldsymbol{p}, \boldsymbol{z}) = \prod_{j=1}^{n} p_j^{\boldsymbol{\eta}_j' \boldsymbol{z}} \tag{2.15}$$

2.3.3 Price and Expenditure Elasticities

Based on the estimation results of the QUAIDS model, we can further estimate the demand elasticity of each commodity, which could help us to understand how household demand responds to changes in prices and total expenditure. Following Banks et al. (1997), the expenditure and price elasticities with demographics included can be obtained by differentiating Equation (2.14) with respect to $\ln y$ and $\ln p_j$, respectively:

$$\mu_i \equiv \frac{\partial w_i}{\partial \ln y} = \beta_i + \boldsymbol{\eta}_i' \boldsymbol{z} + \frac{2\lambda_i}{b(\boldsymbol{p})c(\boldsymbol{p},\boldsymbol{z})} ln \left[\frac{y}{\bar{m}_0(\boldsymbol{z})a(\boldsymbol{p})}\right] \tag{2.16}$$

and

$$\mu_{ij} \equiv \frac{\partial w_i}{\partial \ln p_j} = \gamma_{ij} - \mu_i \times \left(\alpha_j + \sum_k \gamma_{kj} \ln p_k\right) - \frac{(\beta_i + \boldsymbol{\eta}_i' \boldsymbol{z})\lambda_i}{b(\boldsymbol{p})c(\boldsymbol{p},\boldsymbol{z})} \left\{ln \left[\frac{y}{\bar{m}_0(\boldsymbol{z})a(\boldsymbol{p})}\right]\right\}^2 \tag{2.17}$$

Then the expenditure elasticities are computed by:

$$e_i = \frac{\mu_i}{w_i} + 1 \tag{2.18}$$

The compensated (Hicksian) price elasticities which are derived from the Hicksian demand function and assume that the consumer will be given enough wealth (compensation) to maintain the same level of utility before the relative price changes. The compensated price elasticity of commodity i with respect to the change in the price of good j can be obtained by:

$$e_{ij}^c = e_{ij}^u + e_i w_i \tag{2.19}$$

The uncompensated (Marshallian) price elasticities which are derived from the Marshallian demand function, assumes that expenditures remain unchanged when relative prices change. The uncompensated price elasticity of commodity i with respect to the change in the price of good j are computed as:

$$e_{ij}^u = \frac{\mu_{ij}}{w_i} - \delta_{ij} \tag{2.20}$$

where δ_{ij} equals to 1 if $i = j$ and 0, otherwise (Poi, 2012).

2.3.4 Engel's method

For the purposes of comparison, I also use the Engel's method to estimate household equivalence scales. Following Deaton and Muellbauer (1986), the Engel equation takes the

Working-Leser form, in which the food share is linear in the logarithm of household total expenditure:

$$w_f = \alpha + \beta ln\left(\frac{x}{n}\right) + \sum_j \gamma_j n_j + \varepsilon \tag{2.21}$$

where w_f is the expenditure share on food, x is total expenditure, n is household size, thus $\frac{x}{n}$ denotes the per capital expenditure, n_j is the number of individuals in different groups, and ε is the error term. α, β and γ are parameters to be estimated.

The Engel's method assumes that two households are in the same welfare level if their expenditure shares on food are identical. Therefore, for the household h and the reference household r, the equality of their food share implies:

$$\alpha + \beta ln\left(\frac{x_h}{n_h}\right) + \sum_j \gamma_j n_{hj} = \alpha + \beta ln\left(\frac{x_r}{n_r}\right) + \sum_j \gamma_j n_{rj} \tag{2.22}$$

Solving for the ratio of the total expenditure of household h and the reference household r gives us the equivalence scale, which is:

$$E_h = \frac{x_h}{x_r} = \left(\frac{n_h}{n_r}\right) exp\left[\sum_j \left(-\frac{\gamma_j}{\beta}\right)\left(n_{hj} - n_{rj}\right)\right] \tag{2.23}$$

2.4 Data

2.4.1 China Household Finance Survey (CHFS)

The household expenditure data and demographic characteristics used in this study are taken from CHFS, one of the nationally representative surveys of Chinese households. CHFS provides a wide variety of information on the sample of Chinese households, including demographic characteristics, assets and debts, insurance, income and expenditure, as well as other financial characteristics. All of these information is collected

by asking questions to the respondents directly through in-person interviews. The households included in this survey are randomly selected from 29 provinces in China.[6] The initial CHFS was conducted in 2011 and it has been conducted every two years since then. This study is based on the 2017 CHFS data. The main reason why I choose to use cross-sectional rather than panel survey data is that there are some discrepancies among the setting of the survey questions on household expenditure in different waves of the survey, which makes it difficult to keep the definitions of all commodity categories perfectly consistent. For example, the expenditure on internet is included in "communication" in the 2013 and 2017 surveys, while it is included in "entertainment" in 2015 survey; fitness expenditure is incorporated in "education" in the 2013 and 2015 surveys, while in 2017 survey it is excluded (For a detailed comparison of all the expenditure categories across the 2013-2017 CHFS surveys, see Appendix 1).

Household total expenditure is disaggregated into four categories, which are food, clothing, alcohol and tobacco, and others, similar with Chen (2006). Note that the expenditure on food also includes agricultural products self-produced by the households which are consumed by themselves. The reason why I use these four commodity categories is to make the expenditure on commodity categories in the survey perfectly comparable with corresponding prices (further details of the pricing data are introduced in section 2.4.2). More specifically, in the survey, households are invited to state their expenditure on

[6] To guarantee the representativeness of the sample households in the survey, the households in the CHFS survey are chosen by the following process: in stage one, randomly selecting cities/districts for each province; in stage two, randomly selecting communities in urban areas and villages in rural areas in each city/districts chosen in stage one; in stage three, randomly selecting households in each community and village chosen in stage two.

various general commodities in the previous year. However, it is found that there exist some discrepancies between the definitions for certain commodity categories in the survey and those in the pricing data. Taking the category of healthcare as an example, fitness expenditure is also included in this category, while for the pricing index, fitness is included in the category of entertainment. Only for the category of food, alcohol and tobacco, as well as clothing, are they perfectly consistent. Therefore, the total expenditure is disaggregated into food, clothing, alcohol and tobacco, and "others" which includes the expenditure on all other categories.[7]

For the demographic variables used in this study, two different ways are used to measure household composition characteristics. The first is where household members are split into adults and children by their ages (above 18 or not), and then the household composition is measured by the number of adults and children in what I will refer to as the baseline model. Furthermore, children are further divided into different groups by age and gender, respectively. More specifically, in other model specifications, children are divided into four age-cohorts: 0-2; 3-6; 7-12 and 13-17 years. I also distinguish female children from male children in order to analyse whether there exists gender inequality among children in China. The second measure of household composition characteristics uses a set of dummies variables corresponding to different household types, which include A1C0 (1 adult; the reference household), A1C1 (1 adult and 1 child); A2C0 (2 adults),

[7] There are many zero shares of alcohol and tobacco in the survey data. Concerned about the impacts of those 0 values on the estimation process, we also estimate a three-commodity QUAIDS model by incorporating expenditure on alcohol and tobacco into "others". The results are not very different from those obtained from the four-commodity QUAIDS model. Thus, the results from the three-commodity QUAIDS model are not shown in this chapter.

A2C1 (2 adults and 1 child), A2C2 (2 adults and two children), A3C0 (3 adults), A3C1 (3 adults and 1 child) A3C2 (3 adults and 2children), A4C0(4 adults), A4C1(4 adults and 1 child), and A4C2 (4 adults and 2 children). Similar to Chen (2006), other less common types of households are left out in this study, which takes accounts for 3,209 (out of 40,011) observations in the survey.

In addition, three groups of regional dummies are also included in this study, considering the significant differences in economic conditions, food tastes, geography, etc., between households located in different regions in China. The first group of regional dummies is urban and rural dummy. The second group includes four economic region dummies: east, middle, west, and northeast.[8] According to the status of development, China can be divided into four economic regions: east, middle, west, and northeast, with eastern areas most developed and western areas relatively under-developed. In addition, I also include the north/south dummies to control the difference between the north and south residents, as well as the impact of China's central heating policy, called the Huai River policy. There exists a significant difference between the north and south China, such as climate, customs, the dietary habitat, etc. In particular, residents in northern cities could receive central heating and heating subsidies in winter, which often starts from November to March every year, while the central heating does not exist in southern cities. This is due

[8] Four economic regions in China:
East area: Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan;
Middle area: Shanxi, Anhui, Jiangxi, Henan, Hubei and Hunan
West area: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang
Northeast area: Heilongjiang, Jilin and Liaoning

to the central heating policy started in the 1950s.[9] Table 2.2 displays the definitions of all variables.

2.4.2 Pricing data

To estimate the QUAIDS model, the price of each commodity category is also required. In this study, I use the provincial level consumer price indices of each commodity as prices in the demand system. These are collected by the National Bureau of Statistics. The price index in 2010 are assumed to be 100 in order to calculate the price indices in 2016.

As introduced above, household expenditure is aggregated into four categories: food, clothing, alcohol and tobacco, and others. The price indices for food, clothing, as well as alcohol and tobacco are directly calculated given 2010 as the base year. For the price of "others", it is obtained by treating the general CPI for each province as the weighted average of the four (or three) commodity categories, using the expenditure share of each category as the weight.[10]

The scaling function used in this study to incorporate household demographic variables is defined by Eq. (2.10) - (2.12), as shown in section 2.3. Once the prices for all commodities are set to be equal, Eq. (2.12), the logarithm of $\boldsymbol{\phi(p, z, u)}$ will equal to 0, and

[9] The central heating policy was implemented in 1950s. Due to the extreme shortage of energy at that time, the government decided to install central heating systems only in the northern cities where is colder in winter. The dividing line between the north and south is defined by the Qin Mountains and the Huai River (Qin-Huai line), and thus, this policy is also called the Huai River policy. It is noteworthy that only residents in the cities in northern China receive central heating, while residents in rural areas in the north do not receive central heating. Therefore, the Huai River heating policy could affect the consumption patterns of households by affecting their expenditure on heating.

[10] Taking Beijing as an example, for the four-commodity demand system, the general CPI in 2016 is 118.18, the pricing indices for food, clothing, alcohol and tobacco are 120.394, 109.62, and 107.702, respectively. The expenditure shares of food, clothing, alcohol and tobacco, as well as "others" are 0.42, 0.04, 0.06 and 0.48, respectively. Therefore, the price index of "others" is $(118.18 - 120.394 \times 0.42 - 109.62 \times 0.04 - 107.702 \times 0.06)/0.48 = 118.266$.

thus the scaling function will just equal to $\bar{m}_0(\boldsymbol{z})$ [Eq. (2.13)]. Therefore, I further set the pricing indices for all commodities in Beijing to be 100, and the relative pricing indices for the remaining 28 provinces are then calculated.

Table 2.2 Variable definitions

Variable	Definition
Food	Expenditure share of food, including home-made agricultural products consumed by the households
Alcohol & tobacco	Expenditure share of alcohol and tobacco
Clothing	Expenditure share of clothing and footwear
Others	Expenditure share of all other commodities, including local transport; communication; water, electricity and fuel; entertainment; education; tourism; healthcare.
Demographic variables	
NA	Number of adults (aged above 18) in the household
NMA	Number of male adults in the household
NFA	Number of female adults in the household
NC	Number of children (aged 0-17)
NC1	Number of male children aged 0-2
NC2	Number of male children aged 3-6
NC3	Number of male children aged 7-12
NC4	Number of male children aged 13-17
NMC	Number of male children
NFC	Number of female children
Urban (Reference)	=1 if the household resides in urban area; =0 otherwise
Rural	=1 if the household resides in rural area; =0 otherwise
West (Reference)	=1 if the household resides in the west; =0 otherwise
Middle	=1 if the household resides in the middle; =0 otherwise
East	=1 if the household resides in the east; =0 otherwise
Northeast	=1 if the household resides in the northeast; =0 otherwise
South (Reference)	=1 if the household resides in the south; =0 otherwise
North	=1 if the household resides in the north; =0 otherwise

2.4.3 Descriptive Statistics

The sample averages of all variables used in the analysis are reported in Table 2.3. column (1)-(3) shows the averages for the whole sample, subsample of urban and rural households, respectively. Column (4) displays the difference of the sample means of each variable between urban and rural households. The mean of the total expenditure of the

households in the survey is RMB50,715.89 (about USD7,500). For the structure of households' consumption, the expenditure share on "others" is the largest, which is reasonable because this category is defined by the aggregation of a large number of commodities. Food expenditure accounts for about 43% of household total expenditure. Alcohol and tobacco, as well as clothing, on the other hand, only account for relatively small proportions of the total expenditure, 6% and 4%, respectively. For the household composition variables, the mean household size for the whole sample is 2.91. In particular, the average number of adults and children within the household in the sample are 2.49 and 0.42, respectively. For other household demographic variables, 30% of the sample households reside in rural areas and 70% in urban ones. Those households located in the north (45%) are slightly fewer than those located in the south. In addition, about 45% households are located in east China which is the most developed region in China, and households residing in central and northeast China account for 19% and 13%, respectively. About 23% of the sample households are from the western China, the most under-developed area.

Table 2.3 Sample means of variables (whole, urban and rural)

Variable	Whole	Urban	Rural	Diff. between urban and rural
Sample size	29,087	20,322	8,765	
Total Expenditure (Yuan)	50,715.89	58,235.25	33,281.96	24,953.29***
share of food	0.427	0.433	0.413	0.021***
share of alcohol & tobacco	0.06	0.051	0.08	-0.029***
share of clothing	0.04	0.042	0.038	0.003***
share of others	0.473	0.474	0.469	0.005**
Household Demographics				
Household size	2.909	2.826	3.104	-0.278***
NA	2.487	2.424	2.632	-0.208***
NC	0.422	0.401	0.471	-0.07***
NC1	0.036	0.039	0.03	0.009***
NC2	0.098	0.097	0.098	-0.001
NC3	0.159	0.149	0.182	-0.033***

NC4	0.13	0.117	0.161	-0.045***
NMC	0.233	0.219	0.267	-0.048***
NFC	0.189	0.182	0.204	-0.022***
Rural	0.299	0	1	-
North	0.448	0.452	0.439	0.013**
Middle	0.188	0.159	0.253	-0.093***
East	0.451	0.495	0.347	0.148***
Northeast	0.131	0.135	0.122	0.013***

***, ** and * indicate significance at 1%, 5% and 10% respectively.

It can also be seen that the mean of the total expenditure of urban households is much greater than that of rural households. The average household size for urban households is lower than that for rural households, and for each group of household members, the average number of household occupants in urban areas is significantly fewer than those in rural, except for children aged 3-6.

Table 2.4 displays the means of expenditure shares for different types of households covered in this analysis. Childless couples account for the largest proportion, while only a very few of the sample households are with a single adult plus a child. The mean of total expenditure is the highest for the household with two adults and one child, while the food share for this type of household is the lowest.

Table 2.4 Sample means of expenditure shares by household types

	A1C0	A1C1	A2C0	A2C1	A2C2	A3C0	A3C1	A3C2	A4C0	A4C1	A4C2
Sample size	2,498	160	10,099	2,765	1,055	4,997	1,919	552	1,838	2,014	1,190
Share to whole sample	0.086	0.006	0.347	0.095	0.036	0.172	0.066	0.019	0.063	0.069	0.041
Total Expenditure (Yuan)	31,308	39,732	41,631	66,186	63,704	55,208	59,294	53,998	54,238	64,326	59,874
share of food	0.467	0.435	0.468	0.378	0.404	0.413	0.383	0.401	0.394	0.388	0.385
share of alcohol & tobacco	0.041	0.03	0.062	0.052	0.055	0.064	0.059	0.058	0.067	0.065	0.071
share of clothing	0.029	0.036	0.029	0.053	0.049	0.049	0.048	0.045	0.048	0.048	0.048
share of others	0.463	0.499	0.44	0.517	0.492	0.474	0.51	0.496	0.492	0.498	0.496

2.5 Results

2.5.1 Estimation results for simple QUAIDS model

The parameters of the QUAIDS model are estimated by iterated nonlinear seemingly uncorrelated regression method, which takes into account the correlations that exist across different share equations in the demand system in order to produce consistent estimates. To avoid obtaining the singular error-covariance matrix, one of the expenditure share equations has to be dropped from the demand system. Throughout the standard errors are clustered by provinces. The consumption patterns of households located in the same region could potentially be affected by the same factors, such as climate, local customs, etc. However, we can only tell which province each household is located in from the survey data, and thus the standard errors are clustered at province level.

Table 2.5 displays the estimated coefficients of the simple QUAIDS model without incorporating any demographic variables. The coefficients on the linear (β) and quadratic expenditure (λ) are statistically significant for all commodity groups. The statistical significance of both the linear and quadratic expenditure values suggests that the total consumption budget plays an important role in determining households' consumption patterns. It also implies that the Engel curves of expenditure share with respect to the logarithm of total expenditure are non-linear for each commodity. To check whether the specification of the QUAIDS model is more appropriate than the AIDS model in this study, a Wald test (for λ's are significantly different from 0 or not) is conducted, which

suggests that the null hypothesis of linear Engel curves (AIDS) is rejected in favour of the QUAIDS model.[11]

Table 2.5 Estimation results of the simple QUAIDS model (without household demographics)

Variable	symbol	Food	Alcohol & tobacco	Clothing	Others
Constants	α_i	0.152***	0.007	0.039***	0.802***
		(0.035)	(0.008)	(0.006)	(0.024)
Expenditure	β_i	-0.229***	-0.042***	-0.006**	0.276***
		(0.011)	(0.003)	(0.002)	(0.010)
Quadratic expenditure	λ_i	-0.039***	-0.008***	-0.003***	0.050***
		(0.003)	(0.001)	(0.001)	(0.003)
P_food	γ_{i1}	0.082	0.057	-0.010	-0.129
		(0.169)	(0.051)	(0.030)	(0.132)
P_alcohol & tobacco	γ_{i2}		-0.013	-0.011	-0.033
			(0.049)	(0.038)	(0.040)
P_clothing	γ_{i3}			0.035	-0.014
				(0.031)	(0.031)
P_others	γ_{i4}				0.176**
					(0.083)

***, ** and * indicate significance at 1%, 5% and 10% respectively; standard errors clustered at province level are reported in parenthesis.

Table 2.6 displays both the price and expenditure elasticities of each commodity group for the whole sample calculated from the estimated coefficients of the simple QUAIDS model shown in Table 2.5. The expenditure elasticities are all positive and significant at the 1% level. As can be seen from Table 2.6, with a 1% increase in total expenditure, expenditure on food, alcohol and tobacco, clothing, as well as others would increase by 0.811%, 0.787%, 1.117% and 1.188%, respectively. The own-price elasticities for all commodity groups are negative, which is as expected. However, none of the price elasticities are significant at the 10% level, except for the commodity category of "others", due to the insignificance of parameters for price variables.

[11] The Wald test rejects the null hypothesis of AIDS model in favour of the QUAIDS model, with the test statistic and corresponding p-value equals to 493.13 and 0.000, respectively.

There are very few studies that employ the AIDS/QUAIDS model to analyse household demand for broad commodity categories in China, which makes it difficult to make comparisons between the elasticities estimated in this study and other studies. Fan et al. (1995) use the AIDS model for five general commodity groups (i.e., food, clothing, fuel, housing, and others) to analyse household demand in rural China based on provincial level data, mainly focusing on food commodities. Similar to Fan et al. (1995), Li et al. (2020) also estimate the AIDS model for eight commodity groups (i.e., electricity, food, clothing, housing, appliances, medical, transportation and education) using provincial level household expenditure data in China, but they focus on analysing household electricity demand. Food and clothing, the two commodity categories included in this analysis, are also covered in these two analyses. In Fan et al. (1995), household food consumption also includes items produced by households, and their estimated own-price elasticities for food and clothing are -0.526 and -0.248, respectively, while the expenditure elasticities are 0.707 and 0.694, respectively. In Li et al. (2020), the estimated uncompensated own-price elasticities for food and clothing are -0.078 and -1.311, respectively, while the expenditure elasticities for the two goods are 0.843 and 0.826, respectively. It can be seen that the estimated price elasticities for food and clothing are different between this and the two previous studies. This can be due to many reasons, such as different definitions of commodity categories, different datasets, and different model specifications. My estimation of food price elasticity is closer to Fan et al., (1995) since household self-grown food items consumed by themselves are both included in household food consumption.

Besides, my estimation of the expenditure elasticity for food lies between the two studies, whilst the estimated expenditure elasticity for clothing is larger than both previous studies.

Table 2.6. Price and expenditure elasticities from simple QUAIDS model

	Price elasticity				Expenditure elasticity
	Food	Alcohol & Tobacco	Clothing	Others	
Uncompensated					
Food	**-0.843**	0.124	-0.017	-0.074	0.811***
	(0.399)	(0.118)	(0.071)	(0.31)	(0.007)
Alcohol & Tobacco	0.899	**-1.236**	-0.18	-0.27	0.787***
	(0.843)	**(0.818)**	(0.633)	(0.682)	(0.016)
Clothing	-0.322	-0.287	**-0.145**	-0.363	1.117***
	(0.747)	(0.934)	**(0.769)**	(0.778)	(0.021)
Others	-0.228	-0.058	-0.035	**-0.868*** **	1.188***
	(0.281)	(0.086)	(0.067)	**(0.177)**	(0.006)
Compensated					
Food	**-0.497**	0.173	0.015	0.309	
	(0.399)	(0.118)	(0.071)	(0.311)	
Alcohol & Tobacco	1.235	**-1.189**	-0.148	0.102	
	(0.845)	**(0.819)**	(0.633)	(0.68)	
Clothing	0.155	-0.22	**-0.1**	0.165	
	(0.749)	(0.934)	**(0.769)**	(0.777)	
Others	0.28	0.013	0.013	**-0.306**	
	(0.281)	(0.086)	(0.067)	**(0.177)**	

***, ** and * indicate significance at 1%, 5% and 10%, respectively; standard errors are reported in parenthesis.

2.5.2 Estimation results for household equivalence scales

Table 2.7-2.10 present the results of the QUAIDS model with different specifications of the scaling function. Each table displays the results for the whole sample, as well as the subsample of urban/rural households. To save space, only the results of the scaling parameters are displayed.

Table 2.7 shows the estimate equivalence scales with the household members just divided into children (NC) and adults (NA) according to whether the age of a household member is over 18 or not. The single-adult household is chosen as the reference household, which explains why the number of adults is reduced by 1 in the scaling function. Compared

Table 2.7 Estimation results for QUAIDS model with scaling function

$$m_0 = 1 + \rho_1(NA-1) + \rho_2 NC + \rho_3 Rural + \rho_4 North + \rho_5 Middle + \rho_6 East + \rho_7 Northeast$$

	Whole	Urban	Rural
$\rho_1_(NA-1)$	0.185***	0.207***	0.266***
	(0.023)	(0.029)	(0.061)
ρ_2_NC	0.108***	0.071**	0.344***
	(0.025)	(0.025)	(0.073)
ρ_3_Rural	-0.415***		
	(0.038)		
ρ_4_North	-0.133***	-0.06	-0.343***
	(0.049)	(0.064)	(0.078)
ρ_5_Middle	0.015	-0.017	0.062
	(0.055)	(0.068)	(0.11)
ρ_6_East	0.262***	0.211***	0.538***
	(0.07)	(0.073)	(0.166)
$\rho_7_Northeast$	0.161**	0.141*	0.279*
	(0.067)	(0.08)	(0.165)
Obs.	29,087	20,322	8,765

***, ** and * indicate significance at 1%, 5% and 10% respectively; standard errors clustered at province level are reported in parenthesis.

to the single-adult household, a two-adult household need 18.5% more income in order to obtain the same utility level as the reference household. Adding a child into a single-adult household increases its cost of living by about 11%. Given the same composition of household, rural households cost much less than the urban households to obtain the same level of welfare. It is interesting to note that the cost of an additional child for rural households is much greater than that for urban households. This might be explained by the fact that housing expenditure, which is likely to have large effect on the scale is excluded from household total expenditure. The housing cost of an additional child for an urban household is much higher than that for a rural household due to the high housing prices in urban areas in China.

In addition, households in north China would spend about 13% less if they resided in south China. For rural households, living in the north the total cost would decrease by about 34%, while there is no significant difference in the cost of living between northern and southern urban households. Besides, compared to the household in the west, the cost would significantly increase if they were located in the east, the more developed area in China to reach the same utility level, while the total cost would not change significantly if they were located in the middle or northeast.

Table 2.8 shows the estimated equivalence scales with children (NC) in the household divided into different age cohorts. The cost of an additional infant (aged 0-2) is the highest, followed by children aged 13-17 and 3-6, while the cost of children aged 13-17 and 3-6, while the cost of children aged 7-12 is lowest. More specifically, for a reference household containing only one adult, the cost of an infant (aged 0-2) is about 17% that of an adult. The cost of a child aged 3-6 is about 13% of that of an adult. With a child aged 7-12 entering the household, the total cost need increase by about 6%, while the cost rises by about 14% for the household containing an adult plus a child aged 13-17. For rural households, the household with an adult plus a child aged 0-2, 3-6, 7-12 or 13-17 need spend about 75%, 49%, 25% or 34%, respectively in order to reach the same level of utility as the reference single-adult household. For urban households, the household equivalence scales for children are lower than those for rural households for each of the four age cohorts.

Table 2.8 Estimation results for QUAIDS model with scaling function

$m_0 = 1 + \rho_1(NA - 1) + \rho_2 NC1 + \rho_3 NC2 + \rho_4 NC3 + \rho_5 NC4 + \rho_6 Rural + \rho_7 North + \rho_8 Middle + \rho_9 East + \rho_{10} Northeast$

	Whole	Urban	Rural
ρ_1_NA	0.187***	0.215***	0.249***
	(0.024)	(0.029)	(0.061)
ρ_2_NC1	0.173***	0.143*	0.748**
	(0.059)	(0.076)	(0.251)
ρ_3_NC2	0.133***	0.086**	0.489***
	(0.047)	(0.039)	(0.18)
ρ_4_NC3	0.063**	0.003	0.252***
	(0.032)	(0.026)	(0.077)
ρ_5_NC4	0.136***	0.117**	0.336***
	(0.037)	(0.048)	(0.107)
ρ_6_Rural	-0.416***		
	(0.038)		
ρ_7_North	-0.133***	-0.061	-0.34***
	(0.049)	(0.064)	(0.082)
ρ_8_Middle	0.02	-0.01	0.07
	(0.055)	(0.067)	(0.109)
ρ_9_East	0.262***	0.213***	0.53***
	(0.07)	(0.072)	(0.164)
ρ_{10}_Northeast	0.165**	0.146*	0.282*
	(0.067)	(0.079)	(0.165)
Obs.	29,087	20,322	8,765

***, ** and * indicate significance at 1%, 5% and 10% respectively; standard errors clustered at province level are reported in parenthesis.

Household equivalence scales can also be used to check whether there still exists gender inequality against female children in China, which is due to the son preference tradition. This can be achieved by comparing the equivalence scales of families with different compositions. Basically, if the household need a similar compensation for the arrival of a boy or girl to keep household welfare unchanged, there is no gender inequality; on the contrary, if families need a significant higher compensation for the arrival of a boy than that of a girl, it implies that there exists gender discrimination against girls. In order to do this, children are further divided by the gender into male children (NMC) and female children (NFC).

The results do not provide strong evidence on gender discrimination in China. From Table 2.9, it can be seen that the equivalence scale for one adult plus a boy is close to that for one adult plus a girl. In urban areas, the cost of the arrival of a boy rises by the similar amount of that of the arrival of a girl. For rural households, the equivalence scale for a boy is slightly larger than that for a girl, but the difference is relatively small (about 4%) and insignificant.

Table 2.9 Estimation results for QUAIDS model with scaling function

$m_0 = 1 + \rho_1(NA-1) + \rho_2 NMC + \rho_3 NFC + \rho_4 Rural + \rho_5 North + \rho_6 Middle + \rho_7 East + \rho_8 Northeast$

	Whole	Urban	Rural
$\rho_1(NA-1)$	0.187***	0.208***	0.264***
	(0.023)	(0.029)	(0.061)
$\rho_2 NMC$	0.101***	0.058**	0.361***
	(0.026)	(0.029)	(0.08)
$\rho_3 NFC$	0.117***	0.086***	0.319***
	(0.032)	(0.03)	(0.09)
$\rho_4 Rural$	-0.416***		
	(0.038)		
$\rho_5 North$	-0.132***	-0.059	-0.344***
	(0.049)	(0.064)	(0.085)
$\rho_6 Middle$	0.015	-0.018	0.063
	(0.055)	(0.069)	(0.099)
$\rho_7 East$	0.262***	0.212***	0.538***
	(0.07)	(0.073)	(0.166)
$\rho_8 Northeast$	0.161**	0.14*	0.278*
	(0.067)	(0.08)	(0.165)
Obs.	29,087	20,322	8,765

***, ** and * indicate significance at 1%, 5% and 10% respectively; standard errors clustered at province level are reported in parenthesis.

By using the number of members in different gender-age groups to measure household composition as shown in the scaling functions defined in above regressions, it implicitly assumes the marginal effects of adding one more person into the household to be the same. This, however, may not be true especially when considering the existence of household economies of scale, which means that the household's total expenditure does not need to be doubled as its household size doubles, in order to keep household welfare level unchanged. This is because some commodities can be shared between household members.

Therefore, the cost of second child might be less than the first child, because some commodities (such as toys) can be shared between the two children, and thus total expenditure on children does not need to be doubled with the arrival of the second child. Therefore, to see whether there indeed exists household economies of scales, household composition is then measured by a set of dummies for different types of households: A1C0, A1C1, A2C1, A2C2, A3C0, A3C1, A3C2, A4C0, A4C1, and A4C2.

The results are shown in Table 2.10, which provide strong evidence regarding the existence of household economies of scale. More specifically, using the household of a single adult as the reference, the scale value rises by about 36% if a new child enters the household. it needs to increase by around 56% for the second adult, while for the arrival of the third adult, the equivalence scale just rises by about 16%. For households with two adults, the scale value increases by 18% for the first child, while the marginal effect of the second child on the scale value is just 14%. Economies of scales are also found in the households with three and four adults.

Table 2.10 Estimation results for QUAIDS model with scaling function:

$m_0 = 1 + \rho_1 A1C1 + \rho_2 A2C0 + \rho_3 A2C1 + \rho_4 A2C2 + \rho_5 A3C0 + \rho_6 A3C1 + \rho_7 A3C2 + \rho_8 A4C0 + \rho_9 A4C1 + \rho_{10} A4C2 + \rho_{11} Rural + \rho_{12} North + \rho_{13} Middle + \rho_{14} East + \rho_{15} Northeast$

	Whole	Urban	Rural
ρ_1_A1C1	0.36**	0.359**	0.691
	(0.154)	(0.163)	(0.584)
ρ_2_A2C0	0.557***	0.673***	0.454***
	(0.06)	(0.068)	(0.135)
ρ_3_A2C1	0.739***	0.748***	1.482***
	(0.066)	(0.088)	(0.322)
ρ_4_A2C2	0.877***	1.000***	1.235***
	(0.109)	(0.13)	(0.352)
ρ_5_A3C0	0.717***	0.824***	0.832***
	(0.077)	(0.106)	(0.216)
ρ_6_A3C1	0.826***	0.913***	1.111***
	(0.097)	(0.117)	(0.278)

ρ_7_A3C2	0.95***	1.007***	1.527***
	(0.148)	(0.193)	(0.363)
ρ_8_A4C0	0.733***	0.787***	0.97***
	(0.089)	(0.099)	(0.218)
ρ_9_A4C1	1.126***	1.196***	1.686***
	(0.111)	(0.145)	(0.412)
ρ_{10}_A4C2	1.242***	1.091***	2.11***
	(0.138)	(0.142)	(0.449)
ρ_{11}_*Rural*	-0.485***		
	(0.045)		
ρ_{12}_North	-0.124**	-0.06	-0.353***
	(0.058)	(0.073)	(0.096)
ρ_{13}_Middle	0.033	-0.014	0.08
	(0.067)	(0.092)	(0.118)
ρ_{14}_East	0.314***	0.261***	0.594***
	(0.084)	(0.085)	(0.188)
ρ_{15}_Northeast	0.185**	0.163	0.333*
	(0.085)	(0.104)	(0.175)
Obs.	29,087	20,322	8,765

***, ** and * indicate significance at 1%, 5% and 10% respectively; standard errors clustered at province level are reported in parenthesis.

2.5.3 Estimation results from Engel's method

The estimated equivalence scales from the alternative, the Engel's method are shown in Table 2.11-2.13 (for full results of the Engel's model, see Appendix 2). The equivalence scales are estimated for both the whole sample households, as well as the subsample of urban and rural households. In table 2.11, household individuals are simply divided into adults and children. Taking a single-adult household as the reference, with the presence of a child, the household's total expenditure needs to rise by 36% in order to leave the welfare level unchanged, which is similar with that estimated from the QUAIDS model (see Table 2.10). The equivalence scale for another adult added into the reference household is 0.34, which is lower than the QUAIDS scale. It is noticeable that the estimated Engel's scales for households with three adults and one or two children are somewhat unexpectedly smaller than the scales for 3 and 4 adult households. This might be due to the problems

of the method itself, which are discussed in section 2. The equivalence scale for adults is similar between urban and rural households, while the scales for children in rural households are much greater than urban households.

Table 2.11 Engel's equivalence scales: individuals divided into adults and children

Household type	Equivalence scale		
	Whole	Urban	Rural
A1C0	1	1	1
A1C1	1.356	1.245	1.645
A2C0	1.343	1.357	1.281
A2C1	1.366	1.267	1.58
A2C2	1.234	1.051	1.732
A3C0	1.353	1.381	1.23
A3C1	1.223	1.147	1.349
A3C2	1.036	0.892	1.387
A4C0	1.211	1.249	1.051
A4C1	1.026	0.972	1.08
A4C2	0.835	0.726	1.066

Table 2.12 reports the estimated equivalence scales from the Engel's method with children divided into different groups according to their ages. It can be seen that the estimated equivalence scale for adults is lower than that for children. The cost of an additional child aged 3-6 is similar with that of an additional infant (aged 0-2), which are 1.65 and 1.63, respectively. The equivalence scales of one more child aged 7-12 and 13-17 are relatively lower. Besides, the cost of one more child in each age groups added to a 2-adult household in rural are greater than those in urban.

Table 2.12 Engel's equivalence scales: children divided into different age cohorts

Household composition	Equivalence scale		
	Whole	Urban	Rural
1 adult (reference)	1	1	1
1 adult, 1 child aged 0-2	1.651	1.651	1.651
1 adult, 1 child aged 3-6	1.634	1.634	1.634
1 adult, 1 child aged 7-12	1.538	1.538	1.538
1 adult, 1 child aged 13-17	1.448	1.448	1.448
2 adults	1.343	1.35	1.267
2 adults, 1 child aged 0-2	1.404	1.309	1.894

2 adults, 1 child aged 3-6	1.417	1.343	1.632
2 adults, 1 child aged 7-12	1.366	1.255	1.563
2 adults, 1 child aged 13-17	1.292	1.182	1.48
2 adults, 2 child aged 0-2	1.305	1.129	2.517
2 adults, 1 child aged 0-2, 1 child aged 3-6	1.317	1.158	2.169
2 adults, 1 child aged 0-2, 1 child aged 7-12	1.269	1.082	2.077
2 adults, 1 child aged 0-2, 1 child aged 13-17	1.2	1.019	1.967
2 adults, 2 child aged 3-6	1.329	1.188	1.869
2 adults, 1 child aged 3-6, 1 child aged 7-12	1.281	1.11	1.79
2 adults, 1 child aged 3-6, 1 child aged 13-17	1.211	1.045	1.695
2 adults, 2 child aged 7-12	1.104	1.037	1.713
2 adults, 1 child aged 7-12, 1child aged 13-17	1.167	0.976	1.623
2 adults, 2 child aged 13-17	1.104	0.92	1.537

To investigate the existence of gender discrimination using the Engel's method, children are also divided into two groups by gender. The estimated Engel's equivalence scales are reported in Table 2.13, which are broadly consistent with the results from the QUAIDS model shown in Table 2.9. There are no significant differences between the equivalence scales for male and female children in urban households, while for rural households the equivalence scale for male children is relatively larger than that for female children but the difference is not statistically significant.

Table 2.13 Engel's equivalence scales: children divided into male and female children

Household composition	Equivalence scale		
	whole	urban	rural
1 adult	1	1	1
1 adult, 1 male child	1.368	1.245	1.699
1 adult, 1 female child	1.331	1.234	1.575
2 adults	1.343	1.357	1.281
2 adult, 1 male child	1.378	1.267	1.632
2 adult, 1 female child	1.341	1.256	1.513

2.6 Discussion

This section focuses on making comparisons between this and previous studies which also employ the demand system model to estimate household equivalence scales. I first compare my results with Chen (2006), the only study in China using the demand system

model to estimate household equivalence scales, for the purpose of measuring the poverty line for households with different compositions. Chen (2006) estimates the AIDS model for four commodity categories, food, clothing, tobacco and alcohol, as well as others, which are the same groups used in this this. My results suggest that adding the second adult raises the cost of living of the single-adult household by 56%, which is close to Chen's estimation (53%). For the household with one adult and one child, my estimation of the equivalence scale is 1.36, while with children split into different age cohorts, Chen estimates the equivalence scales for the household with one adult and one child aged 0-2, 3-6, and 7-18, which are 1.46, 1.39 and 1.33, respectively, which are also similar with the findings in this study.

Although the existing literature has not used the QUAIDS model to estimate household equivalence scales with Chinese data, it has been applied for such analyses in other countries. Table 2.14 displays household equivalence scales for other countries found in previous studies. The single-adult household is taken as the reference in all these studies, except for Balli and Tiezzi (2010), which instead use two-adult household as the reference in their analysis. The findings in this study fall somewhere in the middle, compared with these studies. The equivalence scales are lower in Australia, as estimated by Blacklow et al., (2010), while Michelini (2001) and Dudel et al., (2020) report greater estimates for New Zealand and Germany, respectively. The estimated equivalence scales in this study are closer to those obtained by Dudel et al., (2020). Note also that the household composition dummies used in this study are set in the similar way with Dudel et al., (2020) according to different number of adults and children in the household, while for Michelini (2001) and

Blacklow et al., (2010), children are split into different age groups when estimating household equivalence scales for different types of households. Michelini (2001) divides children into infants (aged 0-2), young children (aged 3-7) and teenage children (aged 8-14), and the author estimates the equivalence scales for two adults and an infant, as well as two adults with two infants, which are 1.98 and 2.36, respectively. The equivalence scale for households with two adults, one infant, and one young child is 2.43 (which is not displayed in the Table). Children are split into three age groups, 0-4, 5-14, and 15-23 in Blacklow et al., (2010). They report equivalence scales in their paper for households with two adults and an infant (aged 0-4), two adults and one infant and one child aged 5-14, which are 1.53 and 1.69, respectively. Furthermore, one needs to be careful when comparing household equivalence scales across these studies as different methods are applied to incorporate household demographic characteristics into the demand system among these studies. For instance, Blacklow et al., (2010) also employ Ray's scaling method, but their scaling function specifications are totally different from these used in this study. Besides, as noted above, children are split into different age cohorts across these studies.

Table 2.14 Comparisons with previous QUAIDS model applied in different countries

	Country	A1C0	A2C0	A2C1	A2C2
Balli and Tiezzi (2010)	Italy	-	1.00	1.2	1.41
Michelini (2001)	New Zealand	1.00	1.54	1.98	2.36
Blacklow et al. (2010)	Australia	1.00	1.36	1.53	1.69
Dudel et al. (2020)	Germany	1.00	1.58	1.76	2.11
This study	China	1.00	1.56	1.74	1.88

Note: A1C0 refers to households with single-adult, A2C0 refers to households with two adults, A2C1 refers to households with two adults and one child, and A2C2 refers to households with two adults and two children. For Michelini (2001), A2C1 is the household with two adults and one infant (aged 0-2), A2C2 is the household with two adults and two infants; For Blacklow et al. (2010), A2C1 is the household with two adults and one infant (aged 0-4), and A2C2 is the household with two adults and one infant, and one child aged 5-14.

2.7 Conclusion

Using China Household Financial Survey (CHFS) data, this study estimates household equivalence scales from the estimation of the QUAIDS model for four commodity categories. The equivalence scales for different household types are computed. There exist big differences in the equivalence scales between urban and rural households. The findings also provide evidence on the existence of significant economies of scales in household consumption expenditure. All of these results can be applied to compare the welfare level among households with various sizes and compositions, as well as to modify the poverty line for different types of households.

The findings in this study also provide some policy insights. Since China implemented the three-child policy in 2021, more and more local governments have decided to give families subsidies to encourage couples to have children. For example, the government provides RMB500 a month for the second child and RMB1,000 a month for the third child until the child is 3 years old in Harbin; In Hangzhou, the government provides couples with a one-time subsidy of RMB7,000 and RMB25,000 for their second and third child, respectively. The subsidy policies are different between different regions, such as the amount of subsidy, and how to pay subsidies to couples (by year, by month, one-time payment, etc.). The estimated equivalence scales in this study can be used by the policy makers to determine the amount of subsidy to couples to encourage them to have more children. Besides, since in many cities, the subsidies are mainly provided to the second and the third child, considering the existence of economies of scale in household expenditure,

the cost of the first child in a family should be the highest. Therefore, the government should also consider offering subsidies to couples to encourage them to have their first child. On the other hand, the estimated equivalence scales can also be applied to adjust the current poverty standards considering the cost of living are different between households with different sizes and/or different compositions.

The main limitation of this study is related to the price aggregation. In order to make commodity categories and their pricing perfectly consistent, household total expenditure is disaggregated into four categories, which are food, tobacco and alcohol, clothing, and others. The prices for the former three commodity groups are measured by the corresponding pricing indices, while the price for "others" is approximated by using the weighted average of pricing indices. This approximation can be potentially problematic, and it is hard to tell how this approximation of pricing would affect the estimation of household equivalence scales. In future work, one may consider analysing household equivalence scales in China by using subjective approaches, which enables a comparison with the findings from expenditure-based approach used in this study.

Chapter 3 Household size, economies of scale in household energy use in India

Abstract

India has been experiencing a decrease in average household size over the past few decades. Given that previous studies have suggested the existence of economies of scale in household energy use, shrinkage of household size might result in a loss of economies of scale in energy consumed by households in India. In this study, I investigate the economies of scale in household energy consumption by using a dataset taken from a nationally representative household survey in India. In order to do this, two different but not unrelated concepts are introduced for analysis. The empirical results indicate that there exist significant economies of scale in household energy expenditure in India. In particular, the economies of scale vary across different sorts of energy. With one extra person added into the household, per capita expenditure on modern energy, i.e., electricity and liquefied petroleum gas (LPG) decline by 24% and 26.5%, respectively, which are greater than that for traditional energy. Based on the empirical findings, It is projected that a 0.5 person decline in household size by around 2030 would result in a 12% and 13% increase in per capita expenditure on electricity and LPG.

3.1 Introduction

Household economies of scale means that households' cost of living does not need to increase proportionately as the number of household members increases, keeping household utility level constant. This can arise when households with multiple members share public goods, and thus making larger households better off at the same level of per capita resources (Deaton and Paxson, 1998). Here public goods refer to those that can be easily shared among household members. Energy is one good example of public goods, considering that households consume energy goods mainly for the purposes such as cooking, lighting, heating and cooling. Similarly, researchers in existing literature define the economies of scale in household energy consumption as household energy expenditure does not need to increase proportionately with one more person added into a household. Some studies find the evidence on the impact of household size on household energy use (such as Karimu et al., 2016; Song et al., 2018; Mulenga et al., 2019; Kostakis, 2020). However, it is worth emphasizing at the beginning that although the term "economies of scale" is consistently used in these studies (which will be reviewed in section 3.2.2), its underlying measurement or definition are different from one study to another.

Many countries in the world are experiencing a reduction in average household size (United Nations, 2019). This is also the case in India, the most populous country in the world, which has recently overtaken China in April 2023, according to the latest estimates of global population from the United Nations. The average household size in India increased steadily until it reached its peak (5.5) in 1991 and it has fallen sharply since then.

Over 2001-2011, the average household size decreased from 5.3 to 4.9 (Census of India, 2001; 2011). More specifically, India experienced a decadal growth rate of 17.7% for its entire population, while the total number of households rose by 28.9%, which corresponded to the decline in average household size over 2001-2011. Table 3.1 displays the percentage change of households by household size over 2001-2011, from which we can see a significant decrease in the share of households with 7 or above members, while the share of households with 1 to 6 members all increase, in particular for households with 2 and 3 members. Nayak et al. (2014) find that household size varies across different regions in India.

Table 3.1 Percentage change in households by household size during 2001-2011

Household size	2001	2011
1	3.9%	4.2%
2	8.2%	10.5%
3	11.7%	13.4%
4	19.2%	20%
5	18.6%	19.1%
6	13.6%	13.8%
Above 7	24.8%	19%
Total	100%	100%

Data source: Census of India, 2001 and 2011; calculated by authors

India is the third largest carbon emitters in the world, after China and the United States. India has been experiencing a significant increase in carbon emissions over the last few decades. Figure 1 shows India's total carbon emissions has increased from 530 million tons in 1990 to 2,075 million tons in 2020. Residential energy consumption is one of the main sources of greenhouse gas (GHG) emissions, the chief cause of global climate change. Therefore, knowledge of the economies of scale in energy consumed by households can

be used to investigate the impact of the decline in household size on GHG emissions, which is the main motivation of this research.[12]

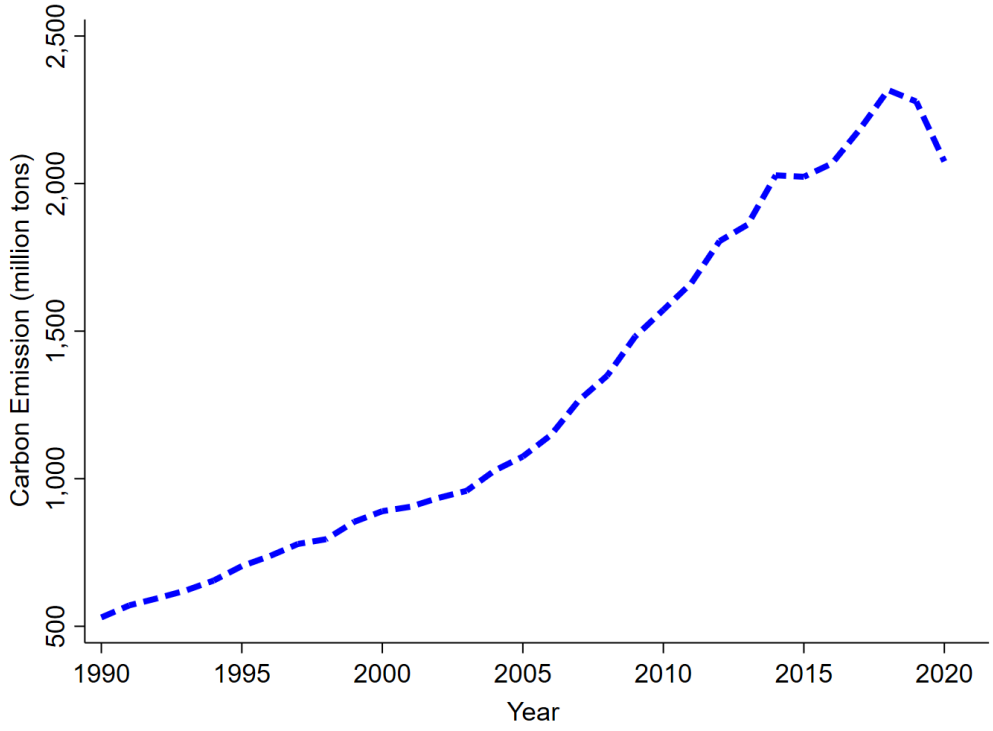


Figure 3.1 India's total carbon emissions from 1990-2020
Data source: International Energy Agency (IEA)

This study aims to quantify the household economies of scale in energy consumption in India. In order to do this, I investigate two different but not unrelated concepts. More specifically, following Kakwani and Son (2005), I attempt to estimate the extent to which energy consumption considered by the household to be a public rather than a private good. Furthermore, I also estimate the per capita elasticity of demand with respect to household size for different types of energy by employing the double hurdle model, including traditional fuels (firewood and dung cake), modern fuels (electricity and liquefied

[12] Another motivation of this study is that in addition to GHG emissions, residential energy use can also lead to indoor air pollution, in particular when traditional solid fuels, such as firewood and animal wastes are burned. In India, there were over 300 million Indian people unable to access to clean fuels in 2015, the majority of which reside in rural areas (Maji et al., 2021).

petroleum gas (LPG)), as well as transition fuel (kerosene) which stands somewhere in between. The results from Kakwani-Son's method suggest that holding household utility level unchanged, household energy expenditure would increase by 0.7% with a 1% increase in household size, indicating that there exist economies of scale in household energy consumption. The results from the double hurdle model show that: first, the factors influencing the household's decision to consume each type of energy differ from those influencing how much is spent; second, the estimated elasticities of per capita expenditure with respect to household size are the largest for modern fuels, while those for traditional fuels are smallest. With one more person added into the household, per capita expenditure on electricity and LPG would decrease by 24% and 26.5%, respectively, while that on firewood would decrease by only 3.7%. The per capita expenditure on dung cake would even increase by 9.3%; third, the results also provide empirical evidence on the energy ladder hypothesis that with the increase in household income, households would move from traditional solid fuels to modern and cleaner energy.

This study contributes to existing literature in two ways. First, this study enriches the scare literature on investigating the economies of scale in household energy consumption in India by using two different measures of economies of scale. Second, this is the first study to examine household economies of scale for different sorts of energy in India, and the estimation results provide empirical evidence that the scale effects vary across different types of energy.

The remainder of this chapter is organized as follows. Section 3.2 reviews the relevant literature. Section 3.3 and 3.4 introduce the data and methodology used in this study. Section 3.5 and section 3.6 present the estimated results and discussion. Section 3.7 concludes this chapter.

3.2 Literature review

In this section, I first review the previous studies that have investigated the determinants of household energy consumption in India. Then I move on to review the studies that focus on examining economies of scale in household energy consumption in any country. Besides, considering the methodology used in this study, I also review the studies that employ Kakwani-Son's method to investigate household economies of scale, as well as those using the double-hurdle model to analyse household energy consumption. Finally, I provide a critical assessment and discussion of these literature reviewed in this section.

I mainly rely on Econlit to search for the relevant literature.[13] More specifically, different combinations of keywords are used to search for the literature which are as follows: "household energy" and "India" (60 results); "household size" and "energy consumption or "energy demand" or "energy expenditure" (52 results); "household economies of scale" and "energy" (15 results). By using those combinations for searching, some relevant studies

[13] The Econlit database provides a comprehensive library of economics literature which covers economics literature published over 130 years from leading institutions in 74 countries. Therefore, the resulting literature by using Econlit are more targeted.

are obtained. Based on these studies, some other studies are further found either from their lists of references, or from those which have cited them.

3.2.1 Determinants of household energy consumption in India

Those factors which are found to significantly affect households' energy consumption in the literature can be summarized into three categories: socio-demographic factors (e.g., household size, household income), building factors (e.g., number of rooms, dwelling type), as well as appliance factors (e.g., number of refrigerators) (Jones et al., 2015). Considering the extensive body of existing research on the determinants of household energy consumption, I focus on reviewing solely those studies conducted in India. Among these studies, some of them focus on investigating the factors affecting by how much energy consumed by Indian households, while more studies focus on analysing the factors that affect (the probability of) Indian households' adoption of some type of energy, in particular, cleaner energy (i.e., LPG and electricity).

I find three research that analyse the factors affecting household energy consumption. Pachauri (2004) conducts a rather unusual study, which explores the factors affecting household energy requirements, based on the cross-sectional survey data from India's Household Consumption Expenditure Survey (CES) in 1993-1994. Household energy requirements in this study are calculated by multiplying household expenditure for each consumption category by the corresponding energy intensity value taken from the

estimation from Pachauri and Spreng (2002).[14] A set of dummy variables corresponding to household size are included in the analysis. The results show that household per capita energy requirements decrease with one more person added into the household, except for the case where household size increases from one to two, which is very odd, but the author does not provide any explanation.

Singh et al. (2018) analyse the variation in electricity consumption by urban households in India and the factors resulting in those variations. They conduct a questionnaire survey to collect household level data. 1,026 households in total are surveyed, located in different cities in three different climatic regions of India, viz., hot-dry, warm-humid, as well as a composite climate region. Those three regions cover the majority of both the total geographical area and population of India. The factors affecting household electricity consumption considered in their model are comprehensive, including dwelling characteristics, demographic variables, ownership of home appliances (air conditioner, refrigerator and water heater), as well as climate regions each household resides in (represented by a set of dummy variables). The results indicate that household size has a significant positive effect on household electricity consumption. With an extra person added, household electricity expenditure increases by 4.7%. As expected, the effects of the ownership of appliances on electricity consumption are positive. In particular, the effect

[14] Pachauri and Spreng (2002) estimate energy intensities for 100 producing sectors in India. In order to calculate household energy requirements, these estimates are used by Pachauri (2004) in conjunction with the micro survey data on household expenditures. Household energy requirements are measured by the sum of indirect and direct energy requirements. Indirect energy requirements are calculated by multiplying household expenditure for each consumption category (except for energy) by the corresponding energy intensity value, while direct energy requirements are obtained by first converting the consumption quantities of each type of fuels in the survey into uniform energy units (MJ) and then multiplied by energy intensity values for the energy sector.

of the ownership of an air conditioner is much stronger than those of demographics and other appliances. One problem of this research however is that single-adult households are excluded from their survey without any explanation.

Conevska and Urpelainen (2020) examine the effect of weather on rural households' electricity consumption. Household energy consumption data is collected from a one-year experiment between 2016-2017, which is conducted in a very small area, 7 habitations (which are set up by an Indian energy service company) in one district of the state of Uttar Pradesh in India. Households covered in the experiment have no access to electricity before the study. By participating, households in the habitations are provided electricity only for essential needs from the installation of solar micro grids. Their analysis only includes weather variables and a set of seasonal dummies, without any household demographic characteristics. The estimation results suggest that temperature plays a significant role in household electricity consumption. With 1°F increases in temperature, weekly electricity consumption will increase by about 5.2Wh on average. Precipitation has a significant positive effect on household electricity consumption, while the effect of humidity is negative.

More studies in the case of India focus on studying the factors that affect households' probability of adopting some type of energy. Among them, Rao and Reddy (2007) and Mottaleb and Rahut (2021) conduct their analysis for both traditional and cleaner modern energy. Rao and Reddy (2007) empirically investigate the determinants of household choice of a particular type of energy using India's Household CES survey data in 1999-2000. The

energy consumed by sample households are divided into four categories: firewood, LPG, kerosene, and other fuels. A multinomial logit model is estimated for rural and urban households separately in this study. Firewood is chosen as the base category and thus the estimated coefficients of the model for the other three categories of energy measures the effect of a change in each corresponding factor on the probability that households will choose the alternative energy relative to the probability of choosing firewood. A group of factors considered to affect the probability of choosing each type of energy are included in the model, such as per capita total expenditure (to proxy income), household size (measured by the number of household members and its square), religion, education level, and other household demographic characteristics. The results indicate that total expenditure and household size exert a non-linear relationship on the probability of choosing cooking fuels.

Mottaleb and Rahut (2021) employ the probit model, which is different from the method used by Rao and Reddy (2007) to investigate the factors affecting urban households' choice of fuel based on the CES survey data in 1991-1992, 2006-2007 and 2011-2012. The probit model is estimated for the choice of three different types of fuels, which are firewood, kerosene, as well as cleaner fuels (LPG and electricity). Their results suggest that household size has a significant and positive effect on both the probability of choosing each of the three fuels, as well as their shares to total energy expenditure. Education and economic status measured by household total expenditure play significant roles in determining the choice of fuels. More specifically, the education level of both

household heads and their spouse has significant positive effect on the adoption of cleaner energy, and the higher the educational level of household heads and their spouse, the lower the probability of adopting dirty energy (such as firewood) and fuels. Similarly, the economic household total expenditure also has significant positive effect on the choice of cleaner energy and negative effect on the adoption of less clean energy.

Kishore and Spears (2014), Saxena et al. (2018), as well as Swain and Mishara (2020), on the other hand, only analyse the factors affecting households' adoption of cleaner energy, but they focus on different factors. Kishore and Spears (2014) focus on examining the effect of the sex of the first child on the adoption of clean cooking fuels. Two possible mechanisms are analysed in this study. First, having a son could increase household status of women who do more cooking than men, and thus increase the possibility of adopting clean cooking fuels which are healthier than traditional fuels. Second, Indian households prefer healthier boys than girls and thus they are more likely to adopt clean fuels rather than traditional fuels for cooking. The results suggest that the probability of adoption of clean fuels is about 2% higher for urban households with a male first child, while the effect of sex of the first child is insignificant for rural households. Household size is also included in their model as a control variable, the effect of which, however, is not reported. The analysis in this study is mainly based on the linear probability model. It is a simple model and coefficients of which are easy to interpret, but at same time it also has some problems, one of which is the estimated probability can be negative or over unity.

Using the CES survey data in 2011-2012 (the same dataset used in this study), Saxena et al. (2018) estimate a probit model to examine the effects of household's caste, tribe and religion on the access to LPG. Meanwhile, a simple OLS model is also used to analyse the effects of these factors on household electricity usage measured in terms of kWh. The authors do not estimate the same probit model for the access to electricity considering that access to electricity is almost universal in India. The sample households are categorized into four social groups, which are scheduled caste, scheduled tribe, Muslim, and upper caste. Similar to other studies in the literature, household total consumption expenditure is also used to proxy household income. Compared to upper caste households, households in the scheduled tribe, scheduled caste and Muslim have significant lower probability of access to LPG and electricity usage.

Swain and Mishara (2020) investigate the factors affecting the adoption of LPG as cooking fuel for rural households in the state of Odisha, who are the beneficiaries of one scheme introduced by Indian government, namely Pradhan Mantri Ujjwala Yojana, which provides subsidized connections of LPG to women members of households living below the poverty line. One survey is conducted for only 106 rural households. which collects data on the sources of cooking energy as well as social demographic characteristics. The adoption of LPG is measured indirectly by the decline in the use of firewood, the main source of cooking energy in the sample households. The results show that the education of household heads has a significant positive impact on the adoption of LPG, while the probability and the extent of LPG adoption are both higher for upper caste households.

This analysis however, is based on the survey information collected from only 106 households. Besides, the author fails to explain why they choose the decline in the use of firewood over 50 percent rather than a lower or higher percentage as the proxy of LPG adoption.

3.2.2 Economies of scale in household energy consumption

Among the existing literature on the determinants of household energy consumption, some studies treat household size as the primary focus to investigate the economies of scale effect which are directly related to this study. Meanwhile, there are also some studies, in which although household size is not the focus, their results also provide clear evidence on the economies of scale in household energy consumption. Both sorts of studies are reviewed in this section and are summarized in Table 3.2. It is worth highlighting that although the term "economies of scale" appears consistently in these studies, its underlying measurement or concept differs from one study to another. Therefore, the speech marks for this term will be used during the review of these studies.

Some studies use the negative impact of household size on per capita energy consumption as the measure of "economies of scale" in household energy use. Ironmonger et al. (1995) investigate "economies of scale" in energy use for adult-only households in Australia. The "economies of scale" in energy use for households with a different number of occupants is measured by the difference of mean annual energy use from the reference household (two-adult household). The results suggest that compared to the two-adult household (reference household), single-adult households suffer diseconomies of scale

while significant "economies of scale" are found in three- and four-adult households. More specifically, single-adult households spend $345 per adult on average compared to $247 per adult for two-adult households, while per adult energy expenditure are $224 and $190 for three- and four-adult households, respectively.

O'Neill and Chen (2002) conduct a case study on demographic factors affecting household energy consumption in the U.S. based on the cross-sectional data drawn from two national surveys in 1993-1994, the Residential Energy Consumption Survey and the Residential Transportation Energy Consumption Survey. The authors analyse the impact of householder age, household composition, as well as household size on residential and transportation energy use, respectively. For household composition which is measured by the number of children and the number of adults, the results suggest that compared to households with adults only, households with children consume 44%, 35% less residential and transportation energy per capita, respectively. They also find that per capita energy use decreases with household size. More specifically, compared to single-person households, per capita energy use is 17% less for two-person households, while for three-person households, per capita energy use is 30% lower than that for single-person households.

Using the cross-sectional data on 300,000 households in the Netherlands, Brounen et al. (2012) examine the impact of dwelling and demographic characteristics on residential gas and electricity consumption. Household size is found to have a significantly negative impact on per capita gas and electricity consumption. More specifically, with each additional person added into the household, per capita gas and electricity consumption

decreases by about 26% and 18%, respectively, which confirms the existence of "economies of scale" in household energy consumption in the Netherlands. Furthermore, they also examine the impact of children on energy consumption: with one more child added into the household, per capita consumption on gas and electricity decreases by about 23% and 17%, respectively. The author, different from other studies in which both of these two types of characteristics are incorporated into one regression, examines the impact of dwelling and demographic characteristics in two separate models. As a robustness check, they also estimate the model which includes both sets of characteristics and there is no big difference in estimated coefficients.

Using household panel data from the UK, Longhi (2015) investigates the impact of dwelling and household demographic characteristics on household energy expenditures using three waves of data from the 2009-2012 UK Household Longitudinal Survey. Three separate analyses are conducted for household expenditures on all types of energy, electricity only and gas only. The results suggest that household size has a negative impact on per capita energy expenditure, indicating the presence of "economies of scale" in residential energy consumption. In particular, the "economies of scale" diminishes as household size increases. For example, moving from a 1 to 2-person, 2 to 3-person, 3 to 4-person household decreases per-capita energy expenditure to about 47%, 40%, 33% and 26%, respectively. Similar results are also found when analysing expenditure on electricity and gas separately.

Wu et al. (2021) estimate the impact of household size on household per capita electricity consumption based on household-level panel data drawn from China Family Panel Studies. Their results suggest that household size has significantly negative impact on per capita household electricity consumption, which is interpreted by the author as the existence of "economies of scale" in household electricity consumption in China. With the presence of each additional person within the household, electricity expenditure per capita would decrease by 18.4%. Moreover, household members are also divided into adults and children, and it is found that the scale effects on per capita electricity expenditure is 19.6% and 15.4% for adults and children, respectively. The relatively smaller scale effect for children is explained by the possible fact that an extra child needs a separate room and thus makes the "economies of scale" from children weaker. Considering that electricity generation in China is predominantly coal-fired, the authors further estimate the environmental impact of decreasing "economies of scale" due to the shrinking of household size. Taking the decline in average household size to be 0.5 as an example, other things equal, electricity-related CO_2 emissions will increase by 42.2-58.7 million tons, while water withdrawal will increase by 1,543-2,142 million m^2. Smoke and ash, SO_2, NO_X emission will increase 4.4-6.2, 22.2-30.8 and 19.9-27.7 million tons, respectively. Similar to Wu et al. (2021), this study is also motivated by the facts on the shrinking household size, as well as the environmental concerns related to household energy consumption in India.

In addition to using the negative impact of household size on per capita energy consumption as the measure of household "economies of scale" in energy consumption,

there are some other studies employing different measures. Schroder et al. (2015) examine household "economies of scales" in residential energy-related expenditure in Japan. Energy-related expenditure refers to total expenditure on electricity, gas, water, and sewage, since it is unable to disaggregate further from the survey data used in their analysis. Similar to Pachauri (2004), they also use a set of dummy variables for household size, but these dummies are defined in a different way. The authors first use the dummies of DN^n (=1 if the number of members is n or larger. Their results suggest that the increase in total energy-related expenditure for the second household member is smaller than that for the first member, and that for the third is smaller than the second, and so on.

Kavousian et al. (2013) analyse the determinants of residential electricity consumption in the US based on a high-resolution electricity consumption data obtained from a dataset for 1,628 households subject to smart-metering, which is collected at 10-min intervals over 238 days. Their results show that the square root of household size has a significant and positive impact on households' daily electricity consumption, which is interpreted as the existence of "economies of scale" by the authors. More specifically, as the number of household size doubles, household electricity consumption increases at a lower rate (1.4 times). However, the authors fail to justify the use of the square root of household size. One possible reason is that they used this because household size is not the focus of their analysis. Besides, due to the large number of explanatory variables, the stepwise selection method is employed to determine those variables to include in the analysis, which can be problematic, such as overfitting the data and biased estimation (Smith, 2018).

Mori and Yepez-Garcia (2020) investigate the determinants of household energy expenditure based on a cross-sectional data of 183,000 households obtained from national household expenditure surveys in 13 Latin American countries. They focus on estimating the relationship between energy spending and income, as well as the existence of "economies of scale" in household energy expenditure. The number of household members in different age groups (aged below and above 12) within the household and their square terms are added into their regression model. The results suggest that the linear terms for the number of children (aged below 12) and the number of people older than 12 are both positive, suggesting that greater household size is associated with higher energy spending, while their quadratic terms are significantly negative, which are interpreted by the authors as proving the existence of "economies of scale" in household energy expenditures.

Table 3.2 Household economies of scale in energy consumption from previous studies

Sources	Country	Period	Commodity	Economies of scale
Ironmonger et al., (1995)	Australia	1985-1986	Electricity	Compared to 2-adult household, energy expenditure per adult is 40% higher, 9% and 23% lower for single-adult, three- and four-adult households, respectively.
O'Neill and Chen (2002)	US	1993-1994	Gross energy	Compared to the single-person households, the per capita energy use is 17% and 30% less for two-person and three-person households, respectively.
Pachauri (2004)	India	1993-1994	Gross energy	Using a set of dummies variables of household size, it is found that household per capita energy requirements decrease with one more person added into the household.

Brounen et al., (2012)	Netherlands	2008-2009	Gas Electricity	With each additional person added into the household, per capita gas and electricity consumption decreases by about 26%, 18%, respectively.
Kavousian et al., (2013)	US	2010	Electricity	The square root of household size has a significant and positive impact on households' daily electricity consumption, which is interpreted as the existence of economies of scale by the authors.
Schroder et al., (2015)	Japan	2004-2012	Gross energy	Household size is measured by a set of dummy variables, and substantial economies of scale are found in residential energy related expenditure.
Longhi (2015)	UK	2009-2012	Gross energy	Household size has a negative impact on per capita to energy expenditure, suggesting the presence of economies of scale in residential energy consumption
Mori and Yepez-Garcia (2020)	13 Latin American countries	2004-2013[15]	Gas Electricity Transport fuel	The quadratic terms of household size have negative impact on household energy , which are interpreted as the existence of economies of scale
Wu et al., (2021)	China	2010-2016	Electricity	Electricity expenditure per capita decreases by 18.4% with an extra person added.

3.2.3 Empirical applications of Kakwani-Son's method

Kakwani and Son (2005) develop a general framework to estimate both household overall economies of scale, as well as the economies of scale for individual commodities consumed by households. Their intuition of economies of scale in one commodity is that an increase in λ percent of household size requires the increase of less than λ percent of expenditure on that commodity, (i.e., the elasticity of expenditure on this good with respect

[15] Different surveys for these 13 Latin American countries are conducted in different period of time from 2004-2013.

to household size need to be less than 1), to maintain the same level of utility given household total income constant. The details of Kawani-Son's framework will be introduced in Section 3.3. To empirically estimate the elasticities of household expenditure on different commodities with respect to household size, they employ a Working-Leser demand system model, which is a simple linear demand system model based on a cross-sectional household survey data in Australia. Household total expenditure is disaggregated into 13 commodity categories, one of which is fuel and power.[16] The results show that with 1% increase in household size, expenditure on fuel and power only increases by 0.5%, indicating existence of economies of scale in households' energy consumption in Australia.

Using Kakwani-Son's method, Mok et al. (2011), and Karbasi and Mohammadzadeh (2018) investigate household economies of scale in Malaysia and Iran, respectively. Since there is no one commodity category for energy in both studies, it cannot be known whether economies of scale exist or not for household expenditure on energy goods. One recent study by Rahman et al. (2020) examine household size economies for household consumption goods in Bangladesh by using Kakwani-Son's method. Household total expenditure is divided into food, clothing, education, health, transportation, shelter, utilities, and miscellaneous goods. They find that the economies of scale exist in household utility expenditure. As household size increases by 1%, utility expenditure rises by only 0.5%.

[16] These commodity categories include food; alcohol; tobacco; fuel and power; clothing and footwear; housing cost; housing furnishing; housing services; medical and services; transport and communication; education and recreation; personal care; miscellaneous expenditure.

3.2.4 Applications of double-hurdle model in household energy consumption

The double-hurdle model which is a variation of tobit model can be used to investigate whether factors influencing the probability of fuel choice differ from those affecting fuel expenditures. It assumes that two separate hurdles must be passed before the positive expenditure is observed. The first hurdle is the decision of spending or not (participation decision), and the second hurdle is the decision of how much to spend. More details of the double-hurdle model are introduced in Section 3.4.

Adusah-Poku and Takeuchi (2019) employ the double-hurdle approach to identify the determinants of household energy expenditures in Ghana based on the Ghana Living Standards Survey in 2012-2013. They mainly focus on households' expenditure on LPG and charcoal, two main cooking fuels in Ghana. The results suggest that the factors considered have different effects on the probability of participating in LPG or charcoal market, and actual amount of LPG or charcoal expenditure. For example, household size is negatively related to the probability of participating in LPG market, but positively related with LPG expenditure conditional on LPG market participation. Having a male as head of the household negatively affects only charcoal expenditure conditional on charcoal market participation. As well as the double-hurdle model, a Heckman model is also employed in their analysis as a robustness check. The Heckman model is a better choice if the zero expenditures result from non-observable responses. It assumes that the participation decision dominates the expenditure decision, which implies that zero expenditures reflect

households' decision to not participate in the market. Results from Heckman model are very similar to those from the double-hurdle model.

There are two very recent empirical applications of the double-hurdle model analysing the factors affecting household energy expenditure in India. Gould et al. (2020) investigate the factors influencing both the probability of participating in the LPG market and LPG expenditure for rural households based on a panel data obtained from the survey called Access to Clean Cooking Energy and Electricity in 2015 and 2018. It is found that household size, total expenditure, education level of the household head, and households belonging upper caste category have a significant and positive relationship with both the probability of adopting LPG and LPG expenditure conditional on adoption. Daripa and Dinda (2022) employ the double-hurdle model to examine the effects of a household's socio-economic factors on each of the two decision stages in firewood expenditure for rural households based on CES survey data in 2011-2012. Household size increases the probability of selecting firewood but decreases firewood expenditure conditional on adoption.

3.2.5 Literature discussion

Looking back over the studies reviewed in this section, it is found that: first, while there are a few studies on investigating the factors affecting household energy consumption in India, more studies focus on analysing the factors that affect Indian households' adoption of some type of energy. For research on the determinants of energy consumed by households in India, most of them are about investigating the factors influencing the

adoption of cleaner energy, such as LPG and electricity, but what matters is not merely the adoption of particular forms of energy but the extent to which they are used and how it changes with household size.

Second, together with household size, other household socio-demographic characteristics such as household income and the demographic characteristics of the household head (e.g., age, education, marriage), are frequently incorporated into the analysis of the determinants of household energy consumption. Some studies directly added income (or its logarithm) into the analysis (e.g., Ironmonger et al., 1995; Meier and Rehdanz, 2010; Brounen et al., 2012; Kavousian et al., 2013; Longhi, 2015; Wu et al., 2021), while other studies just use total expenditure to proxy household income (e.g., Pachauri, 2004; Schroder et al., 2015; Mori and Yepez-Garcia, 2020). Together with household socio-demographic characteristics, dwelling characteristics such as dwelling size, number of rooms, type of dwellings and appliance factors such as the dummy variables for the ownership of different types of home appliances (e.g., refrigerator, air conditioner) are also included in previous studies.

Third, as what we have seen, the economies of scale in household energy consumption is defined by researchers as households' energy consumption does not need to increase in proportion to household size. However, different studies employ different ways to measure it although the same term appears in those studies. For example, it has been indicated by the negative impact of household size on per capita energy expenditure (Ironmonger et al., 1995; O'Neill and Chen, 2002; Brounen et al., 2012; Longhi, 2015; Wu

et al., 2021), or the negative coefficient on the square of household size (Mori and Yepez-Garcia, 2020). The framework developed by Kakwani and Son (2005) provides another way to measure household economies of scale for individual commodities.

In this study, I estimate the household economies of scale in energy consumed by households in India. For this purpose, I investigate two different but not unrelated concepts. More specifically, following Kakwani and Son (2005), I attempt to estimate the extent to which energy consumption considered by the household to be a public rather than a private good. Furthermore, I also estimate the elasticity of household per capita energy demand with respect to household size. To my knowledge, this study is the first to employ Kakwani-Son's method to investigate the household economies of scale in India. Besides, although there have been two very recent studies which employ the double hurdle model to analyse the determinants of household energy consumption in India, this study is different from them in several ways. First, both studies only focus on one type of energy, LPG, or firewood, while this study focuses on different types of energy, which can be grouped into three categories: traditional fuel (dung cake and firewood), transition fuel (kerosene), as well as modern fuel (LPG and electricity), which allows us to examine whether the impact of the same factors are different or not on different types of energy. Second, only rural households in India are included in these two studies, while this analysis is conducted for both urban and rural households. Third, I focus on analysing the impacts of household size on per capita expenditure on different types of fuels, in order to investigate the existence of economies of scale in household energy consumption.

3.3 Methodology

In this study, I employ Kakwani-Son's method to measure the economies of scale in household energy consumption. Meanwhile, the double hurdle model is used to estimate the elasticity of per capita energy consumption with respect to household size for different types of fuels used by Indian households.

3.3.1 Kakwani-Son's method

Kakwani and Son (2005) develop a general framework for defining the economies of scale for individual commodities consumed by households from Barten's model (Barten, 1964) of household's utility maximization with household characteristics included. Barten (1964) analyses the effect of household composition on consumption by assuming that a household has a utility function which takes the form:

$$u = u\left(\frac{q_1}{m_1}, \frac{q_2}{m_2}, \dots, \frac{q_n}{m_n}\right) \tag{3.1}$$

where q_i is the quantity consumed for the ith good, and m_i is the function of household composition vector $\mathbf{z}$, which measures the effect of household composition on utility:

$$m_i(\mathbf{z}) = m_i(z_1, z_2, \dots, z_D) \tag{3.2}$$

where z_d is the number of household members in demographic category d ($d = 1,2,\dots,D$), and thus the household size $h = \sum_{d=1}^{D} z_d$.

Kakwani and Son's intuition of the economies of scale for the ith commodity is that an increasing in α percent of members in each of the D demographic categories requires the

increase of less than α percent of the consumption of good i, holding household utility level unchanged.

According to microeconomic theory, the maximization of Barten's utility function (Eq. 3.1) under the budget constraint yields the Marshallian demand function for good i:

$$q_i = m_i g_i(x, p_1^*, p_2^*, \dots, p_n^*) \tag{3.3}$$

where x is total expenditure. Substituting (3.3) into (3.1) gives the cost function:

$$x = e(u, p_1^*, p_2^*, \dots, p_n^*) \tag{3.4}$$

and by substituting (3.4) into (3.3), the Hicksian demand for good i can be derived as follows:

$$q_i^* = m_i g_i^*(u, p_1^*, p_2^*, \dots, p_n^*) \tag{3.5}$$

where q_i and q_i^* are Marshallian and Hicksian demand for good i, respectively, and $p_i^* = m_i p_i$.

Differentiating the Hicksian demand function with respect to z_d yields the elasticity of q_i^* with respect to z_d:

$$\eta_{id}^* = \delta_{id} + \sum_{j=1}^{n} \varepsilon_{ij}^* \, \delta_{jd} \tag{3.6}$$

where η_{id}^* is the utility compensated elasticity of the ith good with respect to the number of members in the dth composition category. δ_{id} is the elasticity of m_i with respect to z_d and ε_{ij}^* is the compensated price elasticity of good i with respect to the price of good j.

Eq. (3.6) suggests that a change in the number of members in the dth composition category has both income and substitution effects on the consumption of good i, which is because an increase in z_d affects the effective prices of all commodities. The income effect is reflected by δ_{id}, which is identical to all households irrespective of their utility level and household composition, while the substitution effect is reflected by $\sum_{j=1}^{n} \varepsilon_{ij}^{*} \delta_{jd}$, which in contrast depends on both utility level and household composition.

Summing η_{id}^{*} over all composition categories gives the index of the economies of scale for the ith good, ϕ_i^{*}:

$$\phi_i^{*} = \sum_{d=1}^{D} \eta_{id}^{*} = \theta_i + \sum_{j=1}^{n} \varepsilon_{ij}^{*} \theta_j \tag{3.7}$$

where $\theta_i \equiv \sum_{d=1}^{D} \delta_{id}$. If $\phi_i^{*} = 1$, the ith good does not provide economies of scale to the household; if $\phi_i^{*} < 1$, the ith good provides economies of scale to the household, and in particular, if $\phi_i^{*} = 0$, the ith good is a perfect public good. In addition, the ith good can also be diseconomies of scale if $\phi_i^{*} > 1$.

Kakwani and Son also derive the overall index of economies of scale from the cost function in Eq. (3.4). There exist the economies of scale in a household if an increasing in α percent of members in each of the D demographic categories requires the increase of less than α percent of total income or cost, keeping the household staying at the same level of utility. Thus, the overall index of economies of scale, denoted by ϕ^{*}, equals to the sum of the elasticity of cost with respect to the number of members in category d over all composition categories:

$$\phi^* = \sum_{d=1}^{D} \frac{\partial e}{\partial z_d} \frac{z_d}{e} \tag{3.8}$$

The elasticity of cost with respect to the number of members in category d can be obtained by differentiating (3.4) with respect to z_d:

$$\frac{\partial e}{\partial z_d} \frac{z_d}{e} = \sum_{i=1}^{n} w_i \, \delta_{id} \tag{3.9}$$

where w_i is the expenditure share of good i. Therefore, the overall index of economies of scale takes:

$$\phi^* = \sum_{d=1}^{D} \sum_{i=1}^{n} w_i \, \delta_{id} = \sum_{i=1}^{n} w_i \, \theta_i \tag{3.10}$$

(3.10) suggests that the overall index of economies of scale is the weighted average of the indices of economies of scale for individual commodities, with the weight equal budget shares of each commodity.

The economies of scale indices shown in (3.7) and (3.10) are related to ε_{ij}^* , the elasticities of Hicksian demand which however cannot be observed. Only Marshallian demands can be observed and estimated from household consumption data. Hicksian demand elasticities can nevertheless be obtained from Marshallian demand elasticities through the Slutsky equation:

$$\varepsilon_{ij}^* = \varepsilon_i w_i + \varepsilon_{ij} \tag{3.11}$$

where ε_i is expenditure elasticity, ε_{ij} is Marshallian price elasticity of good i to the price of good j.

Substituting (3.11) into (3.7) yields:

$$\phi_i^* = \varepsilon_i \phi^* + \theta_i + \sum_{j=1}^{n} \varepsilon_{ij}\, \theta_j \tag{3.12}$$

It is obvious that price information is needed to calculate Marshallian price elasticities (ε_{ij}) in (3.12). However, Kakwani and Son show that the economies of scale index can also be obtained even we do not know the prices paid by households.

Differentiating the Marshallian demand function (3.3) with respect to z_d yields the elasticity of demand for good i to the number of members in composition category d:

$$\eta_{id} = \delta_{id} + \sum_{j=1}^{n} \varepsilon_{ij}\, \delta_{jd} \tag{3.13}$$

Letting $\phi_i = \sum_i \eta_{id} = \theta_i + \sum_{j=1}^{n} \varepsilon_{ij}\, \theta_j$ and substituting into (3.12) yields:

$$\phi_i^* = \varepsilon_i \phi^* + \phi_i \tag{3.14}$$

(3.14) is the main result of Kakwani-Son's model, based on which we can estimate both the overall and individual commodity's economies of scale, i.e., ϕ_i^* and ϕ^* from household survey data even without price information. Note that since ϕ^* is the weighted average of ϕ_i^* (Eq. 3.10), an additional assumption on the nature of commodities is required in order to estimate the index of economies of scale.[17]

Following Kakwani and Son (2005), I also employ a simple demand system model, namely the Working-Leser demand system model to empirically estimate the economies of scale index for household energy consumption, which takes the form as follows:

[17] If it is assumed that good i is a pure private good, i.e., ϕ_i^*=1, then the overall economies of scale index can be obtained and based on which ϕ_i^* for all i can be calculated.

$$w_i = \alpha_i + \beta_i \ln x + \sum_{d=1}^{D} \gamma_{id} z_d + \mu_i \tag{3.15}$$

where w_i is expenditure share of good i, x is total expenditure, z_d is the number of members in category d, μ_i is error term, α_i, β_i and γ_{id} are parameters to be estimated. The corresponding Marshallian income elasticity of good i equals to:

$$\varepsilon_i = 1 + \frac{\beta_i}{w_i} \tag{3.16}$$

The elasticity of demand for good i with respect to z_d equals to:

$$\eta_{id} = \frac{\gamma_{id} z_d}{w_i} \tag{3.17}$$

Thus,

$$\phi_i = \sum\nolimits_i \eta_{id} = \sum\nolimits_i \frac{\gamma_{id} z_d}{w_i} \tag{3.18}$$

3.3.2 The double-hurdle model

Since there are many zero expenditures for each type of fuel consumed by households in the survey data (as will be shown in Table 3.4), ordinary least squares regression cannot be used to examine the determinants of household energy expenditure, which could result in biased estimation. households' energy expenditure is a representative example of a corner solution response, which is also often called a censored response. More specifically, households' energy expenditure has a continuous distribution over strictly positive values but also has some probability of being zero (the corner is at 0 in this case). The tobit model developed by Tobin (1958) is appropriate to analyse a censored dependent variable.

The standard tobit model can be specified as follows:

$$y_i^* = x_i\beta + \varepsilon_i, \quad \varepsilon_i \sim (0, \sigma^2) \tag{3.19}$$

$$y_i = \begin{cases} y^*, & if\ y_i^* > 0 \\ 0\ , & if\ y_i^* \leq 0 \end{cases} \tag{3.20}$$

where y_i^* is a latent variable that cannot be observed. y_i is the actual observed level of expenditure for household i. x_i is the vector of control variables. ε_i is the error term and β is the vector of parameters to be estimated.

One limitation of tobit model is that it assumes that the same factors affect the probability of a non-zero observation as well as the level of a positive observation in the same way, which is potentially too restrictive. In order to overcome the shortcoming of the tobit model, Cragg (1971) formulates the double-hurdle model, which is an extension of original tobit model. The double-hurdle model can be interpreted as the combination of probit and tobit model. It assumes that households must pass two separate hurdles before they are observed with a positive level of expenditure. The first hurdle corresponds to the decision of either positive or zero expenditure (participation decision) while the second hurdle is the decision of how much to spend conditional on having made the decision of positive expenditure (expenditure decision). The double hurdle model can be specified as follows:

$$y_{i1}^* = z_i\alpha + u_i, \text{ (Participation decision)} \tag{3.21}$$

$$y_{i2}^* = w_i\gamma + v_i, \quad \text{(Expenditure decision)} \tag{3.22}$$

$$y_i = \begin{cases} w_i\beta + v_i, & if\ y_{i1}^* > 0\ and\ y_{i2}^* > 0 \\ 0\ , & otherwise \end{cases} \tag{3.23}$$

where y_{i1}^*, y_{i2}^* are latent endogenous variables representing households' participation decision and expenditure decision, respectively, which are unable to be observed. y_i is the observed expenditure. z_i is the vector of factors affecting households' participation decision, while w_i is the vector of factors affecting households' expenditure decision. u_i, v_i are the error terms, which are assumed to be uncorrelated.

3.4 Data

3.4.1 Data source

The data used in this study is drawn from India's Household CES survey in 2011-2012, which is conducted by National Sample Survey Office, Ministry of Statistics and Program Implementation, Government of India. For each round of the CES survey, four sub-rounds are conducted throughout the year in July-September, October-December, January-March and April-June. The CES survey in 2011-2012 contains 101,662 households, covering 625 (out of 640) districts drawn from the entire geographical area of India. It provides data on household monthly consumption on a variety of commodity categories, including (1) food, beverages, alcohol and tobacco; (2) clothing and bedding; (3) energy; (4) durable goods; (5) other miscellaneous goods. Within each of the five categories, a detailed item classification is further adopted.

For the category of energy, the survey includes data on household monthly consumption of 15 types of energy (for the purpose of fuel, lighting and running household appliances), including: coke, firewood and chips, electricity, dung cake, kerosene, matches, coal, LPG, charcoal, candle, gobar gas, petrol, diesel and other fuel. According to its sources, kerosene

is divided into PDS kerosene and non-PDS kerosene. PDS Kerosene is an allocated and subsidized product, which is distributed to the customers through the Public Distribution System (PDS) network of the State Governments/Union Territories (UT). [18] Correspondingly, non-PDS kerosene, on the contrary is obtained from non-PDS sources. According to the survey, the primary source of energy for cooking for the households in the survey is firewood and LPG, while the primary source of energy for lighting is electricity and kerosene. In particular, 44.3% households in the survey mainly use firewood for cooking, while 42.8% mainly use LPG. Besides, there are also 4.7% households using dung cake for cooking. The households mainly using electricity for lighting accounts for 87.1%, while there are about 12.1% households mainly using kerosene for lighting. Considering the expenditure on all other types of energy goods are zero for most households in the sample, this study focuses on identifying the determinants of household expenditure patterns for these 6 primary fuels used for cooking and lighting, which are dung cake, firewood, PDS and non-PDS kerosene, electricity as well as LPG.

In addition, the CES survey also provides information on a wide range of household characteristics, as well as demographic characteristics of household members. Table 3.3 lists all variables included in this analysis and their definitions.

Table 3.3 Variable definitions

Variable	Definition
Food	Expenditure share on food and beverages over the last 30 days
Medical and healthcare	Expenditure share on medical goods over the last 30 days

[18] The Public Distribution System (PDS) in India is a nationwide system of predominantly third-party run Fair Price Shops (FPS) (administered at the state level), through which the central and state governments distribute subsidized food, kerosene, and other commodities on the basis of household ration card allocations.

Energy	Expenditure share on energy used for fuel, light and household appliances over the last 30 days
Clothing, bedding and footwear	Expenditure share on clothing, bedding and footwear over the last 30 days
Tobacco and alcohol	Expenditure share on tobacco and alcohol over the last 30 days
Miscellaneous	Expenditure share on miscellaneous goods and services including education, medical, entertainment, minor durables (such as umbrella), household consumables, conveyance over the last 30 days
Dependent variables for double hurdle model:	
- First hurdle:	
Dung cake	Dummy (adoption=1, otherwise=0)
Firewood	Dummy (adoption=1, otherwise=0)
PDS kerosene	Dummy (adoption=1, otherwise=0)
Non-PDS kerosene	Dummy (adoption=1, otherwise=0)
Electricity	Dummy (adoption=1, otherwise=0)
LPG	Dummy (adoption=1, otherwise=0)
- Second hurdle:	
MPC_Dung cake	Monthly per capita expenditure on dung cake (in Indian Rupee, same as follows)
MPC_Firewood	Monthly per capita expenditure on Firewood
MPC_PDS kerosene	Monthly per capita expenditure on PDS kerosene
MPC_Non-PDS kerosene	Monthly per capita expenditure on non-PDS kerosene
MPC_Electricity	Monthly per capita expenditure on electricity
MPC_LPG	Monthly per capita expenditure on LPG
Independent variables	
- Household characteristics:	
Household size	Total number of household members
Adult	The number of adults (aged above 18)
Child	The number of children (aged below 18)
Total expenditure	Household total expenditure (in Rs.)
Urban	Dummy (urban=1; rural=0)
Ration	Dummy (having ration card=1, not=0)[19]
Hindudum (Reference)	Dummy (Hinduism=1; otherwise=0)
Other Religion	Dummy (Other Religion =1; Hinduism =0)
Oucdum (Reference)	Dummy (upper classes=1; otherwise=0)
Stdum	Dummy (scheduled tribe=1, otherwise=0)
Scdum	Dummy (scheduled caste=1, otherwise=0)
Obcdum	Dummy (other classes=1, otherwise=0)
- Household head characteristics:	
Age	The age of household head in years
Male	Dummy (male head=1, female=0)

[19] Ration cards are an official document issued to households that are eligible to purchase subsidized food grain and non-food items (such as kerosene) from the public distribution system.

Edu_non (Reference)	Dummy (no formal schooling=1, otherwise=0)
Edu_primary	Dummy (primary schooling=1, otherwise=0)
Edu_high	Dummy (high schooling=1, otherwise=0)
Edu_university	Dummy (university or above=1, otherwise=0)

3.4.2 Descriptive statistics

Table 3.4 shows the descriptive statistics of all variables used in this study. Food accounts for the largest proportion (53%) in household total expenditure, followed by miscellaneous goods (22%). Per capita expenditure on each type of fuel varies significantly across sample households, with electricity, firewood and LPG having the largest variations. The sample mean of monthly expenditure is 7,564.67 Rupees. For household socio-economic characteristics, the average household size in the survey is 4.57. In particular, the average numbers of adults and children within each household are 3.013 and 1.56, respectively. 41.3% of sample households reside in urban areas while the rest 58.7% reside in rural areas. For the religious beliefs of households, Hinduism is the largest religion in India. 75.8% households in the survey are Hindus, while 24.2% households believe in other religions. Households with different religious beliefs may have different dietary and cooking habits thus affecting, e.g., their demand for fuels for cooking. For the social groups of households, 13.4%, 15.4% and 39.3% of sample households belong to scheduled tribe, scheduled caste and other classes, respectively, while the remaining 31.9% of households belong to upper classes.

For the demographic characteristics of household head, 88% of sample households are male headed. The average age of household heads is 46.6. 22.5% household heads are literate with formal schooling but below primary schooling. 40.7% of household heads

possess education levels that are above primary schooling but below undergraduate, and only 12.9% of household heads have completed undergraduate studies or above. The remaining 23.9% of them have no formal schooling experience.

Table 3.4 descriptive statistics

Variable	Mean	S.D.	Min.	Max.
Food	0.528	0.126	0	1[20]
Medical and healthcare	0.051	0.084	0	0.965
Energy	0.107	0.051	0	0.673
Clothing, bedding and footwear	0.071	0.089	0	0.745
Tobacco and alcohol	0.02	0.037	0	0.798
Miscellaneous	0.223	0.110	0	1
Dung cake	0.218	0.413	0	1
Firewood	0.607	0.489	0	1
PDS kerosene	0.547	0.498	0	1
Non-PDS kerosene	0.201	0.400	0	1
Electricity	0.876	0.330	0	1
LPG	0.499	0.500	0	1
MPC_Dung cake	4.8	12.807	0	87.3
MPC_Firewood	37.2	43.91	0	229.5
MPC_PDS kerosene	5.4	6.592	0	37.3
MPC_Non-PDS kerosene	3.2	8.979	0	86.7
MPC_Electricity	52.4	57.206	0	380.3
MPC_LPG	37.9	46.567	0	209.5
Monthly total expenditure (Rupees)	7,564.7	5,163.544	40	36,329
Household size	4.574	2.247	1	39
Adult	3.013	1.542	0	26
Child	1.560	1.451	0	17
Rural	0.587	0.492	0	1
Urban	0.413	0.492	0	1
Ration	0.796	0.403	0	1
Hinduism	0.758	0.428	0	1
Other Religion	0.242	0.428	0	1
Oucdum	0.319	0.466	0	1
Stdum	0.134	0.341	0	1
Scdum	0.154	0.361	0	1
Obcdum	0.393	0.488	0	1
Age	46.6	13.502	5	105

[20] The outliers of the variables used in the analysis are not removed considering the number of outliers are not too many. For example, there are only 10 and 7 (out of about 100,000) households in the survey that share of food expenditure equals 0 and 1, respectively. Removing those outliers or not does not have big differences in the estimation results.

Male	0.88	0.320	0	1
Edu_non	0.239	0.427	0	1
Edu_primary	0.225	0.417	0	1
Edu_high	0.407	0.491	0	1
Edu_university	0.129	0.335	0	1

3.5 Estimation results from Kawani-Son's model

The Working-Leser demand system model is estimated using the seemingly unrelated regression method. Within each household, members are divided into only two groups, adults (aged 18 or above) and kids (aged less than 18). Standard errors are clustered at the district level.[21] Table 3.5 summarizes the estimated Marshallian elasticities ε_i and ϕ_i, as well as overall economies of scale index (ϕ^*) and economies of scale indices for individual commodities (ϕ_i^*) including energy goods based on the estimated parameters of the demand system model. ε_i and ϕ_i are computed by using Eq. (3.16)-(3.18) at the sample average values of w_i and z_d, while ϕ^* and ϕ_i^* are computed by using Eq. (3.14) together with the further assumption on the nature of some commodity. More specifically, medical and healthcare good is first assumed to be purely private, i.e., $\phi_i^* = 1$ for medical and healthcare good, following Kakwani and Son's judgement in their original article. As an alternative food instead of medical goods is also assumed to be purely private according to the intuition.

[21] I also estimate the demand system model with household members divided into more detailed groups according to both gender and age. There are no big differences in estimated elasticities and economies of scale indices. The standard errors are also clustered at household level, while there are no changes in significance of either estimated parameters or economies of scale indices.

The income elasticity of a particular commodity refers to the percentage change in the demand for that commodity with a given percentage change in household income. As shown in column (1), the income elasticities are all significant and positive for each commodity group, suggesting that all of these commodities are normal goods. The income elasticities of demand for food (0.772), alcohol and tobacco (0.829), and energy (0.586) are less than one, while medical and healthcare, clothing, bedding and footwear, and miscellaneous goods are more responsive to the changes in income (with income elasticities greater than 1).

Column (2) shows the estimated Marshallian elasticities of demand for each commodity with respect to the changes in household size, which reveal how households reallocate their total expenditure among different commodities as household size changes, holding household income constant. The results indicate that as household members increase by one percent, expenditure on food and energy goods would increase by 0.17% and 0.16%, respectively, keeping household income unchanged, while expenditure on other commodity groups would decrease by 0.04% (alcohol and tobacco), 0.4% (medical and healthcare), 0.2% (clothing, bedding and footwear) and 0.33% (miscellaneous goods), respectively.

The computed economies of scale indices for all commodities are reported in column (3) and (4). Under the assumption that medical and healthcare goods are purely private, the overall index of economies of scale is 0.929 which is slightly less than 1, suggesting that there exist small economies of scale in household consumption. That is, as household size

Table 3.5 Estimated economies of scale indices using Kakwani-Son's method

			Index of Economies of scale (ϕ_i^*)	
	(1)	(2)	(3)	(4)
Commodities	Income elasticity (ε_i)	ϕ_i	Method 1	Method 2
Food	0.772***	0.174***	0.892***	1.00
	(0.004)	(0.004)	(0.011)	(-)
Alcohol and Tobacco	0.829***	-0.041*	0.729***	0.845***
	(0.024)	(0.021)	(0.023)	(0.02)
Energy	0.586***	0.163***	0.707***	0.789***
	(0.01)	(0.008)	(0.013)	(0.011)
Medical and healthcare	1.501***	-0.395***	1.00	1.21***
	(0.029)	(0.028)	(-)	(0.023)
Clothing, bedding and footwear	1.586***	-0.202***	1.273***	1.495***
	(0.025)	(0.022)	(0.026)	(0.019)
Miscellaneous	1.452***	-0.332***	1.017***	1.221***
	(0.009)	(0.008)	(0.019)	(0.01)
Overall economies of scale index			0.929***	1.07***
(ϕ^*)			(0.012)	(0.004)

* Column (1) is the income elasticity of each commodity; column (2) lists the estimated Marshallian elasticities of demand for each commodity with respect to the changes in household size; column (3) and (4) shows the economies of scale indices for each commodity, which are obtained by assuming medical goods and food goods to be purely private, respectively.
* Standard errors in parentheses clustered at district level; $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

increases by one percent, about 0.07% of total expenditure can be saved for the household staying at the same level of utility. The individual index of economies of scale are significant and less than one for food (0.892), alcohol and tobacco (0.729), and energy goods (0.707). In contrast, the index of economies of scale for clothing, bedding and footwear, as well as miscellaneous goods are greater than 1, indicating that there are diseconomies of scale in the consumption of these commodities. These results are shown in column (3). Column (4) displays the computed indices of economies of scale under the other assumption, namely that food is purely private. It can be seen that ϕ^* turns to be greater than 1, which means that there are now diseconomies of scale in household total consumption. This can be explained by that whereas food is closer to being a private good it is still not purely

private, and thus the assumption of $\phi_i^* = 1$ for food goods might be greater than its actual index of economies of scale. This then causes the computed indices of economies of scale for overall and other individual goods to be higher than their actual values. Among other individual commodities, only alcohol and tobacco, and energy goods have their own economies of scale, while medical, clothing, bedding and footwear, as well as miscellaneous goods have diseconomies of scale.

From columns (3) and (4) it can be seen that economies of scale vary across different commodity groups. Under different assumptions about the nature of a particular commodity, the computed Kakwani-Son's indices of economies of scale are also different both for individual goods and overall, which is one of the potential problems of Kakwani-Son's method. However, the rankings of goods by their economies of scale indices are at least consistent. Here we are mainly interested in whether there exist economies of scale in households' energy consumption, and these results suggest that energy has economies of scale under both assumptions. Furthermore, the index of economies of scale for energy goods is the smallest among all commodities, which means that energy is indeed, the most public commodity compared to all others.

3.6 Estimation results from the double-hurdle model

The double hurdle model provides very different information from Kakwani-Son's method, specifically on the elasticity of fuel use with respect to household size.[22] The

[22] Kakwani-Son's method can also be used to estimate the economies of scale indices for different types of energy by disaggregating the expenditure on energy into different energy commodities. I did not do like this considering the substantial zero expenditures in each of these energy goods in the survey.

estimated results from the double-hurdle model for different types of fuels and energy are presented in Table 3.6-3.11.

Two different models are estimated for each type of fuels. In the first model specification, household size is measured by the number of occupants within each household while in the second specification, household members are further divided into adults and children. Like the tobit model, the estimated coefficients in the double hurdle model are not directly interpretable. Therefore, the "margins" command in STATA is used to compute the marginal effects of each variable based on the original estimated coefficients. Each table reports the estimated marginal effects of all factors included in this analysis on the decision of households to participate in the energy market (i.e., $P(Y > 0|X)$, the probability of households' decision to participate), the expectation of per capita expenditure conditional on the participation of the market ($E(Y|X, Y > 0)$), as well as the overall marginal effect, $E(Y|X)$. Both state and time fixed effects are controlled, and the standard errors are clustered at district level, considered the error terms of households within one district may be correlated due to, for example the local climate, energy access, etc.

It can be seen that for each type of fuels, the impact of the same factors on the probability of households' decision to participate in the energy market (i.e., $P(Y > 0|X)$, shown in column (1) and (4) in Table 3.6-3.11) differs from that on household energy expenditure conditional on the participate of the market(i.e., $E(Y|X, Y > 0)$, shown in column (2) and (5) in Table 3.6-3.11). These findings suggest that the tobit model is not

appropriate in this analysis because it assumes that the variables influencing the decision to participate in a fuel market are the same as those influencing how much to spend (Adusah-Poku and Takeuchi, 2020).

Next, I will first analyse the marginal effects of household demographic characteristics on household expenditure for each type of fuels, and then focus on the factor of household size, by computing the semi-elasticities of per capita energy expenditure on each type of fuels with respect to household size. That is, the percentage change in per capita expenditure on each type of fuels with one more person added into the household. Since similar semi-elasticities are also estimated in previous studies (such as Wu et al., 2021), this allows us to make comparisons between the results estimated in this study and those in existing literature, which also explains why the semi-elasticities rather than the elasticities are computed in this study.

For the convenience of analysing the effects on each of the 6 types of fuels and energy, those fuels and energy are categorized into three groups. More specifically, dung cake and firewood are classified as tradition fuels while electricity and LPG are classified as modern fuels. Kerosene is classified as a transition fuel.

3.6.1 Effects of household demographics on household energy expenditure

3.6.1.1 Traditional fuels

Table 3.6 and 3.7 present the results for the marginal effects of household demographic characteristics on household per capita expenditure on dung cake and firewood, respectively. It can be seen that household size has significant and positive effect on the

probability of households' adoption of dung cake and firewood. With one more person added, households are 1% and 5.2% more likely to adopt dung cake and firewood for cooking, respectively. However, per capita expenditures on dung cake or firewood are negatively affected by household size if households have decided to consume dung cake/firewood. Once households have made the decision of adoption, per capita expenditure on dung cake and firewood decreases by 1.3 Rupees and 7 Rupees, respectively. The unconditional per capita expenditure on dung cake slightly increases by 0.08 Rupee, while that on firewood decreases by 1.41 Rupees with one more person added into the household. In particular, it is shown that there are no significant differences between the effect of adults and children on either the probability of consuming dung cake/firewood or how much to spend.

Turning to the effect of other household characteristics, I find a negative relationship between household total expenditure and the probability of households' choice of dung cake, which is as expected considering dung is not the most pleasant or convenient of fuels. Households' total expenditure is significantly positive correlated with per capita dung cake expenditure if households have decided to use dung cake as fuels, while the unconditional per capita expenditure on dung cake decreases with total expenditure. Similar results are also found for firewood. Households possessing ration cards have higher probability of using dung cake (2.4%) and firewood (14.5%) as fuels than those without ration cards. This is not surprising as possession of a ration card can be seen as a marker for poverty. Once households have decided to use dung cake or firewood, holding ration cards

decreases per capita expenditure on dung cake or firewood. In contrast, being in possession of ration cards increases households' unconditional per capita expenditure on dung cake and firewood. Compared with rural households, urban households have both lower probability of using dung cake or firewood as fuels and the per capita expenditure on dung cake or firewood. This might be because that it is more difficult to source these fuels in urban areas.

Households' religion and their social group also play a significant role in household expenditure on dung cake as well as firewood. Compared to Hindus, households with other religious beliefs have 1.7% lower probability to choose dung cake as fuels for cooking and 0.27 Rupees lower per capita expenditures on dung cake. In the case of firewood, I do not find a significant difference between Hindu households and households with a belief in other religions in both households' choice of and how much to spend per capita on firewood. As for social groups, compared to households in higher caste, belonging to scheduled caste or other classes increases both the probability of using dung cake and firewood and per capita expenditure on dung cake and firewood. The effect of households in scheduled tribe is quite different between the consumption of dung cake and firewood. Scheduled tribe households are less likely to use dung cake and once having decided to use dung cake, per capita expenditure on dung cake is about 2.1 Rupee less than households in upper caste. For firewood, in contrast, the probability of using firewood increases by 16.5% for scheduled tribe households. The conditional and unconditional per capita firewood expenditure also increase by 3.6 and 12 Rupees, respectively.

Table 3.6 Estimated marginal effects for traditional fuel: dung cake

	Model 1			Model 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Prob.	Cond.	Uncond.	Prob.	Cond.	Uncond.
HHsize	0.01***	-1.309***	0.08***			
	(0.001)	(0.136)	(0.011)			
Adult				0.009***	-1.319***	0.058***
				(0.001)	(0.146)	(0.013)
Child				0.011***	-1.302***	0.098***
				(0.001)	(0.14)	(0.014)
Ln_totalexp	-0.019***	2.618***	-0.138***	-0.000***	2.622***	-0.129***
	(0.003)	(0.365)	(0.05)	(0.000)	(0.367)	(0.000)
Age	-0.0002***	0.0003	-0.003***	-0.000*	0.000	-0.002*
	(0.0000)	(0.006)	(0.001)	(0.000)	(0.007)	(0.001)
Male	0.013***	-0.602**	0.172***	0.014***	-0.596***	0.184***
	(0.003)	(0.266)	(0.049)	(0.003)	(0.273)	(0.05)
Edu_primary	-0.012***	-0.706***	-0.226***	-0.012***	-0.705***	-0.223***
	(0.002)	(0.229)	(0.045)	(0.002)	(0.229)	(0.044)
Edu_high	-0.031***	-0.827***	-0.526***	-0.031***	-0.825***	-0.521***
	(0.003)	(0.233)	(0.175)	(0.003)	(0.231)	(0.064)
Edu_university	-0.074***	-1.669***	-1.223***	-0.074***	-1.666***	-1.218***
	(0.005)	(0.399)	(0.13)	(0.005)	(0.397)	(0.13)
Urban	-0.124***	-1.999***	-2.007***	-0.124***	-1.998***	-2.007***
	(0.006)	(0.353)	(0.184)	(0.006)	(0.353)	(0.062)
Ration	0.024***	-0.304	0.376***	0.024***	-0.304	0.378***
	(0.006)	(0.275)	(0.062)	(0.003)	(0.275)	(0.006)
Other religion	-0.017***	-0.252	-0.27***	-0.017***	-0.254	-0.275***
	(0.005)	(0.43)	(0.083)	(0.005)	(0.429)	(0.083)
Stdum	-0.011*	-2.134***	-0.293**	-0.011*	-2.135***	-0.295***
	(0.007)	(0.729)	(0.119)	(0.007)	(0.73)	(0.01)
Scdum	0.015***	0.843**	0.271***	0.014***	0.84**	0.268***
	(0.003)	(0.358)	(0.058)	(0.003)	(0.36)	(0.058)
Obcdum	0.013***	0.99***	0.255***	0.013***	0.989***	0.253***

	(0.003)	(0.312)	(0.058)	(0.003)	(0.313)	(0.058)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	99,506	99,506	99,506	99,506	99,506	99,506

* Prob: probability of households' decision to participate. Cond: per capita expenditure conditional on the participation of the market. Uncond: overall marginal effect, $E(Y|X)$ which can be decomposed into conditional expectation, $E(Y|X, Y > 0)$ and the probability of a positive value of Y, conditional all other regressors, $P(Y > 0|X)$. Standard errors in parentheses clustered at district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7 Estimated marginal effects for traditional fuel: firewood

	Model 1			Model 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Prob.	Cond.	Uncond.	Prob.	Cond.	Uncond.
HHsize	0.052***	-7.000***	-1.41***			
	(0.002)	(0.22)	(0.191)			
Adult				0.049***	-6.694***	-1.419***
				(0.002)	(0.253)	(0.207)
Child				0.056***	-7.233***	-1.378***
				(0.002)	(0.244)	(0.228)
Ln_totalexp	-0.168***	2.651***	-8.194***	-0.166***	2.534***	-8.201***
	(0.009)	(0.759)	(0.734)	(0.009)	(0.763)	(0.731)
Age	-0.000**	0.022	-0.013	-0.000	0.005	-0.012
	(0.000)	(0.017)	(0.017)	(0.000)	(0.018)	(0.017)
Male	0.064***	-0.711	3.312***	0.065***	-0.872	3.298***
	(0.008)	(0.477)	(0.615)	(0.008)	(0.579)	(0.624)
Edu_primary	-0.065***	-2.327***	-5.347***	-0.065***	-2.37***	-5.337***
	(0.009)	(0.544)	(0.677)	(0.09)	(0.542)	(0.675)
Edu_high	-0.202***	-7.599***	-16.841***	-0.202***	-7.64***	-16.823***
	(0.01)	(0.604)	(0.77)	(0.01)	(0.602)	(0.768)
Edu_university	-0.355***	-15.573***	-30.966***	-0.354***	-15.611***	-30.949***
	(0.013)	(1.313)	(1.23)	(0.013)	(1.312)	(1.233)
Urban	-0.485***	-4.401***	-31.46***	-0.485***	-4.398***	-31.449***
	(0.014)	(0.903)	(1.057)	(0.014)	(0.903)	(1.056)
Ration	0.145***	-3.861***	6.077***	0.145***	-3.900***	-6.064***
	(0.009)	(0.773)	(0.711)	(0.009)	(0.775)	(0.711)
Other religion	-0.005	-0.109	-0.24	-0.006	0.167	-0.245
	(0.011)	(0.88)	(0.914)	(0.011)	(0.88)	(0.916)
Stdum	0.165***	3.59***	12.044***	0.165***	3.59**	12.034***
	(0.018)	(1.428)	(1.54)	(0.018)	(1.43)	(1.539)
Scdum	0.135***	2.485***	9.592***	0.135***	2.519***	9.592***
	(0.01)	(0.761)	(0.824)	(0.01)	(0.761)	(0.824)
Obcdum	0.083***	1.274*	5.693***	0.082***	1.296*	5.686***

	(0.008)	(0.745)	(0.700)	(0.008)	(0.746)	(0.698)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	99,535	99,535	99,535	99,535	99,535	99,535

* Prob: probability of households' decision to participate. Cond: per capita expenditure conditional on the participation of the market. Uncond: overall marginal effect, $E(Y|X)$ which can be decomposed into conditional expectation, $E(Y|X, Y > 0)$ and the probability of a positive value of Y, conditional all other regressors, $P(Y > 0|X)$. Standard errors in parentheses clustered at district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The head of households often make the decisions of expenditures, and thus the characteristics of household head could also affect the adoption of fuels and corresponding spending. In particular, having a male as a household head increases the probability of using dung cake by 1.3% and decreases per capita dung cake expenditure by 0.6 Rupee conditional on the choice of dung cake. The probability of using firewood increases by 6.4% for households with a male head, while there are no significant differences in per capita firewood expenditure conditional on the choice of firewood between households with a male or female head. In addition, the unconditional expenditure on dung cake and firewood is 0.2 and 3.3 Rupees respectively higher for households with a male head. These findings are entirely in line with the expectations because female members are often in charge of cooking in Indian households and thus female headed households may be more concerned about the profoundly negative health impacts of indoor air pollution caused by the burning of traditional solid biomass fuels, including dung cake and firewood. In addition, the results indicate that education levels of household heads significantly decrease the probability of using traditional fuels and also households' expenditures on traditional fuels. In particular, the higher level of education obtained by household head, the greater declines in either probability of using or the amount of expenditure on traditional fuels. The reason behind this finding could be that well-educated household head may have more knowledge on the adverse effects of using traditional fuels.

3.6.1.2. Transition fuels

Table 3.8 and 3. 9 display the results for PDS and non-PDS kerosene for the two model specifications, respectively. The correlation between household size and the probability of participating in the PDS kerosene market, as well as PDS kerosene expenditures, is similar with that for non-PDS kerosene. More specifically, with an additional member entering into the household, the probability of consuming PDS (or Non-PDS) kerosene increases by 4.1% (1.2%), while per capita expenditure on PDS (or Non-PDS) kerosene decreases by 1.7 and 2.8 Rupees respectively, conditional on the adoption. The unconditional per capita expenditure on either PDS or non-PDS kerosene is negatively correlated with household size.

It is found that household total expenditure increases the probability of using either PDS kerosene or non-PDS kerosene. As for the conditional expenditure, the effect of total expenditure on PDS kerosene is significantly negative (-0.5 Rupee), while that on non-PDS kerosene is significantly positive (1 Rupee). Having a male as household head increases the probability of consuming kerosene and decreases the conditional expenditures on kerosene. Similar with traditional fuels (i.e., dung cake and firewood), education levels of household head also have significantly negative effects on the probability of consuming kerosene and the expenditure on kerosene, no matter households have decided to adopt kerosene as fuels.

Possessing a ration card increases the probability of consuming PDS kerosene by 78.7%. In contrast, having ration cards decreases the probability of households consuming non-

PDS kerosene and also decreases non-PDS kerosene expenditure conditional on the consumption of non-PDS kerosene. This finding is reasonable because under the PDS, food and non-food items (including kerosene) are distributed to poor households at subsidized rates, and only the households with ration cards are eligible to purchase those subsidized goods from the public distribution system. Urban households are less likely to consume either PDS kerosene or non-PDS kerosene, while their expenditures on PDS kerosene or non-PDS kerosene are greater than rural households once they have decided to consume PDS kerosene or non-PDS kerosene. For the effects of households' social group, compared to households in upper caste, households belonging to scheduled tribe, schedule caste or other backwards classes are more likely to consume either PDS or non-PDS kerosene, while once having decided to consume kerosene, there are no significant differences in kerosene expenditures between scheduled tribe/caste and upper caste households.

Table 3.8 Estimated marginal effects for transition fuel: PDS kerosene

	Model 1			Model 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Prob.	Cond.	Uncond.	Prob.	Cond.	Uncond.
HHsize	0.041***	-1.741***	-0.494***			
	(0.002)	(0.041)	(0.034)			
HH_adult				0.038***	-1.623***	-0.468***
				(0.003)	(0.049)	(0.043)
HH_kid				0.044***	-1.831***	-0.509***
				(0.002)	(0.039)	(0.036)
Ln_totalexp	-0151***	-0.514***	-1.843***	-0.15***	-0.561***	-1.857***
	(0.009)	(0.092)	(0.109)	(0.009)	(0.091)	(0.109)
Age	-0.001***	0.034***	0.005*	-0.001***	0.027***	0.004
	(0.000)	(0.002)	(0.003)	(0.000)	(0.003)	(0.003)
Male	0.037***	-0.746***	-0.015	0.038***	-0.804***	-0.03
	(0.008)	(0.091)	(0.099)	(0.008)	(0.091)	(0.1)
Edu_primary	-0.04***	-0.063	-0.453***	-0.04***	-0.082	-0.456***
	(0.008)	(0.069)	(0.097)	(0.008)	(0.069)	(0.097)
Edu_high	-0.133***	-0.21**	-1.493***	-0.133***	-0.228**	-1.496***
	(0.008)	(0.083)	(0.100)	(0.008)	(0.082)	(0.1)
Edu_university	-0.253***	-0.483***	-2.881***	-0.253***	-0.493***	-2.881***
	(0.015)	(0.134)	(0.172)	(0.015)	(0.134)	(0.173)
Urban	-0.271***	0.667***	-2.455***	-0.27***	0.66***	-2.458***
	(0.011)	(0.138)	(0.122)	(0.011)	(0.138)	(0.123)
Ration	0.787***	0.228	8.288***	0.788***	0.21	8.285***
	(0.021)	(0.171)	(0.244)	(0.021)	(0.172)	(0.244)
Other_religion	0.015	-0.469***	0.404**	0.014	0.492***	0.41**
	(0.014)	(0.124)	(0.162)	(0.014)	(0.124)	(0.162)
Stdum	0.089***	0.23	1.044***	0.089***	0.231	1.043***
	(0.02)	(0.187)	(0.243)	(0.02)	(0.187)	(0.244)
Scdum	0.117***	0.129	1.286***	0.117***	0.141	1.289***
	(0.012)	(0.101)	(0.131)	(0.012)	(0.101)	(0.131)
Obcdum	0.082***	-0.141	0.775***	0.082***	-0.133	0.776***

	(0.01)	(0.093)	(0.117)	(0.01)	(0.093)	(0.116)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	99,490	99,490	99,490	99,490	99,490	99,490

* Prob: probability of households' decision to participate. Cond: per capita expenditure conditional on the participation of the market. Uncond: overall marginal effect, $E(Y|X)$ which can be decomposed into conditional expectation, $E(Y|X, Y > 0)$ and the probability of a positive value of Y, conditional all other regressors, $P(Y > 0|X)$. Standard errors in parentheses clustered at district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9 Estimated marginal effects for transition fuel: non-PDS kerosene

	Model 1			Model 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Prob.	Cond.	Uncond.	Prob.	Cond.	Uncond.
HHsize	0.012***	-2.841***	-0.233***			
	(0.001)	(0.089)	(0.024)			
HH_adult				0.009***	-2.81***	-0.277***
				(0.001)	(0.118)	(0.031)
HH_kid				0.015***	-2.863***	-0.197***
				(0.001)	(0.104)	(0.029)
Ln_totalexp	-0.012**	1.084***	-0.021	-0.011**	1.075***	-0.007
	(0.005)	(0.267)	(0.087)	(0.005)	(0.27)	(0.088)
Age	-0.000***	-0.01	-0.008***	0.000*	-0.012	-0.005**
	(0.000)	(0.008)	(0.002)	(0.000)	(0.008)	(0.003)
Male	0.025***	-0.73**	0.283***	0.026***	-0.747**	0.301***
	(0.004)	(0.321)	(0.089)	(0.004)	(0.326)	(0.09)
Edu_primary	-0.005	-0.414	-0.141*	-0.004	-0.418	-0.132
	(0.004)	(0.293)	(0.083)	(0.004)	(0.292)	(0.083)
Edu_high	-0.036***	-1.378***	-0.769***	-0.035***	-1.382***	-0.757***
	(0.004)	(0.308)	(0.09)	(0.004)	(0.307)	(0.09)
Edu_university	-0.085***	-3.485***	-1.86***	-0.084***	-3.489***	-1.85***
	(0.007)	(0.512)	(0.15)	(0.007)	(0.512)	(0.15)
Urban	-0.06***	3.011***	-0.486***	-0.06***	3.012***	-0.483***
	(0.005)	(0.47)	(0.121)	(0.005)	(0.47)	(0.12)
Ration	-0.194***	-2.184***	-3.373***	-0.193***	-2.188***	-3.37***
	(0.009)	(0.331)	(0.255)	(0.009)	(0.332)	(00.255)
Other_religion	0.013**	1.592***	0.449***	0.013**	1.596***	0.44***
	(0.006)	(0.349)	(0.111)	(0.006)	(0.348)	(0.11)
Stdum	0.011	0.318	0.226	0.011	0.316	0.226
	(0.01)	(0.505)	(0.186)	(0.01)	(0.506)	(0.186)
Scdum	0.021***	0.19	0.357***	0.02***	0.191	0.351****

	(0.005)	(0.34)	(0.097)	(0.005)	(0.34)	(0.097)
Obcdum	0.018***	-0.658**	0.176*	0.017***	-0.656**	0.172*
	(0.005)	(0.305)	(0.094)	(0.005)	(0.305)	(0.093)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	99,502	99,502	99,502	99,502	99,502	99,502

* Prob: probability of households' decision to participate. Cond: per capita expenditure conditional on the participation of the market. Uncond: overall marginal effect, $E(Y|X)$ which can be decomposed into conditional expectation, $E(Y|X, Y > 0)$ and the probability of a positive value of Y, conditional all other regressors, $P(Y > 0|X)$. Standard errors in parentheses clustered at district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6.1.3 Modern fuels

The results for the effects of different factors on households' expenditure on electricity and LPG are reported in Table 3.10 and 3.11, respectively. It is found that the effect of household size on household consumption of electricity/LPG is very different from traditional and transition fuels. The main difference is that there exists a significantly negative rather than positive relationship between household size and the probability of participating in electricity/LPG market. In particular, with the presence of one more person in a household, the probability of adopting electricity and LPG decreases by 0.5% and 0.4%, respectively, and monthly per capita expenditure on electricity/LPG decreases much more than that on traditional and transition fuels. Splitting household occupants into adults and children, it is found that adding one more child decreases the probability of participation in LPG market slightly more than adding one more adult into a household.

It can be seen that household total expenditure has significantly positive effects on the probability of participating in consuming electricity/LPG, and households' electricity/LPG expenditures also increases with total expenditure once the decision to participate has been made. In particular, the effects of household total expenditure on per capita expenditure on each type fuels also satisfy the energy ladder hypothesis, which means with the increase in household income, households would move from traditional solid fuels to modern and cleaner energy (Leach, 1992). As household total expenditure (which is used as the proxy household income) increases by 1%, per capita expenditure on the traditional fuels, dung cake and firewood would decrease by 0.14 and 9.19 Rupees,

while that on modern fuels, electricity and LPG would increase by 24.7 and 35.3 Rupees, respectively.

Households residing in urban areas are more likely to consume electricity/LPG, and urban households' expenditures on electricity/LPG conditional on electricity/LPG market participation are significantly higher than rural households. More specifically, living in urban areas increases the probability of consuming electricity by 5.3% and per capita electricity expenditure by 14.8 Rupees conditional on the participation. In the case of LPG, the probability of participating in LPG market for urban households is 7.3% higher and per capita expenditure on LPG is 13.1 Rupees greater than rural households.

The characteristics of household heads also have significant effects on household electricity and LPG consumption. Having a male as household head decreases the probability of participating in the electricity market by 1.5% and decreases per capita electricity expenditure by 5.5 Rupees conditional on the participation of electricity market. In the case of LPG, households with a male head are 11.3% less likely to consume LPG and once having made the decision of consumption, per capita LPG expenditure is 4.2 Rupees less compared to female-headed households. In addition, the education level of household head also plays a significant role in the participation of electricity and LPG market and the corresponding expenditures. The higher education level obtained by household heads, the more likely to participate in the electricity and LPG market, and the larger per capita expenditures conditional on the participation of electricity/LPG market. For example, households with heads obtaining undergraduate degree or above are 7% and

51.1% more likely to participate in electricity and LPG market, respectively, and their per capita expenditures on electricity and LPG conditional on the adoption are 14.7 and 11.9 Rupees higher than those households whose heads have not been even to primary school.

There are no significant differences between households with different religious beliefs in the probability of adopting electricity, while households in other religions are 2% less likely to participate in LPG markets which is significant at 10% level compared with Hindu households. Conditional on adopting electricity or LPG, per capita expenditure on electricity or LPG are 2.5 and 1.3 Rupees respectively higher for households with other religious beliefs. Turning to the effect of social group, compared with households in the upper caste, households belonging to scheduled tribe/castes and other classes are less likely to participate in electricity/LPG market and their expenditures on electricity/LPG conditional on market participation are also relatively lower.

Table 3.10 Estimated marginal effects for modern fuel: electricity

	Model 1			Model 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Prob.	Cond.	Uncond.	Prob.	Cond.	Uncond.
HHsize	-0.005***	-10.109***	-9.856***			
	(0.000)	(0.192)	(0.19)			
HH_adult				-0.001	-9.199***	-8.798***
				(0.001)	(0.196)	(0.194)
HH_kid				-0.008***	-10.947***	-10.791***
				(0.001)	(0.212)	(0.209)
Ln_totalexp	0.071***	22.738***	24.743***	0.07***	22.432***	24.367***
	(0.003)	(0.449)	(0.454)	(0.003)	(0.443)	(0.449)
Age	0.001***	0.365***	0.377***	0.000***	0.315***	0.319***
	(0.000)	(0.011)	(0.011)	(0.000)	(0.012)	(0.011)
Male	-0.015***	-5.493***	-5.891***	-0.017***	-5.873***	-6.344***
	(0.003)	(0.407)	(0.399)	(0.003)	(0.408)	(0.4)
Edu_primary	0.029***	3.34***	4.415***	0.028***	3.202***	4.246***
	(0.002)	(0.402)	(0.402)	(0.002)	(0.4)	(0.399)
Edu_high	0.052***	8.16***	10.002***	0.051***	8.02***	9.818***
	(0.003)	(0.435)	(0.454)	(0.003)	(0.431)	(0.449)
Edu_university	0.07***	14.685***	17.028***	0.069***	14.577***	16.872***
	(0.004)	(0.645)	(0.673)	(0.004)	(0.644)	(0.671)
Urban	0.053***	14.762***	16.343***	0.052***	14.678***	16.238***
	(0.004)	(0.577)	(0.582)	(0.004)	(0.571)	(0.576)
Ration	0.009***	1.369***	1.68***	0.008***	1.305**	1.605***
	(0.003)	(0.536)	(0.531)	(0.003)	(0.533)	(0.528)
Other_religion	-0.004	2.525**	2.23***	-0.003	2.642**	2.381***
	(0.004)	(0.515)	(0.529)	(0.004)	(0.509)	(0.523)
Stdum	-0.033***	-5.245	-6.404***	-0.033***	-5.188***	-6.348***
	(0.006)	(0.745)	(0.795)	(0.006)	(0.739)	(0.789)
Scdum	-0.021***	-4.117	-4.825***	-0.02***	-4.044***	-4.724***
	(0.003)	(0.422)	(0.441)	(0.003)	(0.42)	(0.44)
Obcdum	-0.009***	-2.745**	-3.005***	-0.009***	-2.669***	-2.911***

	(0.003)	(0.372)	(0.378)	(0.003)	(0.37)	(0.376)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	99,636	99,636	99,636	99,636	99,636	99,636

* Prob: probability of households' decision to participate. Cond: per capita expenditure conditional on the participation of the market. Uncond: overall marginal effect, $E(Y|X)$ which can be decomposed into conditional expectation, $E(Y|X, Y > 0)$ and the probability of a positive value of Y, conditional all other regressors, $P(Y > 0|X)$. Standard errors in parentheses clustered at district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table3.11 Estimated marginal effects for modern fuel: LPG

	Model 1			Model 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Prob.	Cond.	Uncond.	Prob.	Cond.	Uncond.
HHsize	-0.042***	-12.163***	-8.728***			
	(0.003)	(0.204)	(0.273)			
HH_adult				-0.026***	-11.538***	-7.39***
				(0.003)	(0.226)	(0.286)
HH_kid				-0.055***	-12.809***	-9.933***
				(0.003)	(0.215)	(0.291)
Ln_totalexp	0.438***	11.616***	35.334***	0.433***	11.479***	34.862***
	(0.013)	(0.573)	(0.972)	(0.013)	(0.578)	(0.972)
Age	0.006***	0.293***	0.545***	0.005***	0.253***	0.463***
	(0.000)	(0.016)	(0.02)	(0.000)	(0.018)	(0.021)
Male	-0.113***	-4.197***	-9.697***	-0.119***	-4.497***	-10.256***
	(0.009)	(0.503)	(0.633)	(0.009)	(0.51)	(0.637)
Edu_primary	0.145***	3.692***	11.637***	0.143***	3.616***	11.456***
	(0.007)	(0.632)	(0.57)	(0.007)	(0.632)	(0.571)
Edu_high	0.346***	8.941***	27.755***	0.343***	8.887***	27.548***
	(0.008)	(0.622)	(0.631)	(0.008)	(0.624)	(0.631)
Edu_university	0.511***	11.883***	40.398***	0.509***	11.85***	40.161***
	(0.016)	(0.835)	(1.198)	(0.016)	(0.835)	(1.196)
Urban	0.073***	13.085***	30.575***	0.357***	13.069***	30.487***
	(0.012)	(0.666)	(0.93)	(0.012)	(0.665)	(0.925)
Ration	0.073***	-2.142***	3.929***	0.073***	-2.195***	3.851***
	(0.012)	(0.658)	(0.911)	(0.012)	(0.651)	(0.907)
Other_religion	-0.02*	1.344**	-0.712	-0.017	1.462**	-0.439
	(0.011)	(0.658)	(0.833)	(0.011)	(0.658)	(0.827)
Stdum	-0.179***	1.518	-11.417***	-0.179***	1.546	-11.336***
	(0.017)	(1.216)	(1.348)	(0.017)	(0.205)	(1.343)
Scdum	-0.161***	-0.728	-11.296***	-0.16***	-0.698	-11.135***
	(0.01)	(0.219)	(0.784)	(0.01)	(0.596)	(0.783)
Obcdum	-0.057***	-1.055**	-4.367***	-0.055***	-1.012*	-4.232***

	(0.009)	(0.52)	(0.68)	(0.009)	(0.523)	(0.679)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	99,529	99,529	99,529	99,529	99,529	99,529

* Prob: probability of households' decision to participate. Cond: per capita expenditure conditional on the participation of the market. Uncond: overall marginal effect, $E(Y|X)$ which can be decomposed into conditional expectation, $E(Y|X, Y > 0)$ and the probability of a positive value of Y, conditional all other regressors, $P(Y > 0|X)$. Standard errors in parentheses clustered at district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6.2 Household size and economies of scales in household energy expenditure

Although the negative relationship between household size and per capita expenditures on fuels and energy can be regarded as the evidence of the presence of household economies of scale in energy consumption, this study further calculates the semi-elasticities of per capita expenditure with respect to household size for each type of energy, i.e., the percentage change in per capita expenditure with one extra person added into the household.[23] This is achieved by using the command "margins eydx" in STATA, from which we can obtain the marginal effects of a change in household size on the logarithm of per capita energy expenditure, i.e., the semi-elasticities at sample means. These results are reported in Table 3.12-3.14. For traditional fuels (see Table 3.12), the elasticity of per capita expenditure on dung cake with respect to household size is significantly positive (9.3%), indicating significant diseconomies of scale in household dung cake expenditure. It is interesting that households' per capita expenditure on dung cake increases by 12.4% with an extra child added into the household, which is larger than that with an extra adult added. This might be explained by that households with more children have more labour forces to collect free dung cake. In the case of firewood, the per capita expenditure elasticity to household size is -3.7%, i.e., with one more person entering into the household, per capita expenditure on firewood decreases by 3.7%. This suggests that there exist small economies of scale in household expenditure on firewood. The reason behind either no economies of scale in dung cake expenditure or very small economies of scale in firewood expenditure could be that larger household size means more labour force that can be used for collecting free dung cake and firewood, which could offset the scale effects on household expenditures on dung cake and firewood.

[23] The reason why I did not take logarithm directly in per capita energy expenditure in the double hurdle model is that for some observations in the survey, the per capita energy expenditure is less than 1, so taking logs will result in negative values.

Table 3.12 Estimated semi-elasticities of household size (traditional fuels)

	Dung cake		Firewood	
	(1)	(2)	(1)	(2)
HHsize	0.093***		-0.037***	
	(0.012)		(0.005)	
HH_adult		0.067***		-0.037***
		(0.025)		(0.005)
HH_kid		0.124***		-0.036***
		(0.024)		(0.006)
N	99,506	99,506	99,535	99,535

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Moving on to transition fuels, the results show that the economies of scale in households' kerosene consumption lies between traditional and modern fuels, and there are no significant differences between the per capita expenditure elasticities of PDS and non-PDS kerosene with respect to household size (see Table 3.13). More specifically, with an extra person added to a household, per capita expenditure on PDS kerosene decreases by 9%, while that on non-PDS kerosene decreases by 9.9%. The elasticity to the number of adults is close to the number of children within a household for both PDS and non-PDS kerosene. Meanwhile, they are also similar to the model which does not distinguish between adults and children.

Table 3.13 Estimated semi-elasticities of household size (transition fuels)

	Kerosene (PDS)		Kerosene (Non-PDS)	
	(1)	(2)	(1)	(2)
HHsize	-0.09***		-0.099***	
	(0.006)		(0.012)	
HH_adult		-0.085***		-0.117***
		(0.008)		(0.015)
HH_kid		-0.093***		-0.083***
		(0.006)		(0.013)
N	99,490	99,490	99,502	99,502

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Finally, in the case of modern fuels, the expenditure elasticities to household size are largest among these three groups of fuels, indicating the presence of enormous economies of scale in household expenditures on modern and cleaner fuels (see Table 3.14). For electricity, the per capita expenditure elasticity decreases by 24% with an additional person added into the household. Per capita expenditure on electricity declines by 22.5% with an extra adult added, which is slightly lower than that with an extra child added into the household (i.e., 26.3%). The economies of scale

in LPG consumption are similar to electricity. Per capita LPG expenditure decreases by 26.5% as one more occupant enters into a household. It drops by 22.5% with an additional adult while more declines occur (30.3%) with one more child added to a household.

Table 3.14 Estimated semi-elasticities of household size (modern fuels)

	Electricity		LPG	
	(1)	(2)	(1)	(2)
HHsize	-0.24***		-0.265***	
	(0.006)		(0.008)	
HH_adult		-0.214***		-0.225***
		(0.006)		(0.008)
HH_kid		-0.263***		-0.303***
		(0.007)		(0.008)
N	99,636	99,636	99,529	99,529

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.7 Discussion

In this section, I first make a comparison between the main findings in this study and those in previous studies and then project the loss of scale economies in household energy consumption if current trends in decreases of household size persists in India. As mentioned in the literature review section, there are several studies focusing on investigating the economies of scales in household energy consumption in different countries, but they use different ways to measure the same phrase "economies of scale". In this study, I also employed two different measures of economies of scales based on Kakwani-Son's method, as well as the elasticity of per capita energy consumption with respect to household size from the estimation of the double hurdle model. Besides, the definition of energy goods might be subtly different across studies. Given these considerations, one needs to be careful when comparing empirical results across these studies.

3.7.1 Comparisons with other studies

First, the results from Kakwani-Son's model in this study are compared with other studies using the same method. In their original article, Kakwani and Son (2005) estimate the economies of scale in household consumption of individual commodities in Australia with the assumption that

medical goods are purely private. Their results show that the economies of scale index for energy and power goods is 0.55, while this study finds that the index for household expenditure on energy goods in India is 0.71 based on the same assumption. It may not be suitable to compare the two figures directly, considering that they are for different countries and the commodity categories as well as the definition of energy goods in particular are not consistent. Nevertheless, both studies provide empirical evidence on the existence of economies of scale in household energy consumption in Australia and India. Similarly, Rahman et al. (2020) find that the economies of scale index for household utility expenditure is 0.55 in Bangladesh, a country that is very similar to India, based on a different assumption that transportation service is assumed to be purely private good.[24] Although different goods are assumed to be purely private, their results as well as the findings in this study both suggest that energy/utilities are closest to public goods. Unfortunately, since there is no energy commodity category in their demand system model, we cannot obtain the evidence on the economies of scale index for households' energy consumption in Malaysia and Iran from Mok and Maclean (2011) and Karbasi and Mohammadzadeh (2018), respectively.

We now turn to a comparison of the findings in this study with those in previous studies on empirical applications of double-hurdle model in household energy consumption, as well as those studies on the determinants of household energy consumption in India. Adusah-Poku and Takeuchi (2020) use the double-hurdle model to investigate the determinants of household expenditures on LPG and charcoal in Ghana. Some factors used in their analysis are also considered in this study, such as household income (total expenditure is used as the proxy of income), household size, demographic characteristics of household head such as the gender, age, and education levels. Compared to the results for LPG, the effects of household income, the gender and education levels of household heads on the probability of participating in LPG market

[24] Unfortunately, in their article, Rahman et al. (2020) did not provide any information on the specific components of household utility expenditure. I just assume that energy expenditure is incorporated.

in Ghana are very similar with the findings for Indian households in this study. Gould et al. (2020) as well as Daripa and Dinda (2022) are both very recent studies into the determinants of household fuel consumption in India using the double-hurdle model. Gould et al. (2020) conduct a double-hurdle model analysis to explore the determinants of adoption and consumption quantity of LPG in Rural India. The effects of factors on households' adoption of LPG like household size, social group, total expenditure, as well as the gender, education level of household heads found in this study are also in line with their results. Daripa and Dinda (2022) mainly focus on examining the factors influencing households' consumption of firewood in rural India. The effects of household size, education levels of household head and social groups on the probability of adoption firewood are consistent with the findings in this study.

There are some studies in India mainly focusing solely on investigating the determinants of household choice of cleaner fuels, in particular, electricity and LPG. For example, the effects of households' social group and education level of household heads on the probability of adopting electricity and LPG in this study are consistent with those found by Saxena and Bhattacharya (2018), and Mottaleb and Rahut (2021). That is, compared with households in upper classes, those in other classes are less likely to adopt electricity and LPG, and well-educated household heads are more likely to decide to participate in modern energy markets.

Finally, a comparison is made between the estimated elasticity of per capita energy consumption with respect to household size in this study and that in two other studies (Brounen et al., 2012; Wu et al., 2021). As shown in Table 3.14, the results in this study suggest that with one more person added into the household, per capita electricity expenditure on average decreases by 24% in India. Brounen et al. (2012) find that with an extra person entering into the household, per capita electricity consumption decreases by 18% in Netherlands, while the findings by Wu et al. (2021) show that it decreases by 18.4% in China. Note that the two previous study measure

household electricity consumption using consumption quantities (in kWh) rather than expenditures which are used in this study.

3.7.2 Implications of declines in household size in India

According to the information from Census of India, the average household size has been decreasing since 1991 from its peak of 5.5, to 5.3 in 2001 and to 4.9 in 2011. The literature (e.g., Adams, 2010) argues that the decline in family size is a consequence of increasing urbanization, industrialization, and educational expansion, which are all part of higher economic development. India started to liberalize its economy for higher economic development in 1991, and it experienced rapid growth in its economy between 2001-2011. The average household size in India will likely drop by 0.5 around 2030 if the trend in 1991-2011 persists, and according to the projection by the United Nations, the total population in India will rise to 1.51 billion in 2030. Given this information, we can project how changes in household size would affect households' energy expenditure in India. Table 3.15 reports the projected changes in household expenditure on different energy under a 0.5 decline in average household size in India by 2030 compared to 2011, holding all other variables unchanged (which although is a rather unrealistic assumption). It can be seen that except for dung cake, the expenditure on other energy and fuels will increase as the average household size decreases by 0.5 in 2030 due to the loss of economies of scale in household energy expenditure. In particular, the largest increase appears in modern energy, electricity and LPG. On average, monthly per capita expenditure on electricity and LPG will increase by 12% (4.9 Rupees) and 13.3% (4.4 Rupees) respectively relative to the expenditure in 2011. As a comparison, the average monthly per capita expenditure on electricity and LPG are 52.4 Rupees and 37.9 Rupees respectively in 2011-2012 CES survey. By multiplying the projected population in India by 2030, this results in an increase of monthly expenditure on electricity and LPG by 7.4 billion and 6.59 billion Rupees respectively for the whole country compared to 2011. The increase in per capita energy consumption due to the decline in household size could also

cause an increase in pollutants and/or GHG emissions. However, the unavailability of information on the quantities of energy consumed by Indian households makes it difficult to estimate the environmental implications of decreasing in economies of scale in household energy use.

Table 3.15 Implications of a 0.5 decline in average household size on household energy expenditure

Energy	0.5 decline in average household size	
	(1) % change in per capita expenditure	(2) Changes in total expenditure (In Million Indian Rupee)
Dung cake	-4.7%	-60.4
Firewood	1.9%	86.1
PDS kerosene	4.5%	373
Non-PDS kerosene	5%	175.9
Electricity	12%	7441
LPG	13.3%	6590

* Column (1) is the percentage change in household per capita expenditure on each source of energy, which is computed by using the results in Table 3.6 – 3.11; column (2) is the changes in total expenditure on each energy, which is computed by multiplying the marginal effects in Table 3.12-3.14 with the projected total population in 2030.

3.8 Conclusion

Using a nationally representative household survey for India, this study investigates the economies of scale in household energy consumption by using two different but not unrelated concepts. The results suggest that there exist significant economies of scale in household energy use. In particular, the economies of scale for modern and cleaner energy (i.e., electricity and LPG) are greater than these for traditional and transition energy.

The findings in this study have several policy implications. First, there is no doubt that the Indian government should continue to provide households with stable access to modern and cleaner energy (i.e., LPG and electricity). This is not just because modern energy has a less adverse impact on the environment compared to traditional fuels, but also because there exist much higher economies of scale in household consumption of modern energy than that of traditional energy. Second, to mitigate the loss of economies of scale in household energy use associated with the

decreasing average household size, the government should also consider formulating policies that promote shared living spaces, such as tax deductions for households with a greater number of members. Alternatively, it could also investigate the underlying factors contributing to the decline in household size and develop strategies to either reverse or slow down this trend.

This study also has some limitations. First, the relevant literature suggests that energy price and dwelling characteristics also have significant effects on household energy consumption. Due to the unavailability of this information, this study does not identify the effects of those variables in this analysis. Second, Kakwani-Son's method employed in this study has its own limitations. More specifically, this method presumes a simple functional form for the Engel curves and there is also a somewhat arbitrary decision about which commodity is assumed to be a pure private good. This is shown to materially impact the results albeit not enough to change the conclusions that energy is a quasi-public good.

As for possible future work, It is planned to continue employing Kakwani-Son's method to investigate the economies of scale, using different demand systems (such as the popular Almost Ideal demand system), as well as different methods to incorporate household characteristics into the demand system. In addition, finding a proper way to make assumptions about the nature of individual commodities will be a necessary step. Furthermore, I also plan to estimate the environmental impacts of the reduction in the economies of scale in household energy consumption in India in future studies if household level data on the quantities of energy consumed by households is available.

Chapter 4 The amenity value of climate to India households

Abstract

This study estimates the amenity value of climate to Indian households by incorporating climate variables into a popular system of demand, namely the almost ideal demand system model. The sufficient variation in climate along with the availability of data makes India a particularly suitable country for this analysis. The results suggest that climate plays an important role in determining households' costs of living in India. It is found that higher precipitation is an amenity while heating and cooling degree days, as well as relative humidity are disamenities to Indian households. More specifically, holding others constant, households' cost of living would decrease by about 2.1%-5.7% with one standard deviation increase in annual precipitation; for every standard deviation increase in annual heating and cooling degree days, Indian households would experience an increase by 3.1%-5.3% and 2.4%-8% in living costs, respectively; an increase of 1% in annual humidity will raise households' cost of living by 0.7%-2.4%. Based on these findings, climate equivalence scales are computed for households located in major cities in India.

4.1 Introduction

Climate change is currently one of the greatest threats to the whole world. India is considered highly vulnerable to climate change. According to one recent ranking of 67 countries based on their vulnerability to climate change, India is found to be the most vulnerable country to climate change (Paun et al., 2018). The reasons for the vulnerability of India to climate change include: first, India as a lower-middle-income country, has a large number of poor people. Existing studies reveal that the effects of temperature shocks are more adverse to countries that are both hot and poor, such as India (e.g., Dell et al., 2012; Burke et al., 2015; Acevedo et al., 2020 and Kahn et al., 2021); second, two-thirds of its population lives in rural regions and is economically dependent on agriculture which is sensitive to climate change (Gupta, 2022; Gupta and Saudamini, 2022).

Climate change could affect households through several pathways. Changes in climate can damage agricultural production and lead to a decrease in the supply of agricultural products resulting in higher food prices. Households therefore have to pay more for groceries and food items. Changes in climate patterns can also cause health problems, resulting in loss of productivity and income (Mani and Markandya, 2022). In addition, like other local amenities (such as schools, parks and recreation areas), climate also has direct value to households which is called the amenity value of climate to households. Geographically, climate affects the desirability of different locations and the quality of life they offer. Few people are willing to live in the freezing tundra or hot deserts (Albouy et al., 2016). In this study, the amenity value of climate is measured by how much are households willing to pay in order to possess a desirable climate, holding prices and income unchanged.

The amenity value of climate to households can be explained by household production function (HPF) theory, which is originally proposed by Becker (1965). According to HPF theory, marketed goods and non-marketed goods (such as climate) are demanded as intermediate goods for households which are used to produce direct service flows under a certain "production technology". It is these service flows that enter households' utility function directly.[25] According to HPF theory, climate patterns in some locations can facilitate the production of some types of service flows and thus these households can produce the same bundles of service flows at lower costs (Maddison and Rehdanz, 2020). This explains why households located in different climate might achieve different levels of utilities.

This study estimates the amenity value of climate to Indian households by modifying, in a theoretically consistent manner, a popular set of demand functions, namely the almost ideal demand system (AIDS) model, originally proposed by Deaton and Muellbauer (1980) in order to

[25] One example of such basic commodities provided by Becker (1965) is sleeping, which depends on the input of a bed, a house and time.

incorporate climate variables into the analysis.[26] The value of climate amenities is revealed by the role played by climate in determining households' costs of living. This approach has hitherto been mainly applied to calculate household equivalence scales in existing literature. These scales provide information on how much income a household with, e.g., an extra child require in order to sustain the same level of welfare (such as Michelini, 2001; Blacklow et al., 2010; Vashchilko, 2014; Dudel et al., 2020). This study however, uses this approach to estimate the amenity value of climate to Indian households, and to calculate climatic equivalence scales i.e., how much extra income do households inhabiting a particular climate require to achieve the same utility level as an otherwise identical household residing in some reference climate.

There are several reasons why it is interesting to focus on India. First, according to the latest estimates of global population from the United Nations, India has overtaken China to be the most populous country in the world with over 1.4 billion people, which means that the aggregate impact of climate change on Indian households can be enormous. Second, as mentioned above, India is one of the most vulnerable countries to climate change in the world. Third, the sufficient variation in climate due to its geographical and topographical characteristics makes India a particularly suitable country for this analysis. This variation allows us to identify the amenity value of a broad range of climates to households. India has an extraordinary variety of climatic regions, ranging from tropical in the south to temperate and alpine in the Himalayan north. The nation's climate is strongly influenced by the Himalaya, which acts as a barrier to the frigid winds from Central Asia keeping the Indian subcontinent warmer than most locations at similar latitudes (Attri and Tyagi, 2010). As such, North India has a continental climate with severe summer conditions and cold winters when temperatures drop below freezing point, while in the coastal regions of the country, the warmth is unvarying throughout the year. (Attri and Tyagi, 2010). Precipitation in India is significantly influenced by south-west monsoon season (June-September), which brings about 75%

[26] Here "a theoretically consistent manner" means that the coefficients are interpretable.

of the annual rainfall. There is a large variation in the amount of rainfall across different regions in India. The annual average rainfall less than 130mm in west Rajasthan, while in some areas within the northeast state of Meghalaya, the annual rainfall can be over 10,000mm (Attri and Tyagi, 2010).

The econometric analysis in this study is based on the combination of household survey data and district-level historical climate data. The household data are drawn from India's Consumer Expenditure Survey (CES) 2011-2012, which contains information on 101,662 households, covering 625 (out of 640) districts encompassing the entire geographical area. The district-level climate data are constructed based on the high-resolution Japanese global atmospheric reanalysis (JRA-55) database, which provides gridded weather data from 1958 up to present. Climate is measured by annual averages of precipitation, heating degree days (HDDs) and cooling degree days (CDDs), together with relative humidity. The results show that climate plays a significant role in determining households' costs of living. More importantly, climate change could have a noticeable impact on Indian households' costs of living. It is found that more precipitation is a significant amenity to Indian households, while HDDs and CDDs, as well as relative humidity are disamenities.[27] Based on these findings, climate equivalence scales are further computed for households living in ten major cities in India.

This study contributes to the existing literature in important ways. To the best of my knowledge, this is the first study to estimate the amenity value of climate to households in India. Other than focusing on a different country, this study also differs from previous studies using the similar estimation approach. This analysis utilizes micro-level household expenditure data in conjunction with higher resolution of climate data which provides more accurate measurement of climate, while those previous studies are based on aggregated expenditure and climate data. Finally, this study

[27] Degree days are measures of how cold or warm a location is over a period of time, which are defined as the number of degrees by which the average daily temperature is higher than (HDD) or lower than (CDD) the base temperature.

also contributes to the literature by providing the identification of the equivalence scales of climate for households located in major Indian cities.

The remainder of this chapter is organised as follows: Section 4.2 provides an overview of the existing literature on estimating the amenity value of climate. Section 4.3 introduces the demand system model used in this analysis and explains how it is modified to incorporate climates variables. Section 4.4 describes the data, variable definitions as well as empirical specifications. Section 4.5 presents the estimation results, as well as sensitivity analysis. Section 4.6 discusses the main findings and section 4.7 concludes.

4.2 Literature Review

This study is closely related to two strands of literature: the empirical applications of the AIDS model in India, as well as estimating the amenity value of climate to households. I mainly rely on Econlit database to search for the relevant literature. For the first strand of literature, the existing studies employing the AIDS model are searched by using the phrase "demand system" or "almost ideal demand" and "India", resulting in 34 results, including 28 journal articles, 3 working papers and 3 theses. Studies related to the applications of the AIDS model or its quadratic extension by Banks et al., (1997), namely Quadratic Almost Ideal Demand System model (QUIADS) are reviewed in this section. Those studies using other demand system models or involving comparisons over different demand system models even including the AIDS model are excluded. The second strand of literature on estimating climate amenity values to households is searched by using the phrase "amenity value" and "climate" or "climate amenities", resulting in 36 results, among which 18 are in academic journals while 12 and 6 are working papers and theses, respectively. Some other relevant studies are obtained from the references of these resulting studies.

In some studies, the word "climate" is used incorrectly to refer to the weather.[28] I distinguish between climate and weather and only focus on estimating the amenity value of climate (i.e., long-term patterns of weather). Thus, those studies related to weather are excluded from the list of relevant literature. In addition, some other studies are also excluded where climate is not the main focus. Previous studies into the value of climate amenity are reviewed by different valuation approaches including the hedonic approach and life satisfaction approach, as well as the equivalence scale approach.

4.2.1 Applications of the AIDS model in India

Multiple studies have been undertaken utilising the AIDS model in the context of India for the purpose of analysing consumption patterns across different commodity groups. Abdulai et al. (1999) use the AIDS model to estimate food expenditure patterns of both urban and rural households located in two states of India, Haryana and Karnataka. Household food expenditure is disaggregated into 6 categories: milk and milk products, cereal and pulses, edible oils, meat, fish and eggs, vegetables and fruits, other food products. Their results suggest that only milk and milk products are income elastic in both rural and urban areas. Meat, fish and eggs are income elastic for urban households. All other food groups are income inelastic. Srivastava et al. (2013), as well as Felix and Kumar (2020) find the similar results that the food items with high-value food items (including milk, eggs, nuts, and fruits) are more elastic to income and price changes. Aftab et al. (2017) contribute to the literature by estimating the AIDS model to analyse food consumption patterns across three Asian countries, Pakistan, Bangladesh and India.

Some other studies employ the QUAIDS to analyse household food consumption patterns in India. Meenakshi and Ray (1999) is the first among these studies. They estimate the QUAIDS model using household expenditure data collected by national sample surveys covering 16 states

[28] The difference between weather and climate is a measure of time. Weather refers to short-term conditions of atmosphere, while climate characterizes the long-term average patterns of weather (usually taken over 30 years according to NASA and NOAA).

in India. Both food commodities and non-food commodities (including clothing, fuel and light, as well as other non-food) are included in their analysis. Their results show that cereals and cereal substitutes are less income elastic relative to all other commodities. Fuel and light, as well as other non-food goods are more responsive to their own prices compared to the remaining commodity categories. Umanath et al. (2015) use the QUAIDS model to analyse food consumption patterns among households located in the state of Karnataka in India. They find that the income elasticity of demand for rice in rural regions is quite high, while it is income inelastic in urban regions. Rice in Karnataka is a mild complement to many other food commodities, indicating that increased rice price would reduce the consumption of these complementary food commodities. Subramanian et al. (2019) present an analysis of Indian households' demand for a particular high-value food commodity, milk by using the QUAIDS model. Khanal et al. (2016) estimate the QUAIDS model to analyse food demand among rural households in India, and they also find significant substitutional and complementary relationships between different food groups.

There are a few studies that employ the AIDS model to analyse demand for non-food commodities. Gundimeda and Kohlin (2008) examine energy consumption patterns by estimating the AIDS model on four energy commodities: fuelwood, kerosene, LPG and electricity. They find that electricity is less responsive to income changes than other forms of energy, which could be due to the lack of availability of electricity that constrains the responsiveness. The income elasticities of demand for each type of energy are highest for poor households. Instead of analysing household demand for consumption commodities in India, Moore et al. (2005), Moore and Green (2008) employ the AIDS framework to study financial portfolio choices by household and banking sectors, respectively.

The majority of empirical applications of the AIDS model in India focus on analysing household demand for food commodities, while in addition to food commodities, there are also some studies which employ the AIDS model as a framework to investigate demand patterns of energy goods

(Gundimeda and Kohlin, 2008), as well as financial assets (Moore et al., 2005; Moore and Green, 2008). Among those studies using the AIDS model to explore household food demand, it is found that high-value food items (such as milk, eggs, meat, etc) are more elastic to income and price changes, and there exist significant substitutional and complementary relationships between different food items (e.g., Abdulai et al., 1999; Srivastava et al.; 2013; Felix and Kumar; 2020). Umanath et al. (2015) further estimate households' nutrients demand based on their estimation results from the AIDS model. Meenakshi and Ray (1999) is the only study which explores household demand for different food items as well as general non-food commodity categories (including clothing and footwear, fuel and light, and other non-foods). Meenakshi and Ray (1999) and Srivastava et al. (2013) examine households' consumption pattern of food goods by using nationally representative household data, while Abdulai et al. (1999), Umanath et al. (2015), Felix and Kumar (2020) focus on particular states. In addition, Aftab et al. (2017) conduct a cross-country analysis of food demand for households in Pakistan, Bangladesh and India.

4.2.2 Hedonic studies into the amenity value of climate

The hedonic method is the most widely used technique to investigate the amenity value of climate in the literature. The hedonic method assumes that the total value of a good can be separated into the individual values of its attributes. The implicit or hedonic prices of these attributes can be revealed from the observed price of the good and the specific quantities of the attributes associated with it. Rosen (1974) is the first to theoretically analyse the hedonic technique, while Roback (1982) makes another major contribution to the theoretical literature on hedonic technique by arguing that compensating wage and housing price differentials could exist simultaneously across different regions.

Hedonic theory suggests that the costs or benefits of climate are capitalised into housing prices and wage rates, and thus it is expected that households might be willing to pay higher housing prices and accept lower wages to live in areas with desirable climate amenities (Meier and Rehdanz,

2017). The intuition underlying the hedonic technique to value the climate amenities is that if individuals can freely select differentiated localities, they will be attracted to regions with better combinations of amenities. Therefore, housing prices should increase, and wage rates should have to decrease in those regions because of this inward migration. Thus, the implicit value that households place on climate can be inferred from housing price and wage rate differentials across different regions. However, it is notable that the impact of climate amenities on wage rates can be unclear if climate amenities are 'productive' (Roback, 1982) and as will see, empirical evidence from hedonic wage rate studies suggests that the impact of climate on wage rates are ambiguous.

Since Rosen (1974) and Roback (1982), the hedonic technique has been widely used in estimating the value of non-market goods. The hedonic studies on estimating the amenity value of climate are mainly conducted in the US. Hoch and Drake (1974) are the first to study the amenity value of climate. They examine the relationship between wage rates and climate variables for different occupations in the US. Climate is specified by average January and July temperature, average annual precipitation, as well as wind speed. It is found that the climate has significant effect on wage rates and in particular, wage rates are lower for higher summer temperatures. Most of these US studies are not consistent with each other in terms of their use of climate variables. Temperature, precipitation, and relative humidity are the most widely used to estimate the amenity value of climate among these studies in the case of the US, and these three variables are also measured in different ways across studies. In addition, sunshine, windspeed are also used in some studies. Therefore, the review of these hedonic studies in the US in this section are organized according to these three widely used climate variables.

The hedonic literature on hedonic studies in the US suggests a strong relationship between temperature and housing and/or labour markets. There have been two specifications of temperature widely used in the literature, the first of which is average winter and summer temperatures, including Hoch and Drake (1974), Cragg and Kahn (1997; 1999), Kahn (2009),

Koirala and Bohara (2014), Sussman et al. (2014), Sinha et al. (2021). These studies generally suggest that residents in US prefer higher winter temperatures and lower summer temperatures. Cragg and Kahn (1997), Kahn (2009), Sussman et al. (2014) and Koirala and Bohara (2014) find that January temperature significantly raises housing rents, while July temperature significantly decreases housing rents, indicating higher temperature in winter is an amenity while higher temperature in summer is a disamenity. In Sussman et al. (2014), temperature is measured by county-level historical average January and July temperatures and their quadratics. They estimate a hedonic housing price model using Census Bureau's micro-level data for individual houses from the Public Use Microdata Sample (PUMS) in 2000, finding that higher temperature in January is an amenity while, although July temperature also significantly raises housing prices, its quadratic term is significantly negative indicating that above certain level of temperature, residents in US prefer milder summers. For hedonic wage rate analysis, some studies suggest that January temperature significantly increases wage rates while the increased July temperature lowers wages (Cragg and Kahn, 1997; 1999; Koirala and Bohara, 2014). This is of course possible depending on whether climate is considered productive and on the balance between capitalisation in the housing market versus in the labour market. Cragg and Kahn (1999) estimate the same hedonic specification using four cross-sections of PUMS in 1960-1990 respectively, finding that the compensating climate-differentials capitalized into rents and wages have changed over time. February temperature significantly lowers wage rates and July temperature significantly raises wage rates in 1960 and 1970 indicating that US workers prefer warmer Februarys and dislike high temperature in July, while February temperature is no longer significant and July temperature turns to be insignificantly negative in the wage rate regression in 1990. Koirala and Bohara (2014) estimate both hedonic housing price and wage rate models by combining the American Housing Survey and city-level climate data. They use a seemingly unrelated regression method for estimation considering the simultaneity problem caused by the potential correlation between housing rent and wage rate. However, from my perspective, I do not think that this is simultaneity

bias so much as noting that the error terms of these two separate regressions are probably correlated and that taking account of this correlation can lead to greater efficiency. Their results also suggest that January temperature significantly increases both housing rents and wage rates, while July temperature significantly decreases both housing rents and wage rates.

The second widely used specification of temperature in hedonic studies for the US is HDDs and CDDs (Roback, 1982; Hoehn et al., 1987; Gyourko and Tracy, 1991 and Gabriel et al., 2003). Hoehn et al. (1987) use housing expenditures and hourly earnings data to estimate the values of a set of location-specific amenities including climate based on the PUMS data in 1980. It is found that full implicit prices of annual HDDs and CDDs measured by the sum of housing expenditures and wages differentials due to a unit change in HDDs and CDDs are negative, indicating that HDDs and CDDs are significant disamenities in the US. Gyourko and Tracy (1991) use both hedonic housing prices and wages to estimate quality-of-life indices for each state. The state-level quality-of-life index is obtained by multiplying the implicit price for each local amenity (including climate) estimated from the hedonic regression with its quantity and then summing across all local amenities. Their results suggest HDDs and CDDs significantly decrease the state-level quality-of-life. Similar results for city-level quality-of-life indices are also found by Gabriel et al. (2003).

Average annual precipitation is most widely used in Hedonic studies for the US (Hoch and Drake, 1974; Hoehn et al., 1987; Gyourko and Tracy, 1991; Gabriel et al., 2003; Koirala and Bohara, 2014; Albouy et al., 2016). However, preferences for annual precipitation are not consistent among these existing studies. Hoehn et al. (1987) find high average annual precipitation to be a significant amenity, while the results from Gyourko and Tracy (1991) and Gabriel et al. (2003) suggest that average annual precipitation is a disamenity. Englin (1996) investigates the relationship between rainfall measured by long-term average annual rainfall and standard deviation of monthly rainfall, and housing prices in a specific, small area in Washington State in US. It is found that households dislike too much rainfall, but holding the volume of annual rainfall constant, they prefer a greater

variation. Koirala and Bohara (2014) find that city-level annual average precipitation significantly reduces monthly housing rents while its impact on monthly wage rates is insignificant. Thus, the implied implicit price of annual average precipitation is negative.

Humidity is often used in hedonic studies for the US to specify climate, together with temperature and precipitation. Hoehn et al. (1987), Gyourko and Tracy (1991), as well as Gabriel et al. (2003) find that the full implicit price of annual average relative humidity is negative. Cragg and Kahn (1997 and 1999) and Sussman et al. (2014) find annual average relative humidity significantly reduces housing rents. All of these studies suggest that high annual average humidity is a significant disamenity to US residents. In addition, sunshine and windspeed are also used in some US hedonic studies, which suggests that sunshine is a significant amenity while windspeed is a disamenity (Hoehn et al., 1987; Gyourko and Tracy, 1991 and Gabriel et al., 2003).

Outside the US, there are also several hedonic studies of the amenity value of climate for EU countries, including the UK (Maddison, 2001; Rehdanz, 2006; Meier and Rehdanz, 2017), Italy (Maddison and Bigano, 2003), Germany (Rehdanz and Maddison, 2009). By using both hedonic housing and hedonic wage regressions, Maddison (2001) employs the hedonic technique to estimate the amenity value of climate in the UK. Climate measured by annual averages of temperature and precipitation is only found to have significant effect on housing prices but not on wage rates. The results suggest that British households prefer higher temperature and less precipitation. Rehdanz (2006) uses January and July averages of temperature and precipitation to measure the climate of the UK, finding that British residents regard higher temperature and less precipitation in January as amenity. Similar results are also found in Germany by Rehdanz and Maddison (2009). Different from previous hedonic studies in which hedonic housing and wage rates regressions are often estimated separately, Rehdanz (2006) combines the two models by regressing the labour income net of housing costs on climate. This method is first used by Maddison and Bigano (2003) who estimate the amenity value of climate to households in Italy.

Their results show that Italian households prefer less precipitation during winters, but warmer temperature in summer would have a negative welfare effect. Instead of using data aggregated at the regional level e.g., Maddison (2001) and Rehdanz (2006), a more recent hedonic study by Meier and Rehdanz (2017) uses micro-level panel data to investigate the amenity value of climate to households in the UK. Climate at district level is specified by annual averages and annual averages squared of HDDs, CDDs, and rain days. It is found that climate amenities are mainly compensated in housing markets, which is consistent with the findings by Maddison (2001), and in particular, there exists an inverted u-shaped relationship between housing prices and HDDs, CDDs as well as rainy days, which suggests that higher HDDs is a significant amenity to households when it is low while HDDs turns to be a significant disamenity when it is already very high. Similarly for CDDs and rainy days.

The hedonic technique is transparent and easy to interpret, and estimated coefficients from hedonic regressions provide clear indications of the contribution of different variables to housing prices and wage rates. These advantages of hedonic technique make it widely used in environmental valuation. The hedonic studies for climate amenities are mainly conducted in the US and EU countries. However, climate is often measured in different ways which make it difficult to make comparison across studies.

There are also some problems when using hedonic technique to estimate the amenity value of climate. The key assumption of the hedonic technique is that people can freely choose where to live within the country and there are low costs of relocation. This assumption may not hold for most countries in the world. It also explains why most hedonic studies are for the US, since it is a particularly appropriate setting in which to use this method as it has a large population that is mobile over areas with diverse climates (Albouy et al., 2016).

4.2.3 Life satisfaction studies into the amenity value of climate

The study of life satisfaction, which is often used interchangeably with happiness, or subjective well-being (SWB) in the literature is initially conducted by the psychologists and sociologists. Easterlin (1974) is the first economist who empirically studies the reported levels of happiness by examining the relationship between happiness and income. Since then, research into SWB has grown rapidly in economics (for overviews, see Dolan et al., 2008; MacKerron, 2012). In particular, economists are quite interested in the determinants of happiness. The growing number of evidence has shown that it is not only income, but many other factors could affect individuals' SWB, such as health, education, employment, and marriage (Dolan et al., 2008).

Life satisfaction approach is built on the recent development of SWB research in economics. The basic idea of this approach is quite simple. The premise of using this valuation technique is that life satisfaction can serve as a valid approximation for individual welfare. In fact, physiological and psychological research has provided the evidence that using life satisfaction scores as the approximation of welfare is valid.[29] Using self-reported life satisfaction (which is often measured by the answers of the respondents to the survey question such as "On a scale from 1 (not at all satisfied) to 5 (very satisfied), how satisfied are you with your life ?") as the approximation of welfare, one can set up a life satisfaction function by incorporating environmental goods, income and other covariates. The estimated coefficient on the environmental goods can provide a direct valuation in terms of subjective well-being and furthermore, together with the coefficient on income, the willingness-to-pay for the environmental goods can be calculated.

There are several existing studies on estimating the effect of climate on life satisfaction. Frijters and Van Pragg (1998) are the first to estimate the value of climate amenities using this approach. They examine the impact of climate on both subjective welfare and SWB in Russia, with the former

[29] For example, Shedler et al. (1993) find that individuals reporting themselves as satisfied with their lives are more often smiling in social interactions, are less likely to commit suicide.

measured by satisfaction with income and the latter measured by satisfaction with one's whole life, based on household survey data in 1993-1994. Their results suggest that both subjective welfare and SWB are negatively related to cold and windy winters. In addition, higher level of humidity together with higher temperature would also significantly reduce SWB. The biggest problem of this study is that it only identifies 35 climate regions in Russia, the country with largest geographic area in the world. These regions must be very large, and thus the climate cannot be accurately measured within them.

Rehdanz and Maddison (2005) employ life satisfaction approach to estimate the amenity value of climate to households based on the data for 185 observations from 67 countries. Climate for each country is represented by the weighted averages of various indices of temperature and precipitation over major cities within the country, including means, extremes, the number of hot, cold, wet and dry months. The results indicate that climate had significant impact on life satisfaction. High-latitude countries would benefit from limited climate change, while low-latitude countries would suffer losses. Maddison and Rehdanz (2011) confirm their earlier research by including what they believe to be a superior representation of climate into their analysis: heating degree months (HDMs) and cooling degree months (CDMs), which are analogous concepts of HDDs and CDDs.[30] The indices of temperature and climate are over two or more major population centres for each country. Their results suggest that for those living where the climate is characterized by a large number of HDMs, higher temperature could increase their subjective levels of happiness, while for those living in where the climate is characterized by larger number of CDMs, their subjective levels of happiness will decrease with higher temperature. The key limitation of both cross-country studies is that climate for each country is simply measured by the averages over several major cities resulting in imprecision in measurement of climate variables.

[30] Maddison and Rehdanz (2008) replaced the daily mean temperature by monthly mean and calculated so called heating degree months and cooling degree months.

Another cross-country analysis of climate and SWB is conducted by Murray et al. (2013), which estimates the impact of climate on households' SWB in European countries. Murray et al. (2013) solve the inaccurate measurement of climate in previous cross-country studies mentioned above by aggregating gridded climate data into NUTS region level, which results in more than 200 variations in climate.[31] Climate variables used in this study include: temperature, precipitation, relative humidity, frost days, rainy days, percentage of possible sunshine and wind speed. It is found that more sunshine would increase SWB, while higher relative humidity, higher standard deviations in temperature and rainy days would decrease SWB.

Beyond the three cross-country analyses mentioned above, there are also two studies that employ the life satisfaction approach to explore the value of climate amenities to households within a single country. Brereton et al. (2008) examine the impact of local climate on households' life satisfaction in Ireland, which is achieved by matching the individual's life satisfaction together with other demographic and social-economic characteristics taken from 2001 Urban Institute Ireland National Survey on Quality of Life to local (electoral division level) amenities including climate.[32] Climate is described by average annual precipitation, January minimum and July maximum temperature, mean annual wind speed and sunshine. They find that wind speed has a negative impact on subjective well-being, while the effect of both minimum temperature in January and maximum temperature in July are significantly positive. The key contribution of this study is that climate is measured at a highly spatially disaggregated level. Using the identical dataset as Brereton et al. (2008), Ferreira and Moro (2010) further estimate the marginal willingness to pay for the changes in climate by using life satisfaction approach. Local climate is measured by January minimum and July maximum temperature, as well as annual average precipitation. Their results indicate that an individual would be willing to pay about €4,200 and €1,600 for a 0.3℃ increase in

[31] NUTS is a classification system for dividing up the EU into regional territories.

[32] An electoral division is a legally defined administrative area in the Republic of Ireland, generally comprising multiple townlands, and formerly a subdivision of urban and rural districts.

minimum temperature in January and maximum temperature in July in Dublin, respectively. Zapata (2022) contributes to existing literature by firstly investigating the impact of climate on SWB in a tropical and developing country, Ecuador. The author conducts a cross-sectional analysis based on household-level survey data and city-level climate data. Climate is measured by monthly averages of temperature, precipitation, and humidity in two months (November and December), respectively when the survey is conducted. The results suggest that the impact of precipitation and humidity on SWB is insignificant, while temperature plays a significant role in determining people's SWB. An inverted U-shaped relationship is found between the temperature and SWB, indicating that higher temperature is an amenity to Ecuadorean households when temperature is low, while it would turn to be a disamenity when temperature is already very high.

Prior studies using life satisfaction approach to estimate the amenity value of climate to households are not too many. Climate is found to have significant impact on life satisfaction. Climate variables commonly used in these studies include temperature, precipitation, and relative humidity, although the measurement of these variables vary across studies. The cross-country analysis conducted by Rehdanz and Maddison (2005), and Maddison and Rehdanz (2011) suggest that people would prefer higher mean temperatures in the coldest month and lower temperatures in the hottest month. Similar results are also found in Russia (Frijters and Van Pragg, 1998) and Ireland (Brereton et al., 2008; Ferreira and Moro, 2010). The impact of precipitation on life satisfaction is found to be insignificant by Brereton et al., (2008), Ferreira and Moro, (2010), as well as Zapata (2022). Frijters and Van Pragg (1998) find that higher level of humidity together with higher temperature would also significantly reduce SWB in Russia, and Murray et al. (2013) also provide evidence that higher relative humidity would decrease people's life satisfaction in Europe. The impact of humidity is found to be insignificant on life satisfaction in Ecuador by Zapata (2022).

4.2.4 Equivalence scales studies into the amenity value of climate

As introduced in Chapter two, household equivalence scales can be identified from the estimation of a demand system model with household demographic characteristics included. Some studies applied this same equivalence scales approach to estimate the amenity value of climate to households and identify climate equivalence scales, which is achieved by incorporating climate variables into the demand system using the similar way as household demographic variables based on the HPF theory mentioned in Introduction section.

Maddison (2003) is the first to apply equivalence scales approach to estimate the amenity value of climate to households. The quadratic almost ideal demand system (QUAIDS) model developed by Banks et al. (1997) is used to analyse the impact of climate on consumption expenditures based on country-level aggregated data. The climate variables are incorporated into the demand system by "demographic translating" and "demographic scaling" (See Pollak and Wales, 1981), which are originally used to incorporate demographic variables into the demand system. The results indicate that those countries located in high latitudes would benefit from climate change, while low latitude countries would suffer great losses. The biggest contribution of this study is that it introduces an innovative approach, traditionally used to estimate household equivalence scales to estimate the amenity value of climate. The main problem of this study, however, is that the analysis is based on country-level aggregated data. It is obviously implausible to assume that people living in different countries still share the same tastes and preferences. Besides, the climate for each country is simply measured by the climate of its capital city, which is wildly inaccurate especially for those large countries with significant climate variations.

The subsequent study by Maddison et al. (2013) analyse the role of climate in determining the consumption patterns of households in Japan using household expenditure data obtained from 51 cities from 2000 to 2009. They employ both the Linear Expenditure System (LES) and AIDS model in their analysis, the results from which are similar to each other. Climate, which is specified

by the annual mean temperature and precipitation, is incorporated into the demand system using the demographic scaling technique. It is found that higher temperatures are associated with a small but significant reduction in households' cost of living. Precipitation, however, has no significant effect on households' cost of living.

The most recent study is conducted by Affuso (2019), which investigates the welfare effect of climate change in a particular area, Greenland. Following Maddison (2003), the author employs the QUAIDS model and demographic scaling method to incorporate climate into the analysis. Climate change is specified by the anomalies of annual mass change in the ice cap, annual temperature, and annual heating degree days. The analysis is based on aggregated time series price and expenditure data from 2003-2017. Due to the number of parameters being much greater than observations, a very different estimation technique is employed in this study, called generalized maximum entropy. The results suggest that the annual mass change anomaly of the ice cap has a negative impact on consumer's welfare. The counterfactual analysis predicted that the society had suffered a welfare loss of $21.63 million over the period 2003 and 2017 due to climate change, which is in contrast to the previous findings of Maddison (2003).

Previous research utilizing the equivalence scales approach to explore the value of climate amenities to households are scarce. The cross-country analysis conducted by Maddison (2003) suggests that high-latitude countries would benefit from the rise in global temperature, while low-latitude countries would suffer great losses. These results are confirmed by later study conducted by Rehdanz and Maddison (2005) using the life satisfaction approach. In particular, the results from Maddison (2003) indicate that the cost of living index for Japan would slightly decrease by 0.4% with a 2.5℃ rise in global mean temperature, which is consistent with the results found by Maddison et al. (2013). Different from other studies that incorporate climate variables into their analyses, Affuso (2019) includes the climate change variables into the demand system, such as the

anomalies of annual temperature and heating degree days. These variables, in fact refer to weather, rather than climate variables.

4.2.5 Literature discussion

By reviewing the existing literature, it suggests that hedonic approach has been most widely used to estimate the amenity value of climate. Meanwhile, more recent studies are mainly based on the SWB approach. Very few studies on the other hand are relied on the equivalence scales approach. Both hedonic and SWB techniques are transparent and easy to be undertaken for estimating the value of environmental amenities. This may explain the wide use of these two methods in literature. However, the two approaches are not without limitations. For hedonic approach, it requires high quality of property transaction data or wage rate data. These data, in particular property transaction data, may not be available in some places or can only be available at high costs. Missing variables of some attributes (or due to data not being available) may lead to misspecification of the hedonic model and thus biased estimation (Chau and Chin, 2003). Besides, the free movement of individuals without cost and over long distances, as required by the hedonic approach, is also difficult to sustain in many countries, including India. The limitations of SWB approach include: first, it assumes that self-reported life satisfaction is regarded as a good proxy for an individual's welfare, which is not universally accepted by researchers; second, there may exist reverse causality between income and SWB. Besides, property prices and wage rates might already have adjusted, and thus the amenity value of climate may not be reflected by the difference in subjective wellbeing between households resided in different regions may not be able to reflect the amenity value of climate.

To the best of my knowledge, this study is the first to investigate the amenity value of climate to Indian households. In this study, I employ equivalence scales approach for this purpose. Different from the very few existing studies using this approach, this study differs from these studies in at least two aspects: first, this analysis is based on micro-level household expenditure

data; second, higher resolution of climate data is used in this study, which provides more accurate measurement of climate. It is undeniable that like the other two approaches, equivalence scales approach also has its own limitations. More specifically, it assumes that households' preferences for climate are not additively separable. Besides, as discussed by (Bradford and Hildebrandt, 1977), evaluating non-marketed amenities (such as climate) based on households' expenditure of marketed commodities also requires additional assumption that there must exist a price vector of marketed commodities at which marginal utility of the non-marketed amenity is zero.

4.3 Methodology

The approach used in this study involves modifying, in a theoretically consistent manner, a popular demand system model, namely the AIDS model in order to reveal the role played by climate in determining household commodity demands. This approach has been mainly applied to estimate household equivalence scales in existing literature (such as Michelini, 2001; Balli and Tiezzi, 2010; Blacklow et al., 2010; Dudel et al., 2020). This study, however, utilises this approach to estimate climate equivalence scales, i.e., how much extra money would a household inhabiting a particular climate need in order to achieve the same level of utility as an identical household residing in some reference climate. This section presents the derivation of the original AIDS model, followed by the technique used to incorporate climate variables into the demand system: Barten's scaling method (Barten, 1964). As will be seen later, a more flexible way of incorporating climate variables is also employed in this analysis, which is Ray's scaling method (Ray, 1983). Since the details of Ray's method has been introduced in chapter 2, this section will not repeat again. Besides, the AIDS model is also extended to allow household preferences to vary with other conditioning variables, namely demand shifters which are also introduced. Finally, this section presents how to calculate price and expenditure elasticities based on the modified AIDS model. These elasticities are themselves also of interest.

4.3.1 The AIDS Model

The AIDS model gives an arbitrary first-order approximation to any demand system which satisfies all standard axioms of consumer's theory and aggregates perfectly over consumers (Deaton and Muellbauer, 1980). This model has been widely used since it has been proposed. Following Deaton and Muellbauer (1980), the AIDS model can be derived from the cost function which takes the following form:

$$ln\, c(\boldsymbol{p}, u) = (1-u)\, ln\, a(\boldsymbol{p}) + u\, ln\, b(\boldsymbol{p}) \tag{4.1}$$

where u is utility and $\mathbf{p}$ is a price vector.

$\ln a(\boldsymbol{p})$ and $\ln b(\boldsymbol{p})$ are functions of prices, which are expressed as:

$$ln\, a(\boldsymbol{p}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i\, ln\, p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}\, ln\, p_i\, ln\, p_j \tag{4.2}$$

$$ln\, b(\boldsymbol{p}) = ln\, a(\boldsymbol{p}) + \beta_0 \prod_{i=1}^{n} p_i^{\beta_i} \tag{4.3}$$

where α_i, β_i and γ_{ij} are parameters. The subscripts $i, j = 1,2,3, \ldots n$ refer to the i^{th} and j^{th} commodity, respectively, of the n commodities.

The AIDS cost function therefore is written as:

$$ln\, c(u, \boldsymbol{p}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i\, ln\, p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}\, ln\, p_i\, ln\, p_j + u\beta_0 \prod_{i=1}^{n} p_i^{\beta_i} \tag{4.4}$$

According to Shephard's lemma, the Hicksian (or compensated) demand for good i can be obtained by taking the partial derivative of the cost function with respect to p_i. We can obtain the budget share of good i by carrying out the following differentiation:

$$\frac{\partial \ln c(u,\boldsymbol{p})}{\partial \ln p_i} = \frac{p_i q_i}{c(u,\boldsymbol{p})} = w_i \tag{4.5}$$

Therefore, differentiating (4.4) with respect to $\ln p_i$ yields:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij}\, ln\, p_j + u\beta_0 \prod_{i=1}^{n} p_i^{\beta_i} \tag{4.6}$$

For a utility-maximizing consumer, total expenditure y is equal to $c(u, \boldsymbol{p})$ and this equality can be inverted to obtain the indirect utility function, $u(y, \boldsymbol{p})$. By substituting it into (4.6), we finally obtain the AIDS demand functions in budget share form:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln \left[\frac{y}{a(\boldsymbol{p})}\right] \tag{4.7}$$

Furthermore, in order to comply with demand theory and to reduce the number of parameters to be estimated, three constraints are further imposed on the model, which are adding-up restriction, homogeneity and Slutsky symmetry:

(1) Adding up restriction means that the sum of expenditure shares should be one, which implies that:

$$\sum_{i=1}^{n} \alpha_i = 1; \sum_{i=1}^{n} \gamma_{ij} = 0, \forall j; \sum_{i=1}^{n} \beta_i = 0; \text{ and } \sum_{i=1}^{n} \lambda_i = 0 \tag{4.8}$$

(2) Homogeneity means that the demand equation should be homogenous of degree zero in prices and total expenditure, which implies that the sum of the parameters of log-prices for each equation must be zero:

$$\sum_{j=1}^{n} \gamma_{ij} = 0, \forall i \tag{4.9}$$

(3) Slutsky symmetry deals with the substitution effects between goods, which implies that the cross-partial price derivatives are equal:

$$\gamma_{ij} = \gamma_{ji}, \forall i \neq j \tag{4.10}$$

4.3.2 Incorporating climate amenities into the demand system

As suggested by Maddison (2003), climate variables can be incorporated into the AIDS model by applying methods which are identical to those used to incorporate household demographic characteristics into the demand system model based on HPF theory. The climate variables used in this analysis include annual averages of precipitation, HDDs and CDDs, as well as relative humidity (details of climate variables will be introduced in section 4.4.2). In this study, Barten's

price scaling method (Barten, 1964) is initially utilized to introduce climate amenities into the demand system model, which is originally proposed for the treatment of household composition effects in the context of demand systems. Under Barten's method, household demographic effects on consumption demand are similar to price effects. More specifically, household demographic variations are incorporated into household utility functions by dividing the quantity of consumed good i $(i = 1, \dots, n)$ by a function $m_i(\boldsymbol{z})$, where $\boldsymbol{z}$ is a vector of household demographic characteristics:

$$u = u(\frac{q_1}{m_1(\boldsymbol{z})}, \frac{q_2}{m_2(\boldsymbol{z})}, \dots, \frac{q_n}{m_n(\boldsymbol{z})}) \tag{4.11}$$

The corresponding cost function has the following form:

$$c = c(u, p_1^*, p_2^*, \dots, p_n^*) \tag{4.12}$$

where $p_i^* = p_i m_i(\boldsymbol{z})$ is the effective price of good i.

In the baseline model, the price scaling function $m_i(\boldsymbol{z})$ is assumed to be identical across commodities for the simplicity of calculating climate equivalence scales. This, however, is often viewed as a rather unrealistic assumption when one is considering demographic scales, but it is necessary when there is no price variation in the data (which although is not in this study). Moreover, following previous studies (such as Muellbauer, 1977; Michelini, 2001; Molina and Gil, 2005) the price scaling function takes the following form:

$$m(\boldsymbol{z}) = 1 + \boldsymbol{\rho}'\boldsymbol{z} \tag{4.13}$$

Thus, household's cost function (c_h) with inclusion of household demographic characteristics can be written as:

$$c_h(\boldsymbol{p}, \boldsymbol{z}, u) = m(\boldsymbol{z}) \times c_r(\boldsymbol{p}, u) \tag{4.14}$$

where $c_r(\boldsymbol{p}, u)$ is the expenditure function of the reference household. Note that to examine the scaling effect of climate, we also need to define the reference household, for which the scaling function is always equal to one (see below for a discussion of the reference household defined in this study). Correspondingly, the modified AIDS model with household demographic characteristics has the form:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln \left[\frac{y}{m(\boldsymbol{z})a(\boldsymbol{p})}\right] \tag{4.15}$$

The specification of the scaling function in Eq. (4.13) also implies a strong assumption that the scaling function is not affected by prices and reference utility level. Therefore, in a later sensitivity analysis, this assumption is relaxed by employing another method to incorporate household demographic characteristics, namely Ray's scaling method (Ray, 1983) which allows the variation of the scaling function with prices and utility level. (For details of the Ray's method, see section 2.3.2 in chapter two). In particular, note that if $\boldsymbol{\phi}(p, z, \boldsymbol{u}) = 1$ in (2.10), the Ray's scaling function would be identical to Barten's scaling function. In other words, the modified commodity-invariant Barten's scaling function is a special case of Ray's scaling function when its price and utility-varying component is equal to 1. The modified AIDS model using Ray's method has the form:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i^* \ln \left[\frac{y}{\bar{m}(\boldsymbol{z})a(\boldsymbol{p})}\right] \tag{4.16}$$

where $\beta_i^* = \beta_i + \boldsymbol{\eta_i}'\boldsymbol{z}$.

4.3.3 Demand shifters

In addition, households' preference over commodities may also vary across regions and over time, or to other household characteristics except for those scaling variables included in the scaling function (z). This is very likely to happen for the big countries like India. Following Blundell et al. (1993), this variability is modelled by allowing parameters in the AIDS model, α and β to vary

according to those characteristics, which are referred to as demand shifters. More specifically, for the ith commodity, the intercept of the demand equation then is defined as:

$$\alpha_i(\boldsymbol{d}) = \alpha_{i0} + \sum_k \alpha_{ik} d_k\,,\ k = 1,2,\dots,K \tag{4.17}$$

where $\boldsymbol{d}$ is the $K \times 1$ vector of demand shifters. To preserve adding-up, two constraints apply:

$$\sum_{i=1}^{n} \alpha_{i0} = 1 \ and \ \sum_{i=1}^{n} \alpha_{ik} = 0 \tag{4.18}$$

Similarly, the expenditure parameter β for the ith commodity is defined as:

$$\beta_i(\boldsymbol{d}) = \beta_{i0} + \sum_k \beta_{ik} d_k \tag{4.19}$$

with two further constraints:

$$\sum_{i=1}^{n} \beta_{i0} = 0 \ and \ \sum_{i=1}^{n} \beta_{ik} = 0 \tag{4.20}$$

Therefore, the modified AIDS model with the incorporation of both scaling variables and demand shifters takes the following form:

$$w_i = \alpha_i(\boldsymbol{d}) + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i(\boldsymbol{d}) \ln \left[\frac{y}{m(\boldsymbol{z})a(\boldsymbol{p},\boldsymbol{d})}\right] \tag{4.21}$$

4.3.4 Price and expenditure elasticities

Based on the parameters estimated from the baseline model in (4.21), we can further calculate the demand elasticities of each commodity, which can help to understand how household demand responds to the changes in prices and total expenditure. As shown in Banks et al. (1997), the expenditure and price elasticities can be obtained by differentiating (4.21) with respect to $\ln y$ and $\ln p_j$, respectively:

$$\mu_i \equiv \frac{\partial w_i}{\partial \ln y} = \beta_i(\boldsymbol{d}) \tag{4.22}$$

and

$$\mu_{ij} \equiv \frac{\partial w_i}{\partial \ln p_j} = \gamma_{ij} - \beta_i(\boldsymbol{d}) \times \left(\alpha_j + \sum_k \gamma_{kj} \ln p_k\right) \quad (4.23)$$

Then the expenditure elasticity of the ith commodity is computed by:

$$e_i = \frac{\mu_i}{w_i} + 1 \quad (4.24)$$

The compensated (Hicksian) price elasticities are derived from Hicksian demand function, which assumes that the consumer will be given enough wealth (compensation) to maintain the same level of utility before the relative price changes. The compensated price elasticity of commodity *i* with respect to the change in the price of good j can be obtained by:

$$e_{ij}^c = e_{ij}^u + e_i w_j \quad (4.25)$$

The uncompensated (Marshallian) price elasticities, which are derived from Marshallian demand function, assume that real expenditures remain unchanged when relative prices change. The uncompensated price elasticities of commodity i with respect to the change in the price of good j are computed as:

$$e_{ij}^u = \frac{\mu_{ij}}{w_i} - \delta_{ij} \quad (4.26)$$

where δ_{ij} is the Kronecker delta, which equals to 1 if $i = j$ and 0, otherwise (Banks et al., 1997).

4.4 Data and empirical specifications

4.4.1 Household survey data

The household data used in this study is drawn from India's Household CES in 2011-2012, which is a nationally representative household survey conducted by National Sample Survey Office, Ministry of Statistics and Program Implementation, Government of India. The initial wave of the CES survey was conducted in 1986-1987 and the survey in 2011-2012 is the most recent round for which data is publicly available. For each round of CES survey, four sub-rounds are conducted throughout the year in July-September, October-December, January-March and April-June,

respectively. Since CES survey is carried out quarterly over 12 months, it is necessary to check whether it is temporally representative. Figure 4.1 shows the number of households covered in the survey in each month. It can be seen that the total number of sample households are almost evenly distributed over the 12 months, except for May 2012. Table 4.1 further displays the distribution of the number of sample households in each month over districts, from which we can also see that the total number of sample households are almost evenly distributed across districts for each month, while consistent to Figure 4.1, the average number of households in May 2012 is obviously more than that in the rest 11 months.

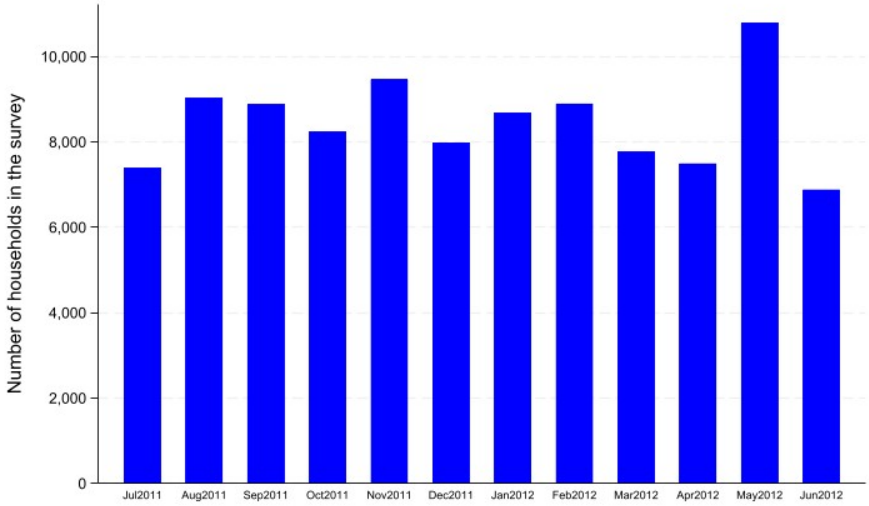


Figure 4.1 Number of sample households in each month in 2011-2012 India's household CES survey

The 2011-12 CES survey contains information on 101,662 households, covering 625 out of 640 districts encompassing the entire geographical area. Along with information on the demographic composition of each household, the CES provides data on household monthly expenditures for a variety of commodity categories: food and beverages, tobacco and alcohol, clothing and bedding, energy, durable goods and other miscellaneous goods. Within each of the five categories, a very detailed item classification is further adopted. In addition, the CES survey also provides the information on a range of household characteristics as well as the demographic characteristics of

household members. The household characteristics include household size, which is the total number of people within each household, principal industry and principal occupation, religion, whether the household owns any land, etc. For the demographic characteristics, the survey provides the data on the gender, age, marital status, level of education, etc.

Table 4.1 Distribution of households across districts in each month

	Mean	S.D.	Min.	Max.
Jul 2011	11.83	12.666	0	86
Aug 2011	14.45	12.485	0	90
Sep 2011	14.22	12.798	0	95
Oct 2011	13.19	12.157	0	74
Nov 2011	15.15	12.306	0	87
Dec 2011	12.78	12.299	0	104
Jan 2012	13.90	12.920	0	88
Feb 2012	14.24	12.063	0	86
Mar 2012	12.44	12.025	0	96
Apr 2012	12.00	12.038	0	81
May 2012	17.27	13.939	0	88
Jun 2012	11.02	12.269	0	93

These household-level survey data are combined with state-level information on prices which is obtained from the Ministry of Statistics and Programme Implementation, Government of India. It provides information on monthly pricing indices for four general categories of commodity in rural and urban areas at state-level, which are: (1) food, beverages, tobacco and alcohol, (2) energy, (3) clothing, bedding and footwear, and (4) miscellaneous goods. Note that there are a few missing values in the original pricing data, which are interpolated in STATA. While for urban areas in state Arunachal Pradesh, since the pricing information are all missing for each month, these missing values cannot be interpolated. Therefore, all the sample households located in urban Arunachal Pradesh (608 in total) are dropped due to this reason.

4.4.2 District-level climate data

District-level climate data are created based on Japanese global atmospheric re-analysis database (JRA-55), which is the second project conducted by the Japan Meteorological Agency (See

Kobayashi et al. (2015) for details of the database).[33] JRA-55 provides historical daily 3-hourly and 6-hourly weather data from 1958 up to the present at a spatial resolution of 0.5625°× 0.5625° latitude-longitude. This makes it very convenient to construct the district-level climate in India (i.e., 66°E-98°E and 4°N-38°N).

To construct district-level climate data, daily total precipitation, daily averages of temperature and relative humidity are first calculated for each grid from 1965-2011, based on which the gridded level climate variables used in this study are computed, including annual averages of total precipitation, temperature, as well as relative humidity throughout the time period. Besides, annual averages of HDDs and CDDs for each grid are also calculated based on daily mean temperature. More specifically, HDDs and CDDs over a period of time are calculated by using the commonly used American Society of Heating, Refrigerating, and Air-Conditioning method, which are defined as follows:

$$\text{CDD} = \sum_{i=1}^{n} (T_d - T_b)^+$$

$$\text{HDD} = \sum_{i=1}^{n} (T_b - T_d)^+$$

where T_d is the daily mean outdoor air temperature and T_b represents the base (threshold) temperature. The base temperatures used to calculate HDDs and CDDs are different for different countries in the world (e.g., 18.3℃ in the US, 15.5℃ in the UK). Following Bhatnagar et al. (2018), the base temperature takes the value of 18℃ in this study. "+" signifies only positive values

[33] JRA-55 is the first comprehensive reanalysis t cover the last half-century since the European Centre for Medium-Range Weather Forecasts 45-year Reanalysis, and is also the first to apply four-dimensional variational analysis to this period. JRA-55 provides a comprehensive atmospheric dataset suitable for the study of multidecadal variability and climate change.

accumulated over n days in the chosen time period (e.g., $n = 365$ if HDDs and CDDs are calculated over one year).

Following previous studies in India, the district-level climate data are constructed from the grid-level climate data by taking the weighted average of all grid points within 100 kilometres of each district's geographic centre, where the weights are the inverse of the distance from the centre.[34] This inverse-distance weighted average approach is consistent with previous studies in India (Guiteras, 2009; Taraz, 2018). The centroids of each district in India are obtained from the shapefiles of Indian districts in 2011. It is notable that the number of districts in India changes over years. For example, the total number of districts in India has increased from 593 in 2001 to 640 in 2011 and reached 773 in 2021. This explains why using the shapefiles in 2011 in this study. Figure 4.2-4.4 shows the Indian district-level annual averages of total precipitation, temperature, and relative humidity, respectively, which are constructed following the steps explained above.

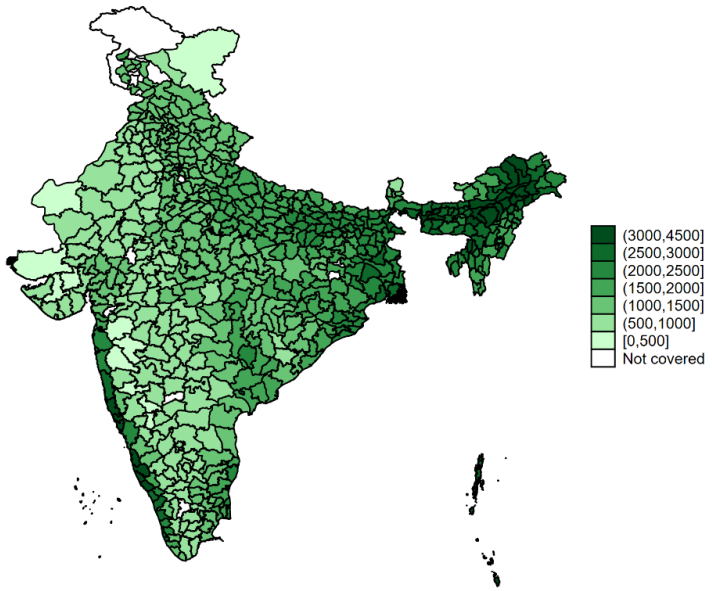


Figure 4.2 District-level annual precipitation in India

Note: The map shows the annual average of precipitation over 1965-2011 for each of 625 districts covered in the household survey. Data source: author's calculation based on JRA-55 dataset.

[34] Alternative radii other than 100km are also used, which does not appreciably affect the district-level climate data.

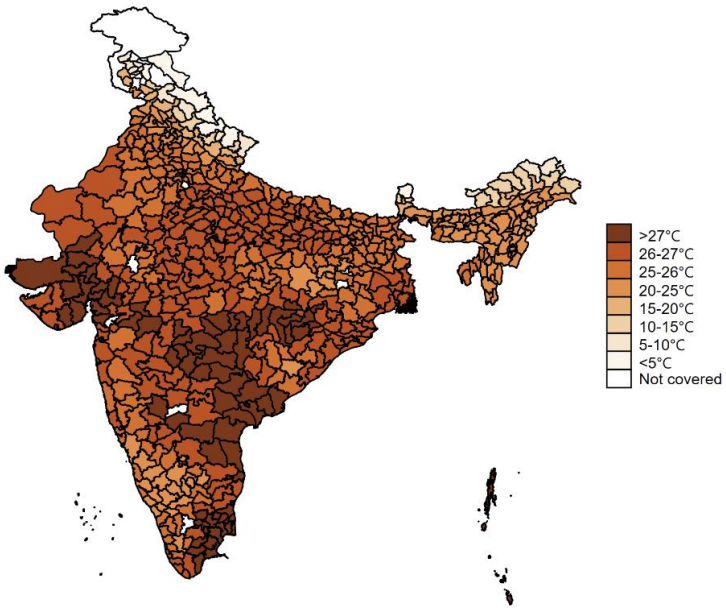


Figure 4.3 District-level annual temperature in India

Note: The map shows the annual average of temperature over 1965-2011 for each of 625 districts covered in the household survey. Data source: author's calculation based on JRA-55 dataset.

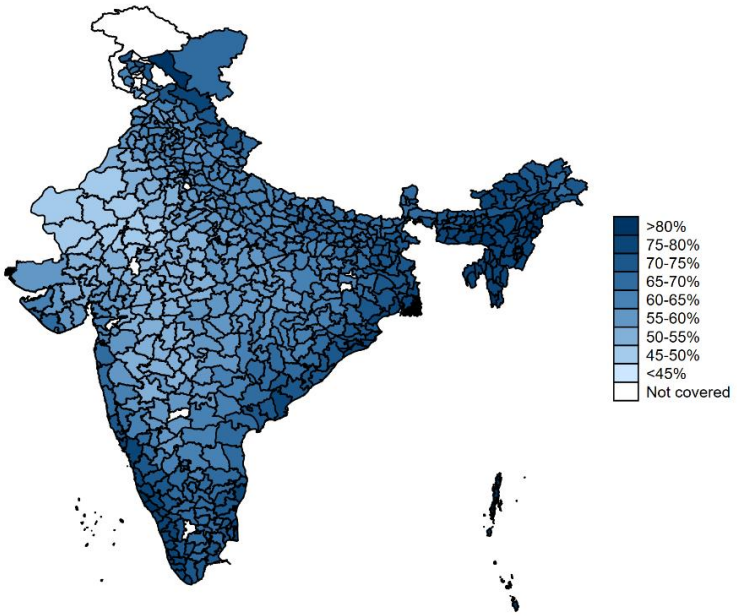


Figure 4.4 District-level annual relative humidity in India

Note: The map shows the annual average of relative humidity over 1965-2011 for each of 625 districts covered in the household survey. Data source: author's calculation based on JRA-55 dataset.

4.4.3 Variable definitions

As mentioned in 4.1, monthly pricing indices are for four general commodity categories, in order to make the expenditure categories in the household survey perfectly consistent with the pricing information, household monthly expenditures are disaggregated into the same categories used in the pricing indices: food, beverage, tobacco and alcohol; energy; clothing, bedding and footwear; and miscellaneous goods. Thus, total expenditure is the sum of expenditures on these four general commodity categories, and expenditures on the rest commodity groups, durable goods and housing are excluded from the analysis.

Climate variables enter into the demand system model through the scaling function. Climate is measured by average annual precipitation, HDDs and CDDs, as well as relative humidity. Besides, considering that the Barten's scaling method is originally proposed to estimate household equivalence of scale, number of adults, number of children in the household, as well as urban dummy which equal to one if the household resides in urban area are together used to define the scaling function.

Regional dummies and time dummies are used as the demand shifters in the demand system model. India can be divided into six regions: north, central, south, west, east and northeast, with northern India set to be the reference. Time dummies are defined according to the survey time of each sub-round: July-Sep (the reference), Oct-Dec, Jan-Mar and Apr-Jun. Table 4.2 displays the definitions of all variables included in the econometric analysis.

Table 4.2 Variable definitions

Variable	Definition
Total expenditure	Total expenditure over the last 30 days
Share of food, beverage, tobacco and alcohol	Expenditure share on food, beverages, tobacco and alcohol over the last 30 days
Share of energy	Expenditure share on energy used for fuel, light and household appliances over the last 30 days
Share of clothing, bedding and footwear	Expenditure share on clothing, bedding and footwear over the last 30 days

Share of miscellaneous goods	Expenditure share on miscellaneous goods and services including education, medical, entertainment, minor durables (such as umbrella), household consumables, conveyance over the last 30 days
Price of food, beverage, tobacco and alcohol	Monthly pricing index of food, beverages, tobacco and alcohol
Price of energy	Monthly pricing index of energy
Price of clothing, bedding and footwear	Monthly pricing index of clothing, bedding and footwear
Price of miscellaneous goods	Monthly pricing index of miscellaneous goods and services
Household demographics	
HH kid	Number of adults (aged below 18) in the household
HH adult	Number of children (aged above 18) in the household
Rural	Rural dummy, =1 if the household living in a rural area; =0 if the household living in an urban area.
Climate variables	
Annual mean prec. (mm)	District level total annual mean precipitation in 1965-2011
Annual mean temp. (℃)	District level annual mean temperature in 1965-2011
Annual mean HDDs	District level annual mean heating degree days in 1965-2011
Annual mean CDDs	District level annual mean cooling degree days in 1965-2011
Annual mean hum. (%)	District level annual mean relative humidity in 1965-2011
Prec_dum1	Dummy variable, =1 if annual mean prec.< 500mm; =0, otherwise
Prec_dum2 (reference)	Dummy variable, =1, if 500mm≤annual mean prec. < 1000mm; =0, otherwise
Prec_dum3	Dummy variable, =1, if 1000mm≤annual mean prec. < 1500mm; =0, otherwise
Prec_dum4	Dummy variable, =1, if 1500mm≤annual mean prec. < 2000mm; =0, otherwise
Prec_dum5	Dummy variable, =1, if annual mean prec. ≥ 2000mm; =0, otherwise
Temp_dum1	Dummy variable, =1 if annual mean temp.< 10℃; =0, otherwise
Temp_dum2	Dummy variable, =1, if 10℃ ≤ annual mean temp. < 15℃; =0, otherwise
Temp_dum3 (reference)	Dummy variable, =1, if 15℃ ≤ annual mean temp. < 20℃; =0, otherwise
Temp_dum4	Dummy variable, =1, if 20℃ ≤ annual mean temp. < 25℃; =0, otherwise
Temp_dum5	Dummy variable,

	=1, if annual mean temp. ≥ 25℃; =0, otherwise
Hum_dum1 (reference)	Dummy variable, =1 if annual mean hum.< 50%; =0, otherwise
Hum_dum2	Dummy variable, =1, if 50%≤annual mean hum. < 60%; =0, otherwise
Hum_dum3	Dummy variable, =1, if 60%≤annual mean hum. < 70%; =0, otherwise
Hum_dum4	Dummy variable, =1, if 70%≤annual mean hum. < 80%; =0, otherwise
Hum_dum5	Dummy variable, =1, if annual mean hum. ≥ 80%; =0, otherwise
Demand shifters	
North dummy (reference)	=1 if households located in Northern region; =0, otherwise
Central dummy	=1 if households located in Central region; =0, otherwise
West dummy	=1 if households located in Western region; =0, otherwise
South dummy	=1 if households located in Southern region; =0, otherwise
East dummy	=1 if households located in Eastern region; =0, otherwise
Northeast dummy	=1 if households located in North-eastern region; =0, otherwise
Sub_Round1 (reference)	=1 if households are interviewed in the first sub-round (Jul2011-Sep2011); =0, otherwise
Sub_Round2	=1 if households are interviewed in the second sub-round (Oct2011-Dec2011); =0, otherwise
Sub_Round3	=1 if households are interviewed in the third sub-round (Jan2012-Mar2012); =0, otherwise
Sub_Round4	=1 if households are interviewed in the fourth sub-round (Apr2012-Jun2012); =0, otherwise

4.4.4 Summary statistics

Summary statistics of all non-climate variables used in this analysis are presented in Table 4.3. The average monthly total expenditure converted into US dollar is $138.1 using the exchange rate between Indian rupee and US dollar at the end of 2011 (1USD=53.266 INR), with great variation across households in the survey ranging from $0.75 to $11,104.64. The expenditure on food, beverage, tobacco and alcohol accounts for the largest proportion of total expenditure, which is 54.8%. The proportion of households' consumption expenditure on miscellaneous goods is 27.4%, which is second to that on food, beverage, tobacco and alcohol. The expenditure on energy, as

well as clothing, bedding and footwear take up relatively small shares, 7.11% and 10.7%, respectively.

The average household size in the survey is about 4.59, which is close to that in 2011 census of India (4.45). In particular, the average number of adults and children in each household is 3.03 and 1.56, respectively. 59.3% of the households surveyed reside in rural areas. 23.1% of sample households reside in south India. Households living in the west and northeast, on the other hand take up relatively small proportion, which is 12.1% and 13.4%, respectively. The proportion of households in the north, central and east India account 16.4%, 17.5% and 17.4%, respectively. Households in the survey are distributed equally across the four sub-rounds.

Table 4.3 Summary statistics for all non-climate variables used in the analysis

	(1)	(2)	(3)	(4)
Variable	Mean	S.D.	Min.	Max.
Total expenditure (US dollar in 2011)	138.062	122.446	0.751	11,104.64
Share of food, beverage, tobacco and alcohol	0.548	0.128	0	1
Share of energy	0.0711	0.0885	0	0.745
Share of clothing, bedding and footwear	0.107	0.0508	0	0.673
Share of miscellaneous goods	0.274	0.127	0	0.982
Price of food, beverage, tobacco and alcohol	113.2	3.994	98.60	135.3
Price of energy	120.3	5.582	98.13	151.9
Price of clothing, bedding and footwear	118.8	6.059	102.2	150.3
Price of miscellaneous goods	112.4	3.903	103	135.2
HH kid	1.562	1.454	0	17
HH adult	3.033	1.532	1	26
Rural	0.593	0.491	0	1
North	0.164	0.370	0	1
Central	0.175	0.380	0	1
West	0.121	0.326	0	1
South	0.231	0.422	0	1
East	0.174	0.380	0	1
Northeast	0.134	0.340	0	1
Sub_Round1	0.251	0.433	0	1
Sub_Round2	0.250	0.433	0	1
Sub_Round3	0.249	0.432	0	1
Sub_Round4	0.250	0.433	0	1

As shown in Table 4.4, the mean of average annual temperature across the 625 districts covered in the household survey is 24.4℃. Ladakh in north India is the coldest district where the average annual temperature is about -10℃ due to its high altitude, while the hottest districts among the covered area is Guntur and Krishna in south India and the annual average temperature in both districts is over 28℃. Most districts in the survey sample (73%) locate in areas with annual average temperature above 25℃. The proportion of districts where annual average temperature ranges from 20-25℃ is about 18%, and districts with average annual temperature below 15℃ only account for about 7% to total districts covered in the survey. The average of annual HDDs and CDDs is 325.7 and 2665, respectively. CDDs is much larger than HDDs as daily average temperature in most districts is greater than the 18.3℃, the base temperature used for calculating degree days.

The average annual precipitation in the sample is 1,588 mm, which varies considerably across districts from only 386mm in Ahmednagar in west India to over 4,000mm in Udipi in south India. Districts with average annual precipitation over 1,000mm take up about 78%, and in particular, annual precipitation is even over 2,000mm for about 26% districts. Only 2% districts are located in drier region with annual precipitation less than 500mm. The average annual humidity across districts ranges from the minimum of 46% to 83%, the mean of which is 65%. About 40% districts have an annual humidity between 60% and 70%. The districts with humidity in 50%-60% or 70%-80% accounts for the similar proportion, which is 27% or 28%, respectively.

Table 4.4 Summary statistics for climate variables

	(1)	(2)	(3)	(4)
Variable	Mean	S.D.	Min.	Max.
Annual mean temp. (℃)	24.434	4.767	-10.065	28.159
Annual mean HDDs	325.697	1052.63	0	10,202.68
Annual mean CDDs	2665.04	796.92	0	3,694.761
Annual mean prec. (mm)	1587.886	708.537	386.386	4,028.788
Annual mean hum. (%)	65.061	8.15	46.36	82.98

Temp1 (<10℃)	0.032	0.176	0	1
Temp2 (10-15℃)	0.024	0.153	0	1
Temp3 (15-20℃)	0.0304	0.172	0	1
Temp4 (20-25℃)	0.1792	0.384	0	1
Temp5 (>25℃)	0.7344	0.442	0	1
Prec1 (<500mm)	0.0176	0.132	0	1
Prec2 (500-1000mm)	0.192	0.394	0	1
Prec3 (1000-1500mm)	0.3216	0.467	0	1
Prec4 (1500-2000mm)	0.2064	0.405	0	1
Prec5 (>2000mm)	0.2624	0.44	0	1
Hum1 (<50%)	0.0208	0.143	0	1
Hum2 (50-60%)	0.2688	0.444	0	1
Hum3 (60-70%)	0.408	0.492	0	1
Hum4 (70-80%)	0.2768	0.448	0	1
Hum5 (>80%)	0.0256	0.158	0	1

4.4.5 Empirical specifications

The analysis in this study starts from firstly estimating the baseline model with its three theoretical restrictions applied (i.e., adding-up, homogeneity and symmetry) using the whole sample of households. Then a set of sensitivity analyses are conducted by modifying the baseline model in different ways. These modifications include: (1) replacing the AIDS by the QUAIDS model in the baseline model; (2) allowing the homogeneity and symmetry assumptions not to be necessarily satisfied (the adding-up restriction is always mechanically satisfied in the AIDS model); (3) estimating the baseline model with weighted sample households considering that households are not evenly distributed across survey time periods and districts; (4) using Ray's scaling method instead of Barten's method. Besides, the baseline model is also estimated using the subsamples of urban/rural, high-/low-income households to investigate whether the scaling effect of climate varies over different groups of households.

As different measures of climate are adopted in this study, there would be too many regression results presented in the analysis if estimating both the baseline model and its modifications introduced above for each specification of climate. Therefore, annual averages of precipitation, HDDs and CDDs, as well as relative humidity are used as the key measures of climate which have

also been widely used in existing literature according to the literature review. The scaling effects of climate measured in other ways are only estimated using the baseline model, viz. replacing the preferred measure of climate by other specifications of climate.

To examine the scaling effect of climate, we also need to define the reference household, for which the scaling function is always equal to one. In this study, the single-adult household in urban New Delhi is defined as the reference. Given the annual average precipitation, HDDs, CDDs and relative humidity, which is 1231.29mm, 156.49, 2959.8 and 59.72%, respectively, the scaling function $\mathrm{m}(\mathbf{z})$ defined in Eq. (4.13), therefore takes the following form:

$$\mathrm{m}(\mathbf{z}) = 1 + \rho_1(Adult - 1) + \rho_2 Kid + \rho_3 Rural + \rho_4(Prec - 1231.29) + \rho_5(HDD - 156.49) + \rho_6(CDD - 2959.8) + \rho_7(Hum - 59.72) \quad (4.27)$$

4.5 Regression results

The parameters of the demand system model are estimated by using the non-linear seemingly unrelated regression package (nlsur) in STATA 17 (Appendix 3 shows the codes for estimating the baseline model). Due to the adding-up restriction, one of the expenditure share equations in the demand system needs to be dropped in order to make the model estimable. No information will be lost when dropping one share equation from the estimation and the estimation results do not depend on which equation is dropped. The parameters of the dropped equation can be obtained by exploiting their dependence on the other parameters of the system. In this study, the expenditure share equation for miscellaneous goods is dropped from estimation. Lastly, the standard errors are clustered at district level considering that the climate variables change across regions rather than individual households.

4.5.1 Results from the baseline model

The complete estimation results for the baseline model are presented in Table 4.5. The intercept and the coefficients on the logarithm of total expenditure (β_i) are found to display spatial and seasonal variation in each share equation. The Wald test strongly rejected the joint zero restrictions on both of them (See Table 4.6). The direct interpretation of the α_i coefficients on demand shifters in each share equation is quite simple. For example, compared to households living in north India, the expenditure shares of food, beverage, tobacco and alcohol, and energy, ceteris paribus, fall by 4.2% and 1.2%, while that of clothing, bedding and footwear, and miscellaneous goods are about 2% and 3.4% higher for the households in Central India. On the other hand, compared to households' expenditure patterns in July-Sep, households' shares of expenditure on food, beverage, tobacco and alcohol, and miscellaneous goods in Oct-Dec are 1.1% and 0.5% less, while the shares on clothing, bedding and footwear, and energy are 1.3%, 0.4% higher. Other regional and time demand shifters in the intercept can also be interpreted in the similar way.

Table 4.5 Estimation results for the baseline model

Variable	Symbol	Food, beverage, tobacco & alcohol	Clothing, bedding & footwear	Energy	Miscellaneous goods
Intercept					
Constant	α_{i0}	0.6166***	0.0508***	0.1348***	0.1977***
		(0.0038)	(0.0027)	(0.0027)	(0.0031)
Central dummy	α_{i1}	-0.0416***	0.02***	-0.0122***	0.0339***
		(0.0041)	(0.0029)	(0.0029)	(0.0034)
West dummy	α_{i2}	-0.0111**	-0.0032	-0.0125***	0.0267***
		(0.0054)	(0.0034)	(0.003)	(0.0043)
South dummy	α_{i3}	-0.0265***	-0.0033***	-0.0399***	0.0697***
		(0.0048)	(0.0008)	(0.0027)	(0.0043)
East dummy	α_{i4}	-0.0256***	0.0241***	-0.0169***	0.0184***
		(0.0047)	(0.0038)	(0.0028)	(0.0037)
Northeast dummy	α_{i5}	-0.0068	0.0164***	-0.0166***	0.007
		(0.0079)	(0.005)	(0.0038)	(0.0058)
Time_dum2	α_{i6}	-0.0111***	0.0126***	0.004***	-0.0054***
(Oct-Dec)		(0.0025)	(0.0013)	(0.0005)	(0.002)
Time_dum3	α_{i7}	-0.0022**	-0.0022**	0.0085***	-0.0041***
(Jan-Mar)		(0.0011)	(0.0009)	(0.0008)	(0.0011)
Time_dum4	α_{i8}	0.0061***	-0.0099***	0.0069***	-0.003**

(Apr-Jun)		(0.0014)	(0.0012)	(0.0007)	(0.0014)
Price of food, beverage, tobacco, alcohol	γ_{i1}	-0.1526***	0.0453***	-0.0444***	0.1517***
		(0.0032)	(0.0024)	(0.0025)	(0.003)
Price of clothing, bedding, and footwear	γ_{i2}		0.0296***	0.0005	-0.0754***
			(0.0047)	(0.0017)	(0.0051)
Price of energy	γ_{i3}			-0.0066**	0.0506***
				(0.0028)	(0.0021)
Price of miscellaneous goods	γ_{i4}				0.1269***
					(0.0051)
Ln (Expenditure)					
Constant	β_{i0}	-0.0007	-0.0003	0.0037	-0.0027
		(0.0045)	(0.0063)	(0.0036)	(0.0062)
Central dummy	β_{i1}	0.0158***	0.0056*	0.0059**	-0.0273***
		(0.0049)	(0.0033)	(0.0028)	(0.0049)
West dummy	β_{i2}	0.0163***	-0.0002	-0.0065*	-0.0096
		(0.0052)	(0.0053)	(0.0038)	(0.0064)
South dummy	β_{i3}	0.0063	0.0089*	0.0042	-0.0194***
		(0.0056)	(0.0047)	(0.0041)	(0.0052)
East dummy	β_{i4}	0.0069***	0.0114***	-0.0041***	-0.0141***
		(0.0024)	(0.0024)	(0.0009)	(0.0032)
Northeast dummy	β_{i5}	0.0147***	0.0007	-0.0075***	-0.0079**
		(0.0026)	(0.0029)	(0.0011)	(0.0033)
Time_dum2	β_{i6}	0.0075***	-0.0053**	-0.0067***	0.0045
		(0.0026)	(0.0025)	(0.0011)	(0.003)
Time_dum3	β_{i7}	0.0793	-0.0689**	0.0163	-0.0267
		(0.0773)	(0.0289)	(0.0222)	(0.0605)
Time_dum4	β_{i8}	-0.0864***	0.0361***	-0.0432**	0.0935***
		(0.0275)	(0.0139)	(0.0192)	(0.0294)
Scaling variables					
HH adult	ρ_1	0.2921***			
		(0.0059)			
HH kid	ρ_2	0.2317***			
		(0.0048)			
Rural	ρ_3	-0.0638***			
		(0.005)			
Annual mean prec.(mm)	ρ_4	-0.00004***			
		(4.05e-06)			
Annual mean HDDs	ρ_5	0.00005***			
		(5.67e-06)			
Annual mean CDDs	ρ_6	0.00007***			
		(5.79e-06)			
Annual mean hum. (%)	ρ_7	0.0132***			
		(0.0009)			

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

Table 4.6 Wald test for joint significance of the intercepts and parameters of expenditure

Wald test	P-value	Conclusion
$\alpha_i(\mathbf{d}) = 0$, for all i	0.000	Rejected
$\beta_i(\mathbf{d}) = 0$, for all i	0.000	Rejected

Turning to the effects of scaling variables on households' costs of living, the regression results suggest that given the single-adult households as the reference, a household with two adults spends about 29% higher to achieve the same level of utility, while one additional child increases the cost by about 23%. These estimated household equivalence scales appear to be plausible. Households with the same composition spend about 6.4% less in rural areas than in urban areas to reach the same welfare level. All of the climate scaling variables are significant at the 1% level. More specifically, average annual precipitation negatively and significantly affects households' costs of living: holding all others constant, one standard deviation increase in annual precipitation will decrease households' spending by about 2.8% keeping the welfare level unchanged. Both HDDs and CDDs significantly increase households' costs of living. For one standard deviation increases in HDDs or CDDs, households' total expenditure will rise by about 5.3% or 5.6%, respectively. Annual humidity also has significant and negative effect on households' living costs: with a one percent increase in annual average humidity, household's living costs will increase by 1.3%. The results suggest that higher annual HDDs and CDDs, as well humidity are disamenities while more annual precipitation is an amenity to Indian households.

The interpretation of the price and total expenditure parameters is best discussed using price and expenditure elasticities, which can be calculated from Eq. (4.22)-(4.26) based on the estimated coefficients obtained in Table 4.5. As presented in Table 4.7, the price and expenditure elasticities of all commodity categories are calculated for the average household where all variables involved are set to be at the sample means. The estimated expenditure elasticities of each commodity group are significantly different from 0 at 1% level. More specifically, the expenditure elasticities of each

of the four commodity categories are all very close to 1. The compensated and uncompensated own-price elasticities are all significant and exhibit the expected negative signs for each of the four commodity groups. This implies that for each commodity group, an increase in the commodity price would result in a decrease in the demand for that commodity. In addition, the estimated cross-price elasticities suggest that there exist complementarity and substitution effects among different commodities. Food, beverage, tobacco and alcohol, clothing, bedding and footwear, as well as miscellaneous goods are substitutes, while energy and food, beverage, tobacco and alcohol seem to be complemented to each other.

Table 4.7 Price and expenditure elasticities using the estimates from the baseline model

	Price elasticity				Expenditure elasticity
	Food, beverage, tobacco & alcohol	Clothing, bedding & footwear	Energy	Miscellaneous goods	
Uncompensated					
Food, beverage, tobacco & alcohol	**-1.2878*** (0.02571)**	0.0816*** (0.0051)	-0.0829*** (0.0065)	0.2543*** (0.0556)	1.0159*** (0.0399)
Clothing, bedding & footwear	0.6758*** (0.0962)	**-0.5796*** (0.067)**	0.0145 (0.0286)	-0.9665*** (0.2105)	0.9344*** (0.1408)
Energy	-0.386*** (0.0546)	0.0074 (0.0167)	**-1.056*** (0.0242)**	0.542*** (0.0998)	0.9511*** (0.069)
Miscellaneous goods	0.5512*** (0.0422)	-0.2755*** (0.0192)	0.1841*** (0.0111)	**-1.6574*** (0.0955)**	1.0045*** (0.063)
Compensated					
Food, beverage, tobacco & alcohol	**-0.7294*** (0.0257)**	0.1541*** (0.0051)	0.0261*** (0.0065)	0.5333*** (0.0556)	
Clothing, bedding & footwear	1.1586*** (0.0962)	**-0.517*** (0.067)**	0.1087** (0.0286)	-0.7252*** (0.2105)	
Energy	0.1335** (0.0547)	0.0748*** (0.0167)	**-0.9546*** (0.0242)**	0.8015*** (0.0998)	
Miscellaneous goods	1.1068*** (0.0422)	-0.2034*** (0.0192)	0.2925*** (0.0111)	**-1.3798*** (0.0955)**	

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

4.5.2 Sensitivity analysis

In this section, the robustness of the scaling effects of climate are tested by: (1) using the QUAIDS model instead, (2) relaxing the theoretical homogeneity and symmetry restrictions, (3)

weighting the sample households, and then I conduct subsample analysis as well as an analysis with alternative measures of climate, (4) using Ray's scaling technique instead, and (5) using different climate specifications.

Table 4.8 presents the estimation results obtained from the QUAIDS model with the same scaling function (Barten's scaling method in Eq. (4.13) used in the baseline model. Note that to save space only the estimated coefficients of the scaling variables which we mainly focus on are reported. It can be seen from Table 4.8 that the estimated cost of adding one more adult or child into the reference single-adult household increases total expenditure by about 27% or 20%, respectively. Households living in urban areas cost about 3.2% higher than living in rural. For the effect of climate on households' living costs, on the other hand, there are no significant differences between the QUAIDS and AIDS model. This result can, or at least partly supports the choice of AIDS model in this study for its simplicity and convenience although the QUAIDS model can better capture the non-linear shape of Engel curves than the AIDS model.

Table 4.8 Selected results for the extended QUAIDS model

Scaling variables (ρ)	
HH adult	0.2684***
	(0.0061)
HH kid	0.1996***
	(0.0048)
Rural	-0.0324***
	(0.0034)
Annual mean prec.(mm)	-0.00004***
	(3.99e-06)
Annual mean HDDs	0.00003***
	(4.20e-06)
Annual mean CDDs	0.00005***
	(5.02e-06)
Annual mean hum. (%)	0.01234***
	(0.001)

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

To investigate whether relaxing the homogeneity and symmetry restrictions shown in Eq. (4.9) and (4.10) can affect the amenity value of climate to households, the baseline model without the two restrictions is estimated, the regression results of which are reported in Table 4.9. It shows that all the scaling variables are still significant and whether imposing or not the two restrictions on the regression model does not significantly affect either the direction or magnitude of the estimated coefficients on climate variables. Besides, the Wald test for both the homogeneity and symmetry hypothesis are also presented in Table 4.9 and the test statistics suggest that neither of the two hypotheses are satisfied. This is not a surprising finding (see for example, Deaton and Mullbauer, 1980; Abdulai et al., 1999; Duffy, 2003). As shown in Duffy (2003), the restrictions of economic theory are more likely to be satisfied in the long-run than in the short-run.

Table 4.9 Selected results for the baseline model without homogeneity and symmetry restrictions

Scaling variables (ρ)			
HH adult			0.296***
			(0.0059)
HH kid			0.2319***
			(0.0049)
Rural			-0.0282***
			(0.0025)
Annual mean prec.(mm)			-0.00005***
			(4.43e-06)
Annual mean HDDs			0.00004***
			(4.95e-06)
Annual mean CDDs			0.00006***
			(5.06e-06)
Annual mean hum. (%)			0.01337***
			(0.0009)
Wald test	Test statistic	P-value	Conclusion
Homogeneity	276.92	0.000	Rejected
Homogeneity and Symmetry	1,638.79	0.000	Rejected

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

The number of sample households varies significantly across districts in the survey, ranging from only 6 households located in New Delhi to 879 households located in West Tripura. This means that each observation has a different probability of being sampled. In order to take this into

account, the baseline model is estimated using weighted observations. More specifically, each observation is weighted by the inverse of its probability of being sampled. In this study, each observation is weighted directly using the sampling weight provided by the survey itself. The weight in survey data simply indicates how many households in the universe is represented by each sample household. For example, if the weight is 100 for a particular household in the survey, it means that the probability of this household being sampled is 1/100, or equivalently, this household represents 100 households in the universe.

The regression results are presented in Table 4.10. It suggests that the estimated equivalence scale of adult and child are similar with that obtained from the baseline model using unweighted observations, while compared to urban households with the same composition, the cost decreases by about 12% (which is greater than 6% obtained from the baseline model) for urban households for a given composition of the household. For climate variables, the effects of climate variables on households' costs of living are still significant and keep the same direction, while the magnitudes are slightly less than those obtained without weighting observations.

Table 4.10 Selected results for the baseline model using weighted observations

Scaling variables (ρ)	
HH adult	0.2794***
	(0.0096)
HH kid	0.2209***
	(0.0073)
Rural	-0.1216***
	(0.0104)
Annual mean prec.(mm)	-0.00003***
	(3.60e-06)
Annual mean HDDs	0.00003***
	(5.52e-06)
Annual mean CDDs	0.00003***
	(5.15e-06)
Annual mean hum. (%)	0.0084***
	(0.0009)

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

The baseline regression results suggest that the cost of living for an urban household is about 6% higher than the household living in a rural area keeping all other characteristics the same. Furthermore, in order to examine whether the effect of climate on households' costs of living are

different among urban/rural households, the baseline model is estimated by using urban and rural subsamples separately. The results for urban and rural households are presented in column (1) and (2) respectively in Table 4.11. The estimated equivalence scale of adults and children for urban households are slightly larger than rural households, while there are no significant differences on the scaling effect of climate on households' cost of living among urban/rural households.

Table4.11 Selected results for the baseline model for urban/rural subsample

Scaling variables (ρ)	(1) Urban	(2) Rural
HH adult	0.3606***	0.2312***
	(0.01)	(0.0059)
HH kid	0.2528***	0.1927***
	(0.0072)	(0.0055)
Annual mean prec.(mm)	-0.000036***	-0.00005***
	(4.45e-06)	(5.76e-06)
Annual mean HDDs	0.00003***	0.00003***
	(3.93e-06)	(5.14e-06)
Annual mean CDDs	0.00009***	0.00005***
	(8.35e-06)	(6.62e-06)
Annual mean hum. (%)	0.01597***	0.0106***
	(0.0012)	(0.0009)

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis

To investigate whether climate change will exacerbate or reduce inequality among the rich and poor households, the baseline model is also estimated using high- and low-income subsamples, respectively. The quantile method is applied to simply divide households into two categories: high-income and low-income, with 50th quantile of monthly household total expenditure equal to about 5,933 Rupees ($111) in the survey data. The regression results are reported in Table 4.12, in which column (1) and column (2) presents the results for high-income and low-income households, respectively. The estimated equivalence scale of adults for low-income households is about 0.35 which is slightly larger than that for high-income households (0.29), while the cost of additional one child is similar between high-income and low-income households. The impact of climate on the costs of living for households on the other hand, varies significantly between high- and low-income households. One standard deviation increases in annual precipitation is associated with a 5.7% and 2.1% reduction in the costs of living for high- and low-income households, respectively. In contrast, for one standard deviation increase in annual CDD, high-income households will

experience a larger increase of 8% in living costs while the costs of living for low-income households will only rise by 4%. Additionally, an increase of 1% of annual humidity will lead to a 2.4% increase in living costs for high-income households and a 0.8% increase for low-income households. This might be explained by that rich households consume more services, the production cost of which is more sensitive to climate.

Table 4.12 Selected results for the baseline model for high-/low-income subsample

Scaling variables (ρ)	(1) High-income	(2) Low-income
HH adult	0.2915***	0.3463***
	(0.007)	(0.0095)
HH kid	0.2324***	0.2543***
	(0.0064)	(0.0072)
Rural	-0.0486***	-0.0293***
	(0.0041)	(0.004)
Annual mean prec.(mm)	-0.00008***	-0.00003***
	(6.97e-06)	(4.78e-06)
Annual mean HDDs	0.00003***	0.00004***
	(3.90e-06)	(6.52e-06)
Annual mean CDDs	0.0001***	0.00005***
	(8.79e-06)	(6.33e-06)
Annual mean hum. (%)	0.0241***	0.0077***
	(0.0016)	(0.0008)

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

Table 4.13 reports the results from the baseline model replacing Barten's scaling function by Ray's scaling function, which allows the scaling function to vary with both income and prices. From Eq. (4.15) and (4.16), it can be seen that the only difference between the modified AIDS using the two scaling techniques lies in the coefficient of expenditure (β). Note that the demand shifters are introduced into the AIDS model only through the intercepts (α) when using Ray's scaling method. The results suggest that there are no big differences in both the magnitudes and significance of estimated coefficients on scaling variables between Barten's scaling method and more general Ray's scaling method.

Table 4.13 Selected results for Ray's scaling method

Scaling variables (ρ)	
HH adult	0.2545***
	(0.0055)
HH kid	0.2099***
	(0.0047)
Rural	-0.0799***
	(0.0058)
Annual mean prec.(mm)	-0.00004***
	(3.75e-06)
Annual mean HDDs	0.00003***
	(3.65e-06)
Annual mean CDDs	0.00003***
	(2.75e-06)
Annual mean hum. (%)	0.0072***
	(0.0006)

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

Finally, the measure of climate in the baseline model are replaced by alternative climate measures, including: (1) using average annual temperature instead of HDDs and CDDs, and (2) a set of binned climate variables. The estimation results are presented in Table 4.14. The estimated costs of adults and children does not vary with different climate measures. It is found that the effect of annual temperature on households' costs of living is significantly negative, suggesting that Indian households regard higher temperature as a disamenity. For the second measure of climate, five binned groups are created respectively based on the distribution of annual averages of temperature, precipitation and humidity in the dataset. To be precise, annual temperature is grouped into below 10℃, 10-15℃, 15-20℃, 20-25℃, and above 25℃. Annual precipitation is categorized into below 500mm, 500-1,000mm, 1,000-1,500mm, 1,500-2,000mm and above 2,000mm. Annual humidity is grouped into below 50%, 50-60%, 60-70%, 70-80% and above 80%. The reference household is located where annual temperature is between 15-20℃, annual precipitation is between 500-1,000mm, and annual humidity is below 50%.

Column (2) in Table 4.14 presents the estimated results of the binned climate variables. All of the coefficients on those climate bins are significant. Annual precipitation less than 500mm raises living costs of the reference household by 1.5%, while annual precipitation exceeding 1,000mm

significantly reduce the cost of the reference household. Annual temperature below 15℃ or above 20℃ will increase the living costs of the reference household, with the largest increase the largest occurring when annual temperature drops below 10℃. In addition, annual humidity between 50-60% will result in a reduction of living costs of the reference household by approximately 0.7%, while annual humidity exceeding 60% significantly raises households' costs of living, and in particular, the reference household will spend about 52% more if annual humidity rising above 80%, which is a very large impact. Districts with relative humidity over 80% are located in either costal states including islands (such as Kerala and Puducherry, Lakshadweep and the Andaman and Nicobar Islands), or the states in northeast with much precipitation, including Nagaland and Meghalaya. It is also found that households in India prefer mild climate. The cost of living is the lowest for households residing in the region where annual precipitation is between 1,500-2,000mm, annual temperature is around 15-20°C and annual humidity ranging from 50-60%, although no one district covered in the survey satisfies this combination of climate bins.

Table 4.14 Selected results for the baseline model with different measure of climate

Scaling variables (ρ)	(1)	(2)
HH adult	0.2869***	0.3142***
	(0.0054)	(0.0058)
HH kid	0.2243***	0.2493***
	(0.0046)	(0.0049)
Rural	-0.0734***	-0.0765***
	(0.0064)	(0.0066)
Annual mean prec.(mm)	-0.00004***	
	(4.14e-06)	
Annual mean tem. (℃)	0.0028***	
	(0.0004)	
Annual mean hum. (%)	0.012***	
	(0.0009)	
Prec_dum1 (<500mm)		0.0153**
		(0.0077)
Prec_dum3 (1000-1500mm)		-0.0146***
		(0.002)
Prec_dum4 (1500-2000mm)		-0.0639***
		(0.0085)
Prec_dum5 (>2000mm)		-0.0464***
		(0.0048)
Temp_dum1 (<10°C)		0.1603***
		(0.0339)
Temp_dum2 (10-15°C)		-0.0059
		(0.0986)
Temp_dum4 (20-25°C)		0.0283***
		(0.0048)

Temp_dum5 (>25°C)	0.0727***
	(0.0054)
Hum_dum2 (50%-60%)	-0.0071***
	(0.0014)
Hum_dum3 (60-70%)	0.0982***
	(0.0076)
Hum_dum4 (70-80%)	0.1699***
	(0.0144)
Hum_dum5 (>80%)	0.5183***
	(0.176)

***, ** and * indicates significance at 1%, 5% and 10%, respectively; standard errors clustered at district level in parenthesis.

4.6 Discussion

The results in this study suggest that for one standard deviation increase in annual HDDs or CDDs, Indian households will experience an increase by 3.1%-5.3% or 2.4%-8% in living costs, respectively, and an increase of 1% in annual humidity will raise households' costs of living by 0.7%-2.4%. This indicates that higher annual HDDs, CDDs and humidity are significant disamenities to Indian households. These findings are similar with those found in other countries in previous studies as shown in section 4.2. It is also found that Indian households regard higher annual temperature as a disamenity. For 1℃ rise in annual temperature, Indian households' costs of living will increase by about 0.4% on average. This is consistent with the finding by Maddison (2003) that any increase in temperatures in low latitude (or tropical) countries is likely to result in a large increase in cost of living. This can be explained by that most areas in India are already very warm, with more than 90% districts covered in the analysis whose annual temperatures are above 20℃, the hotter climate may result in the significant increase in the cost of cooling demand, human health, etc. On the contrary, higher precipitation per year is found to be a significant amenity to Indian households in this study. One standard deviation increase in annual precipitation will result in a reduction of 2.1%-5.7% in households' costs of living. It seems to be a bit difficult to explain why a higher level of annual precipitation is an amenity or a disamenity to households, and the empirical evidence on preferences regarding annual precipitation are also mixed among these existing studies in US and EU countries. In general, the negative effect of higher precipitation per

year on Indian households' cost of living could be due to the fact that warmer regions with more precipitation would have more effects on boosting the amenity value of greater precipitation. Besides, an area with more precipitation would be quite lush as opposed to a desert.

Based on the estimation results, we can further compute the equivalence scale of climate for households located in different cities associated with different types of climate for a given composition of household. Climate equivalence scales provide the information on the changes in households' costs of living due to the differences in climate. Table 4.15 presents the climate equivalence scales for ten major cities in India, which are calculated using the estimated coefficients in the baseline model (Table 4.5), and the single-adult households located in New Delhi is taken as the reference. The ten cities listed in Table 4.15 are located in different regions of India and thus the climate varies significantly across these cities (see Figure 4.5). The reference city New Delhi is located in North India, while Kanpur and Kolkata are in Central and East India, respectively. Mumbai, Surat, Pune, and Ahmedabad are situated in west India. Bangalore, Hyderabad and Chennai are positioned in the southern region of the country.



Figure 4.5 Map of ten major cities in India

All the major cities are quite warm. Average annual temperatures are 25℃ or above for all of the ten cities. Ahmedabad is the hottest amongst the ten cities, with an annual temperature reaching 27.9℃. Thus, their HDDs are very low while CDDs are quite high. Annual precipitation varies significantly from about 400mm to over 2,000mm across the ten cities. It is over 2,000mm in Mumbai, Kolkata and Chennai and less than 1,000mm in Ahmedabad, Bangalore and Pune, while annual precipitation in the rest three cities, Kanpur, Surat and New Delhi lies between 1,000-2,000mm. Besides, annual relative humidity in Mumbai, Kolkata and Chennai are all over 70% and the lowest is for the city of Ahmedabad (54%).

Table 4.15 Estimated climate equivalence scale for major cities in India

City	Prec (mm)	HDD	CDD	Hum (%)	Climate Equivalence scale
Chennai	2,326	0.02	3,398.33	76.43	1.198
Kolkata	2,403	12.34	2,955.14	75.45	1.15
Mumbai	2,476	0.06	3,045.51	70.64	1.089
Bangalore	726	0.01	2,423.33	66.29	1.062
Hyderabad	690	0.12	3,215.81	60.13	1.04
Surat	1,386	0.33	3,276.47	61.78	1.036
Kanpur	1,655	85.47	3,170.6	61.16	1.016
New Delhi (reference)	1,231	156.49	2,959.8	59.72	**1.00**
Pune	412	0.23	2,650.65	58.48	0.989
Ahmedabad	838	1.54	3,609.1	54.19	0.985

Note: climate equivalence scales for each major cities are calculated based on the estimation results from the baseline model (shown in Table 4.5). Taking a household living in New Delhi as the reference, climate equivalence scales tell us by how much will the identical household's cost of living change if moving to other cities.

From Table 4.15, it can be seen that the climate equivalence scale in Chennai is 1.205, the largest among the ten cities, which suggests that the climate of Chennai raises households' costs of living by 20.5% compared with the climate of New Delhi. This is mainly because annual relative humidity in Chennai (76.4%) is much higher than New Delhi (59.7%), even though annual precipitation in Chennai is also higher than New Delhi. According to the findings in this study, Indian households prefer more annual precipitation and less relative humidity, since higher level of humidity makes people feel hotter. The increase in households' costs of living due to the higher relative humidity

dominates the reduction associated with more precipitation in Chennai with the comparison of New Delhi, which explains the larger climate equivalence scales in Chennai than New Delhi. Similar for Kolkata and Mumbai.

The climate in Bangalore will raise the living costs of the reference household by 7.5%. Both annual HDDs and CDDs in Bangalore are lower than those in New Delhi. However, Bangalore has less annual precipitation and higher annual relative humidity, both of which result in an increase in households' costs of living. The climate equivalence scales in Hyderabad, Surat and Kanpur are slightly larger than New Delhi. Compared to the climate in New Delhi, the three cities have similar annual relative humidity, lower annual HDDs but higher annual CDDs. The climate equivalence scale in Hyderabad is a bit larger than Surat and Kanpur, which is because its annual precipitation is less than New Delhi while the precipitation in the other two cities are greater than New Delhi.

The climate equivalence scales in Pune and Ahmedabad are 0.997 and 0.983, respectively, which are less than the reference city. Although annual precipitation in Pune is the lowest among the ten cities, both its annual HDDs and CDDs are lower than New Delhi and thus the temperature in Pune is more pleasant. For Ahmedabad, its climate can reduce the reference households' costs of living by 1.7% which is mainly due to its lower annual relative humidity than New Delhi.

4.7 Conclusions

India is considered as one of the most vulnerable countries to climate change. This study empirically confirms that climate can significantly affect households' costs of living in India by estimating a system of demand equations with climate variables incorporated. It is found that more precipitation is a significant amenity to Indian households, while higher HDDs and CDDs, as well as relative humidity are disamenities. These findings are robust in terms of magnitudes and

significance according to the sensitivity analysis. Based on the estimation results, the climate equivalence scales are computed for households located in major cities in India.

This study seems not to have immediate policy implications. However, the knowledge on the amenity value of climate to households can help contribute towards building a better picture of the global economic effects from climate change (Tol, 2009), considering the impacts of climate change on households are different between high-income countries and in low-income countries which tend to be in tropical zones and are more vulnerable to climate change. This in turn facilitates the cost benefit analysis of policy responses to climate change (such as the strategy of global carbon emissions reductions).

There exist some limitations in this study. First, as mentioned in section 4.2, the equivalence scale approach also has its own limitations. Besides, by using this approach, it is assumed that the impact of climate is the same on households' demand for different commodities, which might be implausible. For example, the effect of a temperature shock on household energy demand is very likely to be different from recreation. In addition, in this study, I incorporate climate variables into the demand system model by using a very simple way, i.e., Barten's scaling method. For simplicity and the convenience of estimating climate equivalence scales, it is assumed that the scaling function is invariant to both prices and income, which might not be realistic.

There is also some future work that I plan to do. First, employing different and more general techniques (such as demographic translating and scaling methods) to incorporate climate variables into the demand system model and to compare the estimated amenity value of climate using these techniques. Second, based on the estimation results, it will be helpful to project the impact of future climate change on households' living costs in India. This can be achieved if gridded level projected weather data are available, and we can construct the district-level projected climate data using the same approach as this study. Such information is of course important to evaluate the overall impacts of climate change on household sector and therefore to the Government to design

climate mitigation and adaption policies. Finally, given the general availability of household expenditure data this equivalence scale approach could readily be applied to other developing countries.

Chapter 5 Conclusions, Limitations and Future Work

5.1 Conclusions

In this thesis, I deal with estimating household equivalence scales, economies of scale in household energy use, as well as the amenity value of climate to households, based on nationally representative household-level survey data from China and India. In chapter two, household equivalence scales in China are investigated by estimating the quadratic almost ideal demand system model with household characteristics included by means of Ray's scaling method. The cross-sectional household-level data used for the analysis are obtained from 2017 China Household Finance Survey. Taking the single-adult household as the reference, how much extra income does the household need is estimated in order to sustain the same level of welfare with one more person added. This is a worthwhile endeavour for two reasons. First, previous analyses for China have relied on the Engel's method which can be problematic. Second, the two-child policy in China has rencently been relaxed (by three-child policy), and thus it is interesting to investigate the cost of extra one more child added into the housheolds.

The estimation results reveal that taking the single-adult household as the reference, the living costs would rise by about 36% if a new child enters into a single-adult household. It needs to increase by around 56% for the second adult, while for the arrival of the third adult, the equivalence scale just rises by about 16%. For households with two adults, the scale value increases by about 18% for the first child, while the marginal effect of the second child on the scale value is just about 14%. These findings suggest that there exist significant economies of scale in Chinese households.

Chapter three continues with the investigation of economies of scale but switch attention from China to India and focus on one commodity. Specifically, the economies of scale in energy consumed by households in India is estimated using household level data obtained from the

Household Consumer Expenditure Survey conducted in 2011-2012. Two different concepts are employed to measure the household economies of scale in energy use, which are the household elasticity of energy demand estimated by using Kakwani and Son's method (Kakwani and Son, 2005), and the elasticity of per capita energy use with respect to household size estimated from the double hurdle model. This is an important undertaking because India is currently the world's most populous country and the average household size in India is declining at a time when there is concern about GHG emissions. Furthermore, there is a longstanding concern about the health impacts of fuel use in India and this too is likely to be impacted by changes in demographic composition. Because energy use in Indian households is diverse, failing to distinguish between different fuels might obscure important tendencies.

The results suggest that there exist significant economies of scale in household energy use in India. The estimated elasticities of per capita energy use for different types of energy indicate that the economies of scale for modern and cleaner energy (i.e., electricity and LPG) are greater than these for traditional and transition energy. In particular, it is found that there are diseconomies of scale in households' use of dung cake, which is interesting because it shows the importance of distinguishing fuels. Besides, the results from double hurdle model also demonstrate that it works better than tobit model since household size, as well as other factors (such as household head's gender and education level) affecting households' energy choice differ from those affecting energy expenditures.

Based on the same household survey data used in chapter three, chapter four estimates the amenity value of climate to Indian households using an approach similar to that used in chapter two, i.e., modifying the demand system model to incorporate climate variables into the analysis. Different from Chapter two however, for simplicity and the convenience of analysis, I choose as the benchmark model the almost ideal demand system model and Barten's price scaling method to include climate into the demand system. The value of climate amenities is thus revealed from

the role played by climate in determining households' costs of living. Insofar as climate is changing due to anthropogenic emissions of GHGs, determining the impact on households' living costs in an already generally hot, poor, and populous country is extremely important. Thus, I then use this approach to estimate climatic equivalence scales i.e., how much extra do households living in a particular climate require in order to achieve the same level of welfare as an otherwise identical household residing in some reference climate.

The results demonstrate that climate plays a significant role in determining households' costs of living in India. It is found that higher precipitation is an amenity, while higher HDDs and CDDs, and relative humidity are disamenities to Indian households. In particular, humidity is found to have surprisingly large impact on households' living costs in India. the corresponding climate equivalence scales for ten major cities in India suggest that taking New Delhi as the reference city, Chennai has the largest climate equivalence scale, indicating that the cost of living of an identical household in Chennai is largest among the ten cities due to its climate, while Ahmedabad has the lowest climate equivalence scale.

5.2 Limitations

While the conclusions drawn from this thesis provide important insights into household equivalence scales in China, economies of scale in energy use, as well as the amenity value of climate to Indian households, it is essential to recognize the limitations of this research. For chapter two, the chief limitation is about how aggregated price indices are processed in order to make them perfectly consistent with commodity categories. Due to the lack of price information on each particular commodity consumed by households in CHFS survey, the provincial level consumer price indices are used as the proxy of prices. There exist numerous discrepancies in the definitions of some commodity categories between the survey and the pricing indices, except for food, tobacco and alcohol, as well as clothing. This explains why households' total consumption

expenditure has to be disaggregated into four commodity groups (i.e., food, tobacco and alcohol, clothing, and others). The pricing indices for the first three commodity groups are used directly as the measure of prices in the demand system, while the pricing index for the others commodity is unavailable, and thus it is measured by treating the general CPI as the weighted average of the four commodity categories, using the expenditure share of each category as the weights. The resulting pricing index for others commodity group, however, may be problematic.

The household survey data used in chapter three and four are obtained from the 2011-2012 India's Household Consumer Expenditure Survey. Although it might be somewhat out of date (probably due to the pandemic), it is still the most recent available nationally representative survey data. Except for the outdated dataset, for chapter three, another limitation is that although it provides very detailed information on the fuels used by the household, the survey data does not provide any information on accommodation characteristics, such as house sizes, number of floors, number of rooms, etc. These factors are found to have significant impact on household energy consumption in the existing literature.

For chapter four, by using the equivalence scales approach to estimate the amenity value of climate, only the costs of climate can be estimated insofar as climate alters the household's pattern of demands. Another limitation is related to the method employed to incorporate climate variables into the demand system. Barten's price scaling method is used because of its simplicity and the convenience of estimating climate equivalence scales. By using this method, however, it is assumed that the scaling function is invariant to both prices and income. In addition, annual averages for rainfall and humidity are used as the measure of climate, which might be considered very simplistic in a country whose climate is governed by the monsoon.

5.3 Possible future work

Regarding possible future work, I believe that it would be worthwhile to continue the research on estimating the economies of scale in household energy use, as well as the amenity value of climate to households particularly in developing countries. Considered that climate change has become one of the biggest challenges in the world, and many countries have been experiencing a decline in the average household size, investigating the economies of scale in household energy use is helpful to understand the impact of the decline in household size on overall energy use, and the environmental implications of the loss of economies of scale in household energy use. Besides, the amenity value of climate to households are definitely not neglectable in the analysis of the overall impact of climate change to humans.

Given the limitations of this thesis discussed above, there are many potential areas for future research. First, it might be worth investigating the household economies of scale in energy use either in India or other developing countries with the incorporation of accommodation characteristics with the availability of relevant data. In addition, it would be interesting to estimate the environmental implications of the changes in energy use due to changes in household size (such as carbon emissions) if the information on the quantities of different types of energy consumed by households are available.

Second, it would be worthwhile to apply the same method used in chapter four to other developing countries to see if we can get similar results. China can be a very suitable country for estimating the amenity value of climate to households (and this was indeed my original intention). China is the second largest country in terms of population and ranks the third in size, second only to Russia and Canada. Its geographical and topological characteristics lead to significant variations in climate cross the country, which is a necessary condition for applying this technique. The main challenge of conducting the similar research in China is to obtain household survey data in

conjunction with the information on the geographical areas where households are located in. More importantly, those geographical areas should be so small that within these areas it is safe to assume that all households enjoy the same climate.

In addition, it would also be interesting to employ different techniques to estimate the amenity value of climate to households in India, such as the life satisfaction approach, from which we can see how different approaches to valuing the amenity value of climate might yield different estimates.

References

Adams, B. N. (2010). Themes and threads of family theories: A brief history. Journal of Comparative Family Studies, 41(4), 499–505.

Abdulai, A., Jain, D. K., & Sharma, A. K. (1999). Household food demand analysis in India. Journal of Agricultural Economics, 50(2), 316–327.

Acevedo, S., Mrkaic, M., Novta, N., Pugacheva, E., & Topalova, P. (2020). The effects of weather shocks on economic activity: what are the channels of impact?. Journal of Macroeconomics, 65.

Adusah-Poku, F., & Takeuchi, K. (2019). Household energy expenditure in Ghana: A double-hurdle model approach. World Development, 117, 266–277.

Affuso, E. (2019). Consumer welfare and climate change in Greenland. Energy Economics, 84.

Aftab, S., Yaseen, M. R., & Anwar, S. (2017). Impact of rising food prices on consumer welfare in the most populous countries of South Asia. International Journal of Social Economics, 44(8), 1062–1077.

Albouy, D., Graf, W., Kellogg, R., & Wolff, H. (2016). Climate amenities, climate change, and American quality of life. Journal of the Association of Environmental and Resource Economists, 3(1), 205-246.

Attri, S. D., & Tyagi, A. (2010). Climate profile of India. Environment Monitoring and Research Center, India Meteorology Department: New Delhi, India.

Balli, F., & Tiezzi, S. (2010). Equivalence scales, the cost of children and household consumption patterns in Italy. Review of Economics of the Household, 8(4), 527–549.

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

Barten, A. P. (1964). Consumer demand functions under conditions of almost additive preferences. Econometrica: Journal of the Econometric Society, 1-38.

Basumatary, S., Devi, M., & Basumatary, K. (2021). Determinants of household electricity demand in rural India: A case study of the impacts of government subsidies and surcharges. International Journal of Energy Economics and Policy, 11(6), 243–249.

Becker, G. S. (1965). A Theory of the Allocation of Time. The economic journal, 75(299), 493-517.

Besagni, G., & Borgarello, M. (2018). The determinants of residential energy expenditure in Italy. Energy, 165, 369–386.

Besagni, G., & Borgarello, M. (2020). What are the odds? On the determinants of high residential energy expenditure. Energy and Built Environment, 1(3), 288–295.

Bhatnagar, M., Mathur, J., & Garg, V. (2018). Determining base temperature for heating and cooling degree-days for India. Journal of Building Engineering, 18, 270-280.

Bhattacharjee, S., & Reichard, G. (2011). Socio-economic factors affecting individual household energy consumption: A systematic review. ASME 2011 5th International Conference on Energy Sustainability, ES 2011, PARTS A, B, AND C, 891–901.

Bhattacharyya, S. C. (2015). Influence of India's transformation on residential energy demand. Applied Energy, 143, 228–237.

Biewen, M., & Juhasz, A. (2017). Direct Estimation of Equivalence Scales and More Evidence on Independence of Base. Oxford Bulletin of Economics and Statistics, 79(5), 875–905.

Biswas, S., & Das, U. (2022). Adding fuel to human capital: Exploring the educational effects of cooking fuel choice from rural India. Energy Economics, 105.

Blacklow, P., Nicholas, A., & Ray, R. (2010). Demographic demand systems with application to equivalence scales estimation and inequality analysis: The australian evidence. Australian Economic Papers, 49(3), 161–179.

Blundell, R., & Lewbel, A. (1991). The information content of equivalence scales. Journal of Econometrics, 50(1–2), 49–68.

Blundell, R., Pashardes, P., & Weber, G. (1993). What do we learn about consumer demand patterns from micro data?. The American Economic Review, 570-597.

Borah, M., Keldenich, C., & Knabe, A. (2019). Reference Income Effects in the Determination of Equivalence Scales Using Income Satisfaction Data. Review of Income and Wealth, 65(4), 736–770.

Bosch-Domènech, A. (1991). Economies of scale, location, age, and sex discrimination in household demand. European Economic Review, 35(8), 1589-1595.

Bradford, D., & Hildebrand, G., 1977. Observable public good preferences. Journal of Public Economics 8, 111–131

Brereton, F., Clinch, J. P., & Ferreira, S. (2008). Happiness, geography and the environment. Ecological Economics, 65(2), 386–396.

Brounen, D., Kok, N., & Quigley, J. M. (2012). Residential energy use and conservation: Economics and demographics. European Economic Review, 56(5), 931–945.

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Global non-linear effect of temperature on economic production. Nature, 527(7577), 235-239.

Chakravorty, S., Goli, S. & James, K.S., 2021. Family demography in India: Emerging patterns and its challenges. Sage Open, 11(2).

Chalal, M. L., Benachir, M., White, M., Shahtahmassebi, G., Cumberbatch, M., & Shrahily, R. (2017). The impact of the UK household life-cycle transitions on the electricity and gas usage patterns. In Renewable and Sustainable Energy Reviews (Vol. 80, pp. 505–518). Elsevier Ltd.

Chau, K. W., & Chin, T. L. (2003). A critical review of literature on the hedonic price model. International Journal for Housing Science and its applications, 27(2), 145-165.

Chen, Z. (2006). Measuring the poverty lines for urban households in China-an equivalence scale method. China Economic Review, 17(3), 239–252.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1975). Transcendental logarithmic utility functions. The American Economic Review, 65(3), 367-383.

Chindarkar, N., Jain, A., & Mani, S. (2021). Examining the willingness-to-pay for exclusive use of LPG for cooking among rural households in India. Energy Policy, 150.

Conevska, A., & Urpelainen, J. (2020). Weathering electricity demand? Seasonal variation in electricity consumption among off-grid households in rural India. Energy Research and Social Science, 65.

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, 829-844.

Cragg, M., & Kahn, M. (1997). New estimates of climate demand: evidence from location choice. Journal of Urban Economics, 42(2), 261-284.

Curtis, C.C., Lugauer, S., & Mark, N.C., 2017. Demographics and aggregate household saving in Japan, China, and India. J. Macroecon. 51, 175–191.

D'Agostino, A. L., Urpelainen, J., & Xu, A. (2015). Socio-economic determinants of charcoal expenditures in Tanzania: Evidence from panel data. Energy Economics, 49, 472–481.

Daripa, S., & Dinda, S. (2022). Hurdles in Fuel Choice and Consumption in Rural India. In Persistent and Emerging Challenges to Development (pp. 109-123). Springer, Singapore.

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

Deaton, A., & Muellbauer, J. (1980). An Almost Ideal Demand System. The American Economic Review, 70(3), 312–326.

Deaton, A. S., & Muellbauer, J. (1986). On measuring child costs: With applications to poor countries. Journal of Political Economy, *94*(4), 720-744.

Dell, M., Jones, B. F., & Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. American Economic Journal: Macroeconomics, 4(3), 66-95.

Djanibekov, U., & Gaur, V. (2018). Nexus of energy use, agricultural production, employment and incomes among rural households in Uttar Pradesh, India. Energy Policy, 113, 439–453.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. Journal of Economic Psychology, 29(1), 94–122.

Dudel, C., Garbuszus, J. M., & Schmied, J. (2020). Assessing differences in household needs: a comparison of approaches for the estimation of equivalence scales using German expenditure data. Empirical Economics.

Duffy, M. (2003). Advertising and food, drink and tobacco consumption in the United Kingdom: a dynamic demand system. Agricultural Economics, 28(1), 51-70.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In Nations and households in economic growth (pp. 89-125). Academic press.

Ekholm, T., Krey, V., Pachauri, S., & Riahi, K. (2010). Determinants of household energy consumption in India. Energy Policy, 38(10), 5696–5707.

Filippini, M., & Pachauri, S. (2004). Elasticities of electricity demand in urban Indian households. Energy Policy, 32(3), 429–436.

Ekholm, T., Krey, V., Pachauri, S., & Riahi, K. (2010). Determinants of household energy consumption in India. Energy Policy, 38(10), 5696–5707.

Engel, E. (1895). Die Lebenskosten belgischer Arbeiter-Familien früher und jetzt. Heinrich.

Englin, J. (1996). Estimating the amenity value of rainfall. The Annals of Regional Science, 30, 273-283.

Fan, S., Wailes, E. J., & Cramer, G. L. (1995). Household demand in rural China: A two-stage LES-AIDS model. American Journal of Agricultural Economics, 77(1), 54-62.

Felix, K. T., & Kumar, P. N. (2020). An economic analysis of food demand in rural and urban households of tamil nadu. Indian Journal of Economics and Development, 16(2), 173–179.

Ferreira, S., Akay, A., Brereton, F., Cuñado, J., Martinsson, P., Moro, M., & Ningal, T. F. (2013). Life satisfaction and air quality in Europe. Ecological Economics, 88, 1–10.

Ferreira, S., & Moro, M. (2010). On the use of subjective well-being data for environmental valuation. Environmental and Resource Economics, 46, 249-273.

Filippini, M., & Pachauri, S. (2004). Elasticities of electricity demand in urban Indian households. Energy Policy, 32(3), 429–436.

Frijters, P., & Van Praag, B. M. (1998). The effects of climate on welfare and well-being in Russia. Climatic Change, 39(1), 61-81.

Gabriel, S. A., Mattey, J. P., & Wascher, W. L. (2003). Compensating differentials and evolution in the quality-of-life among U.S. states. Regional Science and Urban Economics, 33(5), 619–649.

Gould, C. F., Hou, X., Richmond, J., Sharma, A., & Urpelainen, J. (2020). Jointly modeling the adoption and use of clean cooking fuels in rural India. Environmental Research Communications, 2(8).

Gould, C. F., & Urpelainen, J. (2018). LPG as a clean cooking fuel: Adoption, use, and impact in rural India. Energy Policy, 122, 395–408.

Gundimeda, H., & Köhlin, G. (2008). Fuel demand elasticities for energy and environmental policies: Indian sample survey evidence. Energy Economics, 30(2), 517–546.

Gupta, S., (2022). Climate change impacts in India: a review. EAERE Magazine.

Gupta, S., & Saudamini D., (2022). Adapting to a changing climate in India: initiatives, challenges, and future directions. EAERE Magazine.

Gyourko, J., & Tracy, J. (1991). The structure of local public finance and the quality of life. Journal of political economy, 99(4), 774-806.

Hertwich, E. G., and Peters, G. P. (2009). Carbon footprint of nations: a global, trade-linked analysis. Environmental science & technology, 43(16), 6414-6420.

Hoch, I., & Drake, J. (1974). Wages, climate, and the quality of life. Journal of Environmental Economics and Management, 1(4), 268-295.

Hoehn, J. P., Berger, M. C., & Blomquist, G. C. (1987). A hedonic model of interregional wages, rents, and amenity values. Journal of Regional Science, 27(4), 605–620.

Humphrey, D. H. & Howar. S. Q. (1976). Expenditure and household - size elasticities in Malawi: Urban - rural comparisons, demand projections and a summary of East African findings. The Journal of Development Studies, 252–269.

Idrees, M., Raza, K., & Aziz, B. (2013). An Econometric Analysis of Electricity Demand for the Residential Sector of Pakistan. Forman Journal of Economic Studies, 09, 1–18.

IEA, I., 2021. India energy outlook 2021. World Energy Outlook Special Report.

IEA (2022), World Energy Outlook 2022, IEA, Paris https://www.iea.org/reports/world-energy-outlook-2022, License: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A)

Ironmonger, D. S., Aitken, C. K., & Erbas, B. (1995). Economies of scale in energy use in adult-only households. Energy Economics, 17(4), 301–310.

Jayasinghe, M. S., Smith, C., Chai, A., & Ratnasiri, S. (2016). The implications of income dependent equivalence scales for measuring poverty in Sri Lanka. International Journal of Social Economics, 43(12), 1300–1314.

Jimenez Mori, R., & Yepez-Garcia, A. (2020). Understanding the Drivers of Household Energy Spending: Micro Evidence for Latin America. Latin American Economic Review, 1–33.

Jones, R. V., Fuertes, A., & Lomas, K. J. (2015). The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. Renewable and Sustainable Energy Reviews, 43, 901–917.

Kahn, M. E. (2009). Urban Growth and Climate Change. Annual Review of Resource Economics, 1(1), 333–350.

Kahn, M. E., Mohaddes, K., Ng, R. N., Pesaran, M. H., Raissi, M., & Yang, J. C. (2021). Long-term macroeconomic effects of climate change: A cross-country analysis. Energy Economics, 104, 105624.

Kakwani, N., & Son, H. H. (2005). Economies of Scale in Household Consumption: With Application To Australia. Australian Economic Papers, 44(2), 134–148.

Karbasi, A. & Mohammadzadeh, S.H., 2018. Estimating Household Expenditure Economies of Scale in Iran.

Karimu, A., Mensah, J. T., & Adu, G. (2016). Who adopts LPG as the main cooking fuel and why? Empirical evidence on Ghana based on national survey. World Development, 85, 43-57.

Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. Energy, 55, 184–194.

Khanal, A. R., Mishra, A. K., & Keithly, W. (2016). Heterogeneity in Food Demand among Rural Indian Households: The Role of Demographics. Canadian Journal of Agricultural Economics, 64(3), 517–544.

Kishore, A., & Spears, D. (2014). Having a son promotes clean cooking fuel use in urban India: Women's status and son preference. In Economic Development and Cultural Change (Vol. 62, Issue 4).

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... & Takahashi, K. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. Journal of the Meteorological Society of Japan. Ser. II, 93(1), 5-48.

Koirala, B. S., & Bohara, A. K. (2014). Valuing US climate amenities for Americans using a hedonic pricing framework. Journal of Environmental Planning and Management, 57(6), 829–847.

Kostakis, I. (2020). Socio-demographic determinants of household electricity consumption: Evidence from Greece using quantile regression analysis. Current Research in Environmental Sustainability, 1, 23-30.

Koulovatianos, C., Schröder, C., & Schmidt, U. (2005). On the income dependence of equivalence scales. Journal of Public Economics, 89(5–6), 967–996.

Lancaster, G., & Ray, R. (1998). Comparison of alternative models of household equivalence scales: the Australian evidence on unit record data. Economic Record, 74(224), 1–14.

Leach, G. (1992). The energy transition. Energy Policy, 20(2), 116–123.

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

Lewbel, A. (1985). A unified approach to incorporating demographic or other effects into demand systems. The Review of Economic Studies, 52(1), 1-18.

Li, Y., Jiang, Y., & Li, S. (2022). Price and income elasticities of electricity in China: Estimation and policy implications. Regional Science Policy & Practice, 14, 76-90.

Liu, N., Li, X., & Yan, L. (2021) The Child Costs in Chinese Family: A Welfare Estimate Using Equivalence Scale. Population and Economics, 1, 50-67. (in Chinese)

Longhi, S. (2015). Residential energy expenditures and the relevance of changes in household circumstances. Energy Economics, 49, 440–450.

Mackerron, G. (2012). Happiness economics from 35000 feet. Journal of Economic Surveys, 26(4), 705–735.

Maddison, D. (2001). The amenity value of the global climate. Earthscan.

Maddison, D. (2003). The amenity value of the climate: The household production function approach. Resource and Energy Economics, 25(2), 155–175.

Maddison, D., & Bigano, A. (2003). The amenity value of the Italian climate. Journal of Environmental Economics and Management, 45(2), 319–332.

Maddison, D., & Rehdanz, K. (2011). The impact of climate on life satisfaction. Ecological Economics, 70(12), 2437–2445.

Maddison, D., Rehdanz, K., & Narita, D. (2013). The household production function approach to valuing climate: The case of Japan. Climatic Change, 116(2), 207–229.

Maddison, D., & Rehdanz, K. (2020). Cross-country variations in subjective wellbeing explained by the climate. Handbook on Wellbeing, Happiness and the Environment. Edward Elgar Publishing Limited, Cheltenham, UK, 105-126.

Maji, P., Mehrabi, Z., & Kandlikar, M. (2021). Incomplete transitions to clean household energy reinforce gender inequality by lowering women's respiratory health and household labour productivity. World Development, 139.

Malakar, Y., & Day, R. (2020). Differences in firewood users' and LPG users' perceived relationships between cooking fuels and women's multidimensional well-being in rural India. Nature Energy, 5(12), 1022–1031.

Mani M., Markandya, A., (2022). Climate change and living standards in India: hotspots and effective adaption. EAERE Magazine.

Matsumoto, S. (2019). Daily Habits and Energy Consumption: Go to Bed Earlier for Environmental Protection. European Journal of Sustainable Development, 8(4), 54.

Meenakshi, J. V., & Ray, R. (1999). Regional differences in India's food expenditure pattern: A complete demand systems approach. Journal of International Development, 11(1), 47–74.

Meier, H., & Rehdanz, K. (2010). Determinants of residential space heating expenditures in Great Britain. Energy Economics, 32(5), 949–959.

Meier, H., & Rehdanz, K. (2017). The amenity value of the British climate. Urban Studies, 54(5), 1235–1262.

Michelini, C. (1999). The estimation of a rank 3 demand system with demographic demand shifters from quasi-unit record data of household consumption. Economics Letters, 65(1), 17-24.

Michelini, C. (2001). Estimating the cost of children from New Zealand quasi-unit record data of household consumption. Economic Record, 77(239), 383–392.

Mok, T.P., Maclean, G. and Dalziel, P., 2011. Household Size Economies: Malaysian Evidence. Economic Analysis & Policy, 41(2).

Moore, T., & Green, C. J. (2008). Flow of funds and the impact of financial controls on bank portfolio behaviour: A study of India. European Journal of Finance, 14(7), 641–661.

Moore, T., Green, C. J., & Murinde, V. (2005). Portfolio Behaviour in a Flow of Funds Model for the Household Sector in India.

Mottaleb, K. A., & Rahut, D. B. (2021). Clean energy choice and use by the urban households in India: Implications for sustainable energy for all. Environmental Challenges, 5.

Mulenga, B. P., Tembo, S. T., & Richardson, R. B. (2019). Electricity access and charcoal consumption among urban households in Zambia. Development Southern Africa, 36(5), 585-599.

Murray, T., Maddison, D., & Rehdanz, K. (2013). Do geographical variations in climate influence life-satisfaction?. Climate Change Economics, 4(01), 1350004.

Navajas, F. H. (2009). Engel curves, household characteristics and low-user tariff schemes in natural gas. Energy Economics, 31(1), 162–168.

Nayak, D. K., & Behera, R. N. (2014). Changing household size in India: An inter-state comparison. Transactions of the Institute of Indian Geographers, 36(1), 1–18.

Nelson, J. A. (1992). Methods of Estimating Household Equivalence Scales: an Empirical Investigation. Review of Income and Wealth, 38(3), 295–310.

Nelson, J. A., & Nelson, J. A. (1993). Household Equivalence Scales : Theory versus Policy? Journal of Labor Economics, 11(3), 471–493.

Nicholson, J. L. (1976). Appraisal of different methods of estimating equivalence scales and their results. Review of Income and Wealth, 22(1), 1-11.

OECD (2005). What are equivalence scales?

O'Neill, Brian C., & Chen, B. S. (2002). Demographic Determinants of Household Energy Use in the United States. Population and Development Review, 28(2002), 53–88.

Pachauri, S. (2004). An analysis of cross-sectional variations in total household energy requirements in India using micro survey data. Energy Policy, 32(15), 1723–1735.

Pachauri, S., & Spreng, D. (2002). Direct and indirect energy requirements of households in India. Energy policy, 30(6), 511-523.

Park, J., Bangalore, M., Hallegatte, S., & Sandhoefner, E. (2018). Households and heat stress: estimating the distributional consequences of climate change. Environment and Development Economics, 23(3), 349-368.

Paun, A., Acton, L., & Chan, W. S. (2018). Fragile planet: Scoring climate risks around the world. HSBC Bank.

Pendakur, K. (2018). Welfare analysis when people are different. 51(May), 321–360.

Phipps, S., & Garner, T. I. (1994). Are Equivalence Scales the Same for the United States and Canada? Review of Income and Wealth, 40(1), 1–17.

Poi, B. P. (2012). Easy demand-system estimation with quaids. The Stata Journal, 12(3), 433-446.

Pollak, R. A., & Wales, T. J. (1981). Demographic variables in demand analysis. Econometrica: Journal of the Econometric Society, 1533-1551.

Rahman, M.M., Jeon, S.H., & Yoon, K.S., 2020. Estimation of equivalence scale and assessment of its impact on poverty measurement in Bangladesh. Sustainability, 12(21), p.9081.

Rahut, D. B., Behera, B., & Ali, A. (2016). Household energy choice and consumption intensity: Empirical evidence from Bhutan. Renewable and Sustainable Energy Reviews, 53, 993–1009.

Rao, M. N., & Reddy, B. S. (2007). Variations in energy use by Indian households: An analysis of micro level data. Energy, 32(2), 143–153.

Ray, R. (1983). Measuring the Cost of Children: An alternative approach. Journal of Population Economics, 22, 89–102.

Rehdanz, K. (2006). Hedonic pricing of climate change impacts to households in Great Britain. Climatic Change, 74(4), 413–434.

Rehdanz, K., & Maddison, D. (2005). Climate and happiness. Ecological Economics, 52(1), 111–125.

Rehdanz, K., & Maddison, D. (2009). The amenity value of climate to households in Germany. Oxford Economic Papers, 61(1), 150-167.

Roback, J. (1982). Wages, rents, and the quality of life. Journal of political Economy, 90(6), 1257-1278.

Rosen S (1979). Wage-based indexes of urban quality of life. Current issues in urban economics, 74-104.

Rothbarth, E. (1943). Note on a Method of Determining Equivalent Income for Families of Different Consumption. War Time Pattern of Saving and Spending, Occasional Paper, (4).

Salari, M., & Javid, R. J. (2017). Modeling household energy expenditure in the United States. In Renewable and Sustainable Energy Reviews (Vol. 69, pp. 822–832). Elsevier Ltd.

Sankhyayan, P., & Dasgupta, S. (2019). 'Availability' and/or 'Affordability':What matters in household energy access in India? Energy Policy, 131, 131–143.

Saxena, V., & Bhattacharya, P. C. (2018). Inequalities in LPG and electricity consumption in India: The role of caste, tribe, and religion. Energy for Sustainable Development, 42, 44–53.

Schröder, C., Rehdanz, K., Narita, D., & Okubo, T. (2015). The decline in average family size and its implications for the average benefits of within-household sharing. Oxford Economic Papers, 67(3), 760–780.

Sharma, S. V., Han, P., & Sharma, V. K. (2019). Socio-economic determinants of energy poverty amongst Indian households: A case study of Mumbai. Energy Policy, 132, 1184–1190.

Singh, J., Mantha, S. S., & Phalle, V. M. (2018). Characterizing domestic electricity consumption in the Indian urban household sector. Energy and Buildings, 170, 74–82.

Sinha, P., Caulkins, M., & Cropper, M. (2021). The value of climate amenities: A comparison of hedonic and discrete choice approaches. Journal of Urban Economics, 126.

Smith, G. (2018). Step away from stepwise. Journal of Big Data, 5(1), 1-12.

Song, N., Aguilar, F. X., Shifley, S. R., & Goerndt, M. E. (2012). Factors affecting wood energy consumption by U.S. households. Energy Economics, 34(2), 389–397.

Srivastava, S. K., Mathur, V. C., Kumar, N., Hasan, R., & Meena, P. C. Unravelling Food Basket of Indian Households: Revisiting Underlying Changes and Future Food.

Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand. The Economic Journal, 64(255), 511-527.

Subramanian, R., Kakkagowder, C., Perumal, A., & Gurusamy, P. (2019). Consumption, Expenditure and Demand Analysis of Milk and Milk Products in India. Indian Journal of Economics and Development, 15(2), 301.

Sussman, F., Saha, B., Bierwagen, B. G., Weaver, C. P., Cooper, W., Morefield, P. E., & Thomas, J. V. (2014). Estimates of changes in county-level housing prices in the United States under scenarios of future climate change. Climate Change Economics, 5(3).

Swain, S. S., & Mishra, P. (2020). Determinants of adoption of cleaner cooking energy: Experience of the Pradhan Mantri Ujjwala Yojana in rural Odisha, India. Journal of Cleaner Production, 248.

Taale, F., & Kyeremeh, C. (2019). Drivers of households' electricity expenditure in Ghana. Energy and Buildings, 205.

Tian, X., Yu, X., & Klasen, S. (2018). Gender discrimination in China revisited: A perspective from family welfare. Journal of Chinese Economic and Business studies, 16(1), 95-115.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. Econometrica: journal of the Econometric Society, 24-36.

Tol, R. S. J. (2009). The economic effects of climate change. Journal of economic perspectives, 23(2), 29-51.

Tripathi, S. (2013). Do large agglomerations lead to economic growth? Evidence from urban India. Review of Urban & Regional Development Studies, 25(3), 176–200.

Tsakloglou, P. (1991). Estimation and comparison of two simple models of equivalence scales for the cost of children. The Economic Journal, 101(405), 343-357.

Ulman, P. (2012). Equivalence Scale in Terms of Polish Households' Source of Income. Folia Oeconomica Stetinensia, 10(2), 114–127.

Umanath, M., Vijayasarathi, K., Babu, B. N., & Baskar, M. (2015). Food consumption pattern and nutrient intake in rural and urban Karnataka. Indian Journal of Agricultural Economics, 70(4), 487-501.

United Nations, Department of Economic and Social Affairs, Population Division. (2019). Patterns and trends in household size and composition: evidence from a United Nations dataset.

Van Praag, B. M. (1968). Individual welfare functions and consumer behavior: A theory of rational irrationality (Vol. 57). North-Holland Publishing Company.

Van Praag, B. M., & Baarsma, B. E. (2005). Using happiness surveys to value intangibles: The case of airport noise. The Economic Journal, 115(500), 224-246.

Vashchilko, A. (2014). Household Expenditure Patterns, Equivalence Scales, and Poverty in Belarus: Formal Versus Informal Sector. Eastern European Economics, 52(6), 92–108.

WEF (World Economic Forum), The global risk report 2022, 17th edition.

Wiesmann, D., Lima Azevedo, I., Ferrão, P., & Fernández, J. E. (2011). Residential electricity consumption in Portugal: Findings from top-down and bottom-up models. Energy Policy, 39(5), 2772–2779.

WHO, 2020. Indoor air pollution and household energy (https://www.who.int/heli/risks/indoorair/indoorair/en/), Geneva, Switzerland.

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

Wu, W., Kanamori, Y., Zhang, R., Zhou, Q., Takahashi, K., & Masui, T. (2021). Implications of declining household economies of scale on electricity consumption and sustainability in China. Ecological Economics, 184, 106981.

Zapata, O. (2022). Happiness in the tropics: climate variables and subjective wellbeing. Environment and development economics, 27(3), 250-271.

Appendices

Appendix 1 The comparison of all commodity categories across 2013-2017 CHFS survey

Commodity	2013	2015	2017
Food	Monthly expenditures on food, including both food-at-home and food-away-from-home. The expenditure on both food in total and food-away from-home are available in the survey	Same	Same
Clothing	The expenditure on clothing over the last year. There are no detailed explanation of clothing in the survey	same	Same
Transport	Monthly expenditures on local transportation fees, where " local" means within the scope of the households' local counties and cities, such as subway/bus/taxi fares, or self-drive oil/parking/repair/maintenance fees. (Note: the transportation costs for the purpose of tourism are excluded)	Same	Same
Communication	The monthly costs of using telephone/cell phone/internet	The costs of internet are included in the category of entertainment, rather than communication	The same as 2013 survey
Education	The total expenditure on education over the last year, including kindergarten, school, interest classes, **fitness**, textbooks, equipment, adult training, etc.	Same as 2013 survey	Fitness cost is excluded from education
Entertainment	The monthly expenditure on entertainment includes newspapers, magazines, CDs, movie and show tickets, night clubs and so on.	The cost of internet is also included in entertainment, rather than communication, like 2013 and 2017 surveys.	Same as 2013 survey
Tourism	The total expenditure on tourism over the last year	Same	Same
Healthcare	The total expenditure on medical and healthcare. the gross medical expenditure and the amount paid by the government (through social insurance) is listed	Same as 2013 survey, except for listing in different part of the survey.	Different for 2013 and 2015 surveys that fitness cost is put together with

	in the insurance part rather than the expenditure part as in 2015 and 2017 surveys		healthcare, rather than in the category of entertainment.
Housing	The monthly expenditure on water/electricity/fuel, property management, and the expenditure on house decoration/ repair/maintenance/expansion over the last year	Same as 2013 survey	Same as 2013 and 2015 surveys, except that expenditure on house expansion is excluded
Daily necessities and services	The monthly expenditures on daily necessities (such as laundry supplies, toiletries, household textiles, etc. and housekeeping services)	Same	Same

Appendix 2 The estimation results from the Engel's Method

1. Whole sample:

Variables	(1) Model 1	(2) Model 2	(3) Model 3
ln (pc_totalexp)	-0.108***	-0.108***	-0.108***
	(0.005)	(0.005)	(0.005)
adult	-0.043***	-0.043***	-0.043***
	(0.002)	(0.002)	(0.002)
kid	-0.042***		
	(0.004)		
kid0-2		-0.039***	
		(0.006)	
kid 3-6		-0.038***	
		(0.004)	
kid7-12		-0.042***	
		(0.004)	
kid13-17		-0.048***	
		(0.004)	
male Kid			-0.041***
			(0.004)
female Kid			-0.044***
			(0.004)
rural	-0.067***	-0.067***	-0.067***
	(0.007)	(0.007)	(0.007)
north	-0.046***	-0.045***	-0.046***
	(0.011)	(0.011)	(0.011)
middle	-0.013*	-0.013*	-0.013*
	(0.007)	(0.007)	(0.007)
east	0.03**	0.03**	0.03**
	(0.012)	(0.012)	(0.012)
northeast	0.005	0.005	0.005
	(0.015)	(0.015)	(0.015)
constant	1.612***	1.613***	1.612***
	(0.049)	(0.49)	(0.049)
Obs.	36,188	36,188	36,188
R-squared	0.211	0.211	0.211

*** p<0.01, ** p<0.05, * p<0.1; standard errors clustered at province level in parentheses

2. Subsample of urban households:

Variables	(1) Model 1	(2) Model 2	(3) Model 3
ln (pc_totalexp)	-0.116***	-0.117***	-0.116***
	(0.006)	(0.006)	(0.006)
adult	-0.045***	-0.046***	-0.045***
	(0.002)	(0.002)	(0.002)
kid	-0.055***		
	(0.005)		
kid0-2		-0.051***	
		(0.006)	
kid 3-6		-0.048***	
		(0.005)	
kid7-12		-0.056***	
		(0.006)	

kid13-17		-0.063***	
		(0.006)	
male Kid			-0.055***
			(0.005)
female Kid			-0.056***
			(0.004)
north	-0.036**	-0.036**	-0.036**
	(0.014)	(0.014)	(0.014)
middle	-0.012	-0.012	-0.012
	(0.009)	(0.009)	(0.009)
east	0.029**	0.028**	0.029**
	(0.014)	(0.014)	(0.014)
northeast	0.004	0.004	0.004
	(0.017)	(0.017)	(0.017)
constant	1.701***	1.703***	1.701***
	(0.062)	(0.061)	(0.062)
Obs.	25,471	25,471	25,471
R-squared	0.237	0.238	0.237

*** p<0.01, ** p<0.05, * p<0.1; standard errors clustered at province level in parentheses

3. Subsample of rural households:

Variables	(1) Model 1	(2) Model 2	(3) Model 3
ln (pc_totalexp)	-0.092***	-0.092***	-0.092***
	(0.003)	(0.003)	(0.003)
adult	-0.041***	-0.042***	-0.041***
	(0.003)	(0.002)	(0.002)
kid	-0.018***		
	(0.003)		
kid0-2		0.0003	
		(0.009)	
kid 3-6		-0.014**	
		(0.006)	
kid7-12		-0.018***	
		(0.003)	
kid13-17		-0.023***	
		(0.005)	
male Kid			-0.015***
			(0.003)
female Kid			-0.022***
			(0.004)
north	-0.068**	-0.069***	-0.069***
	(0.009)	(0.009)	(0.009)
middle	-0.012	-0.012	-0.012
	(0.007)	(0.007)	(0.007)
east	0.037***	0.037**	0.037***
	(0.014)	(0.014)	(0.014)
northeast	0.003	0.003	0.003
	(0.011)	(0.011)	(0.011)
constant	1.388***	1.39***	1.389***
	(0.03)	(0.03)	(0.03)
Obs.	10,717	10,717	10,717
R-squared	0.176	0.177	0.177

*** p<0.01, ** p<0.05, * p<0.1; standard errors clustered at province level in parentheses

Appendix 3 STATA codes for estimating the baseline model in chapter 4

```
set more off

program nlsurmyreg1

version 17.0

syntax varlist(min=23 max=23) if, at(name)

tokenize `varlist'

args share_food share_cloth share_energy  lnp1 lnp2 lnp3 lnp4 lny hh_adult1 hh_kid rural district_climate_pr1 district_climate_HDD1 district_climate_CDD1 district_climate_hum1 central west south east northeast timedum2 timedum3 timedum4

tempname a10 a20 a30 a40

scalar `a10'=`at'[1,1]

scalar `a20'=`at'[1,2]

scalar `a30'=`at'[1,3]

scalar `a40'= 1-`a10'-`a20'-`a30'

tempname a11 a21 a31 a41

scalar `a11'=`at'[1,4]

scalar `a21'=`at'[1,5]

scalar `a31'=`at'[1,6]

scalar `a41'=-`a11'-`a21'-`a31'


tempname a12 a22 a32 a42

scalar `a12'=`at'[1,7]

scalar `a22'=`at'[1,8]

scalar `a32'=`at'[1,9]

scalar `a42'=-`a12'-`a22'-`a32'


tempname a13 a23 a33 a43

scalar `a13'=`at'[1,10]

scalar `a23'=`at'[1,11]

scalar `a33'=`at'[1,12]

scalar `a43'=-`a13'-`a23'-`a33'


tempname a14 a24 a34 a44

scalar `a14'=`at'[1,13]

scalar `a24'=`at'[1,14]
```

```
scalar `a34'=`at'[1,15]
scalar `a44'=-`a14'-`a24'-`a34'

tempname a15 a25 a35 a45
scalar `a15'=`at'[1,16]
scalar `a25'=`at'[1,17]
scalar `a35'=`at'[1,18]
scalar `a45'= 0-`a15'-`a25'-`a35'

tempname a16 a26 a36 a46
scalar `a16'=`at'[1,19]
scalar `a26'=`at'[1,20]
scalar `a36'=`at'[1,21]
scalar `a46'=-`a16'-`a26'-`a36'

tempname a17 a27 a37 a47
scalar `a17'=`at'[1,22]
scalar `a27'=`at'[1,23]
scalar `a37'=`at'[1,24]
scalar `a47'=-`a17'-`a27'-`a37'

tempname a18 a28 a38 a48
scalar `a18'=`at'[1,25]
scalar `a28'=`at'[1,26]
scalar `a38'=`at'[1,27]
scalar `a48'=-`a18'-`a28'-`a38'

tempname b10 b20 b30 b40
scalar `b10'=`at'[1,28]
scalar `b20'=`at'[1,29]
scalar `b30'=`at'[1,30]
scalar `b40'=-`b10'-`b20'-`b30'

tempname b11 b21 b31 b41
scalar `b11'=`at'[1,31]
scalar `b21'=`at'[1,32]
```

```
scalar `b31'=`at'[1,33]
scalar `b41'=-`b11'-`b21'-`b31'

tempname b12 b22 b32 b42
scalar `b12'=`at'[1,34]
scalar `b22'=`at'[1,35]
scalar `b32'=`at'[1,36]
scalar `b42'=-`b12'-`b22'-`b32'

tempname b13 b23 b33 b43
scalar `b13'=`at'[1,37]
scalar `b23'=`at'[1,38]
scalar `b33'=`at'[1,39]
scalar `b43'=-`b13'-`b23'-`b33'

tempname b14 b24 b34 b44
scalar `b14'=`at'[1,40]
scalar `b24'=`at'[1,41]
scalar `b34'=`at'[1,42]
scalar `b44'=-`b14'-`b24'-`b34'

tempname b15 b25 b35 b45
scalar `b15'=`at'[1,43]
scalar `b25'=`at'[1,44]
scalar `b35'=`at'[1,45]
scalar `b45'=-`b15'-`b25'-`b35'

tempname b16 b26 b36 b46
scalar `b16'=`at'[1,46]
scalar `b26'=`at'[1,47]
scalar `b36'=`at'[1,48]
scalar `b46'=-`b16'-`b26'-`b36'

tempname b17 b27 b37 b47
scalar `b17'=`at'[1,49]
scalar `b27'=`at'[1,50]
```

```
scalar `b37'=`at'[1,51]
scalar `b47'=-`b17'-`b27'-`b37'

tempname b18 b28 b38 b48
scalar `b18'=`at'[1,52]
scalar `b28'=`at'[1,53]
scalar `b38'=`at'[1,54]
scalar `b48'=-`b18'-`b28'-`b38'

tempname c11 c12 c13 c14
tempname c21 c22 c23 c24
tempname c31 c32 c33 c34
tempname c41 c42 c43 c44
scalar `c11' = `at'[1,55]
scalar `c12' = `at'[1,56]
scalar `c13' = `at'[1,57]
scalar `c14' = -`c11'-`c12'-`c13'

scalar `c21' = `c12'
scalar `c22' = `at'[1,58]
scalar `c23' = `at'[1,59]
scalar `c24' = -`c21'-`c22'-`c23'

scalar `c31' = `c13'
scalar `c32' = `c23'
scalar `c33' = `at'[1,60]
scalar `c34' = -`c31'-`c32'-`c33'

scalar `c41' = `c14'
scalar `c42' = `c24'
scalar `c43' = `c34'
scalar `c44' = - `c14'-`c24'-`c34'

tempname g1 g2 g3 g4 g5 g6 g7
scalar `g1'=`at'[1,61]
scalar `g2'=`at'[1,62]
```

```
scalar `g3'=`at'[1,63]

scalar `g4'=`at'[1,64]

scalar `g5'=`at'[1,65]

scalar `g6'=`at'[1,66]

scalar `g7'=`at'[1,67]


tempvar lnscaling

gen `lnscaling' = ln(1 + `g1'*`hh_adult1' + `g2'*`hh_kid' + `g3'*`rural' + `g4'*`district_climate_pr1' +`g5'*`district_climate_HDD1' +`g6'*`district_climate_CDD1' + `g7'*`district_climate_hum1')


tempvar a1

gen `a1' =`a11'*`central' + `a12'*`west' + `a13'*`south' + `a14'*`east' +`a15'*`northeast' +`a16'*`timedum2' + `a17'*`timedum3' + `a18'*`timedum4'


tempvar a2

gen `a2' =`a21'*`central' + `a22'*`west' + `a23'*`south' + `a24'*`east' +`a25'*`northeast' +`a26'*`timedum2' + `a27'*`timedum3' + `a28'*`timedum4'


tempvar a3

gen `a3' = `a31'*`central' + `a32'*`west' + `a33'*`south' + `a34'*`east' +`a35'*`northeast' +`a36'*`timedum2' + `a37'*`timedum3' + `a38'*`timedum4'


tempvar a4

gen `a4'= `a41'*`central' + `a42'*`west' + `a43'*`south' + `a44'*`east' +`a45'*`northeast' +`a46'*`timedum2' + `a47'*`timedum3' + `a48'*`timedum4'


tempvar b1

gen `b1' =  `b11'*`central' + `b12'*`west' + `b13'*`south' + `b14'*`east' +`b15'*`northeast' +`b16'*`timedum2' + `b17'*`timedum3' + `b18'*`timedum4'


tempvar b2

gen `b2' =  `b21'*`central' + `b22'*`west' + `b23'*`south' + `b24'*`east' +`b25'*`northeast' +`b26'*`timedum2' + `b27'*`timedum3' + `b28'*`timedum4'


tempvar b3

gen `b3' =  `b31'*`central' + `b32'*`west' + `b33'*`south' + `b34'*`east' +`b35'*`northeast' +`b36'*`timedum2' + `b37'*`timedum3' + `b38'*`timedum4'


quietly {
```

```
        tempvar lnpindex
    gen double `lnpindex' = 3 + (`a10' +`a1')*`lnp1' +(`a20'+`a2') *`lnp2' + (`a30' + `a3')*`lnp3' + (`a40'+`a4')*`lnp4'
    forvalues i = 1/4 {
    forvalues j = 1/4 {
    replace `lnpindex' = `lnpindex' + 0.5*`c`i'`j''*`lnp`i''*`lnp`j''
      }
    }

    replace `share_food'=`a10' +`a1'+`c11'*`lnp1'+`c12'*`lnp2'+`c13'*`lnp3'+`c14'*`lnp4'+ (`b10' +`b1')*(`lny'-`lnscaling'-`lnpindex')
    replace `share_cloth'=`a20' +`a2'+ `c21'*`lnp1'+`c22'*`lnp2'+`c23'*`lnp3'+`c24'*`lnp4'+ (`b20' +`b2')*(`lny'-`lnscaling'-`lnpindex')
    replace `share_energy'=`a30' +`a3'+ `c31'*`lnp1'+`c32'*`lnp2'+`c33'*`lnp3'+`c34'*`lnp4'+ (`b30' +`b3')*(`lny'-`lnscaling'-`lnpindex')
}
end

nlsur myreg1 @ share_food share_cloth share_energy  lnp1 lnp2 lnp3 lnp4 lny hh_adult1 hh_kid rural district_climate_pr1 district_climate_HDD1 district_climate_CDD1 district_climate_hum1 central west south east northeast timedum2 timedum3 timedum4, ifgnls vce(cluster District_code) parameters(a10 a20 a30  a11 a21 a31 a12 a22 a32 a13 a23 a33 a14 a24 a34 a15 a25 a35 a16 a26 a36 a17 a27 a37 a18 a28 a38 c11 c12 c13 c22 c23 c33 b10 b20 b30 b11 b21 b31 b12 b22 b32 b13 b23 b33 b14 b24 b34 b15 b25 b35 b16 b26 b36 b17 b27 b37 b18 b28 b38 g1 g2 g3 g4 g5 g6 g7)  nequations(3)
```