HAL open science

Déterminants de la demande d'électricité des ménages au Vietnam entre 2012 et 2016

Hoai-Son Nguyen

To cite this version:

Hoai-Son Nguyen. Déterminants de la demande d'électricité des ménages au Vietnam entre 2012 et 2016. Economics and Finance. Université Paris Saclay (COmUE), 2019. English. NNT : 2019SACLA013 . tel-02294630

HAL Id: tel-02294630

https://theses.hal.science/tel-02294630v1

Submitted on 23 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université PARIS-SACLAY

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

NNT : 2019SACLA013

Exploring the determinants of household electricity demand in Vietnam in the period 2012 – 16

Thèse de doctorat de l'Université Paris-Saclay
préparée à AgroParisTech (l'Institut des sciences et industries du vivant et de l'environnement)

École doctorale n°581 ABIES
Spécialité de doctorat: sciences économiques

Thèse présentée et soutenue à Paris, le 24 Juin 2019, par

Hoai-Son NGUYEN

Composition du Jury :

M. Phu NGUYEN–VAN Directeur de Recherche, (BETA & Université de Strasbourg)	Président
M. Michel SIMIONI Directeur de Recherche, (INRA)	Rapporteur
M. Joachim SCHLEICH Professeur (Grenoble École de Management)	Rapporteur
M. Phu LE–VIET Lecturer (Fullbright University, Vietnam)	Examinateur
Mme Carine BARBIER Ingénieure de Recherche (CIRED)	Examinateur
M. Minh HA–DUONG Directeur de Recherche (CIRED)	Directeur de thèse

Thèse de doctorat

ACKNOWLEDGEMENT

First of all, I would like to convey my grateful attitude and great appreciation to my supervisor, Dr. Minh HA-DUONG for granting me the chance to conduct this thesis. He not only offers financial funding for my thesis but also shows his profound belief in my abilities and encourages me to explore my best capacity. He consistently allowed the thesis to be my own work but steered me in the right the direction whenever he thought I needed it. It was very especial that he also provided me with the best place for thesis writing that I can imagine, as major part of my thesis was written in his house in Bagneux, next to a wood-burning stove.

I am extremely grateful to my jury and thesis committee, Dr. Phu NGUYEN-VAN, Dr. Michel SIMIONI, Dr. Joachim SCHLEICH, Mme. Carine BARBIER, Dr. Phu LE-VIET and Dr. Gilles CRAGUE. I am deeply indebted to them for their insightful comments and tough questions, which realized my thesis. I would particularly like to thank Mrs. Elisabeth MALTESE, my French teacher, for her patience and kindness, so that I was able to complete my French course as a required condition for my thesis submission.

I would like to sincerely thank Mrs. Marguerite CZARNECKI for translating the thesis summary to French and Mrs. Pippa CARRON for copyediting the first draft of the thesis. They both performed their works at the light speed to help me submit my thesis in time. I also gratefully acknowledge the assistance of the lecturers, colleagues and staffs at CIRED and ABIES.

My heartfelt appreciation goes to my best friend Hoang Anh and his family who always there encouraging me and make Paris my second home whenever I am in Paris.

The thesis would not have been possible without the great support and motivation I get from my wife, Nha Trang and my little two angels, Lam Bach and Khai An. And my special thanks go to my parents-in-law for helping us with our boys when I was away from Viet Nam for the thesis.

Last but not least, this thesis is dedicated to my beloved father Mr. Ban NGUYEN-VAN and mother Mrs. Lan LE-NGOC for the great and small things they have done for me.

Contents

Chapter 1. Introduction 6
Chapter 2. Background information and policy context 8
2.1 Overview of the residential electricity market 8
2.2 Residential electricity prices 8
2.3 Subsidy in residential electricity 9
2.3.1 Subsidy in electricity prices 9
2.3.2 Subsidy in cash transfer for poor households 10
Chapter 3. Model specification 11
3.1 Short-run versus long-run demand functions 11
3.1.1 Theory models of electricity demand function 11
3.1.2 Empirical models for electricity demand estimation 14
3.2 Model specification 20
Chapter 4. Data 22
4.1 Temperature data 22
4.2 Vietnam Household Living Standard Survey data 23
4.2.1 Overview of VHLSS 23
4.2.2 VHLSS samples 24
4.2.3 Sample rotation, survey time and weight 26
4.2.4 Data process from VHLSS 26
4.3 Merging electricity price with VHLSS 27
4.4 Merging temperature data with VHLSS 28
4.5 Deflating monetary variables to a base year and cleaning data 29
4.6 Data sets for short-run and long-run demand function 29
4.7 Descriptive statistics 30
Chapter 5. Price elasticity of residential electricity demand 33
5.1 Introduction 33
5.2 Literature review 33
5.2.1 Controversy in the type of prices 33
5.2.2 Models for testing price types 36
5.2.3 Endogeneity in block pricing 38
5.3 Model specification and data 39
5.3.1 Short-run model 39
5.3.2 Long-run model 42
5.3.3 Perceived price model 42
5.3.4 Econometric techniques 42
5.3.5 Data 44
5.4 Results and discussion 44
5.4.1 Short-run demand 44
5.4.2 Long-run demand 54
5.4.3 Perceived price model 58
5.5 Conclusion 59
Chapter 6. Income and electricity poverty in Vietnam 2012–16 61
6.1 Introduction 61
6.2 Literature review 61
6.2.1 Direct measurement approach 62
6.2.2 Indirect measurement approach 62
6.3 Methods and data 64
6.4 Results 66
6.4.1 Visualization approach 66

6.4.2 Econometric approach 67
6.5 Conclusion 69
Chapter 7. Economies of scale in residential electricity consumption 71
7.1 Introduction 71
7.2 Literature review 71
7.2.1 Economies of scale 71
7.2.2 Economies of scale for household electricity expenditure 72
7.3 Methods and data 74
7.3.1 Non-parametric method 74
7.3.2 Parametric method 76
7.3.3 Data 78
7.4 Results and discussion 79
7.4.1 Non-parametric analysis 79
7.4.2 Parametric analysis 81
7.5 Conclusion 83
Chapter 8. Heatwaves and residential electricity demand 85
8.1 Introduction 85
8.2 Literature review 85
8.3 Methods and data 87
8.3.1 Methods 87
8.3.2 Data 88
8.4 Results and discussion 89
8.5 Conclusion 91
Chapter 9. Conclusion 92

List of tables

Table 2-1. The three most recent retail electricity prices for residential 9
Table 2-2. Evolution of electricity subsidy in cash transfer 10
Table 3-1. Electric appliances in different research papers 17
Table 4-1. List of weather stations in GHCN data 22
Table 4-2. Available data on each round of VHLSS 2008–16 24
Table 4-3. Sample sizes of VHLSS from 2008–2014 25
Table 4-4. Number of wards/EA in each round of VHLSS 25
Table 4-5. Number of households selected in each EA 25
Table 4-6. Survey plan for VHLSS 2008–2014 26
Table 4-7. Paired t-test for mean comparison between original and calculated kWh 27
Table 4-8. Similarity between data sets for short-run and long-run function 31
Table 4-9. Differences between data sets for short-run and long-run functions 32
Table 5-1. Income per capita at 2012 price over dwelling type 40
Table 5-2. Pairwise comparison of income per capita at 2012 price over dwelling type 40
Table 5-3. Estimated results of the short-run demand function 46
Table 5-4. Endogeneity tests for the short-run demand function 46
Table 5-5. Weak instrument tests for the short-run demand function 47
Table 5-6. Pairwise correlation between prices and price IVs 48
Table 5-7. Weak instrument tests from first stage regression for the short-run function 48
Table 5-8. Sensitivity analysis for different sub-samples 49
Table 5-9. Estimated long-run function of MP model 55
Table 5-10. Estimated long-run function of MP model 56
Table 5-11. Robustness tests for the long-run model with different IVs and estimators 56
Table 5-12. Endogeneity tests for long-run models 57
Table 5-13. Weak instrument tests for long-run models 57
Table 5-14. Short-run estimates with the panel as a pooled data 58
Table 5-15. Estimates for perceived price model 59
Table 6-1. Estimates of kWh per cap on the quantiles of income per cap 68
Table 6-2. Average kWh consumption of households having corresponding income quantiles 69
Table 6-3. Electricity services a household can consume with 50 kWh per month 69
Table 7-1. Correlation between income and demographic variables. 77
Table 7-2. Household types and household groups for non-parametric analysis 78
Table 7-3. The average electricity share weighted by the density of income per capita 80
Table 7-4. Estimation of parametric approach for economies of scale in electricity expense . 82
Table 7-5. Sensitivity analysis for economies of scale in electricity expense 82
Table 8-1. Descriptive of heatwave dummy variable 89
Table 8-2. Estimates of impacts of heatwaves on electricity demand 90

List of figures

Figure 2-1. Increase in electricity demand, 2006–2015 (Unit. KTOE) 8
Figure 4-1. Position of the 14 weather stations 23
Figure 4-2. Rotation for VHLSS sampling 26
Figure 4-3. The construction of data sets for short-run and long-run functions 30
Figure 5-1. Changes of marginal price and demand 34
Figure 5-2. Changes of intra-marginal price and demand 35
Figure 5-3. Scatter density plot with fitted line for prices and price IVs 47
Figure 6-1. The S-shaped relationship between household income and kWh consumption 64
Figure 6-2. Density scatter plots with a series of fitted lines for different income quantiles 67
Figure 7-1. Channels of economies of scale in household electricity expenditure 74
Figure 7-2. Illustration of w_m and $f(z_m)$ over regression grid 76
Figure 7-3. Non-parametric analysis for economies of scale 79
Figure 8-1. Histogram of the difference between desired Tavg and actual Tavg 89

Abbreviations list

AC	Air conditioning
AP	average prices
CDD	cooling degree days
CPI	consumer price index
CSCTWA	cross-sectionally correlated and time-wise autoregressive model
CSHTWA	cross-sectionally heteroscedastic and time-wise autoregressive model
DBT	decreasing block tariff
EA	enumeration area
ECM	error correction model
EOS	economies of scale
EVN	Vietnam Electricity
GADM	Global Administrative Areas
GHCN	Global Historical Climatology Network
GMM	generalized method of moments
HDD	heating degree days
IBT	increasing block tariffs
IVs	instrument variables
kWh	kiloWatt hour
LB	lower bound
LkWh	lower bound kWh
LPG	liquefied petroleum gas
LSDV	least square dummy variables
MoF	Ministry of Finance
MoIT	Ministry of Industry and Trade
MP	marginal prices
MPHS	Multi-Purpose Household Survey
NOAA	National Oceanic and Atmospheric Administration (US)
NOOA	National Centers for Environmental Information
OLS	ordinary least squares
PPS	probability proportionate to size (rule of)
PSU	Primary sample unit
RSP	rate structure premium
SSU	Secondary sample unit
TSU	Tertiary sample unit
TVs	televisions
UB	upper bound
UkWh	upper bound kWh
UHI	urban heat island
VAR	vector autoregressive
VHLSS	Vietnam Household Living Standard Surveys
VLSS	Vietnam Living Standard Survey
VND	Vietnamdong

Chapter 1. Introduction

In recent years, demand-side management in residential electricity markets has been a major tool for developing countries in harmonizing economic growth, energy security, and reduction of CO_2 emissions.

First, demand-side management tools, such as increasing block-tariff schedules have potential in encouraging people to save electricity, which in turn reduces carbon emissions from electricity generation associated with fossil fuel consumption. This is important for Asian countries where a large proportion of electricity supply is still based on coal. Developing countries, particularly in Southeast Asia, use coal to ensure continuity of supply.

Second, demand-side management also helps developing countries ensure energy security by constraining the surging demand in electricity. Developing countries face higher tension in electricity markets than developed countries. In developed countries, the market is well established, and demand is relatively stable. Supply sources in those countries can gradually transition to a structure with a higher proportion of renewable sources while providing for economic growth. In developing countries, the market is growing fast, and demand soars due to economic growth and rapid population increase. The fast-rising demand surpasses electricity supply capacity causing power outages. Demand-side management can constrain the fast-rising demand to be in line with current supply capacity.

Demand-side management implementation in residential electricity markets requires a deep understanding of customer behaviors, and household demand. In the past, much has been done with aggregate data to explore factors impacting on residential electricity demand (Houthakker, Verleger and Sheehan, 1974; Hsing, 1994; Holtedahl and Joutz, 2004; Alberini and Filippini, 2011). However, recently two points have emerged that set new challenges in estimating electricity demand. First, there is a movement from aggregate data to micro data at a household level, but the micro data is often either missing data in price (Branch, 1993; Alberini, Gans and Velez-Lopez, 2011) or is narrowed to a regional level rather than a national level due to the absence of national data on tariff structures (Reiss and White, 2005; Zhou and Teng, 2013). Second, climate change has recently introduced a new factor of heatwaves which has not been carefully investigated in electricity demand in the past.

Therefore, this thesis revisits the story of electricity demand estimation within the context of Vietnam over the period 2012–16. Four reasons justify this context. First, Vietnam is a tropical country with frequent summer heatwaves so is ideal for investigation of the impact of heatwaves on electricity demand. Second, the micro survey in Vietnam is a rotated survey which allows the construction of a panel data set from three rounds in three different years, as well as the construction of a pooled data set from the rest. The separation of data into panel and pooled data is ideal to estimate electricity demand both in short-run and long-run. Third, the residential electricity market in Vietnam is a monopoly with a single seller, Vietnam Electricity (EVN). Electricity tariff schedules are proposed by EVN and set by the government and are thus uniform in national scale. This provides a chance to estimate demand function from national micro survey data with full detail of electricity prices.

Finally, Vietnam is a country carrying full features of a developing country, with increasing role of demand-side management. Due to the rapid pace of economic growth, electricity demand in Vietnam has surged, causing challenges in ensuring energy security as well as in developing renewable energy sources. During the period from 2006 to 2015, national electricity consumption grew at an average rate higher than 10 per cent (MOIT and DEA, 2017). Demand for electricity by 2035 is predicted to grow at an average annual rate of eight per cent (MOIT

and DEA, 2017). Almost half of the new capacity is proposed to be coal fired (MOIT and DEA, 2017). In response to this issue, the Vietnamese government considers energy efficiency as a "first fuel". Electricity saving is estimated potentially at 17 per cent by 2030 (MOIT and DEA, 2017). In practice, the government has implemented measures such as increasing the block tariff schedule for residential consumption to encourage people to save electricity (EVN, 2015). In that context, the results of this thesis not only broaden our understanding of residential electricity demand functions in developing countries, but also provides a reference for policy makers in designing measures to manage demand side in Vietnam.

This thesis aims to explore the factors which impact on Vietnamese residential electricity demand. The exploration focuses on four main factors: increasing block tariffs, income, demographics (including household size and composition), and heatwaves. The approach is to investigate the role of these factors via estimating a common form of demand function, with each factor investigated in detail in separate chapters.

The master data for this thesis are constructed from Vietnam Household Living Standard Surveys (VHLSS) 2012, 2014 and 2016, various legal documents on electricity prices, and temperature data from the National Centers for Environmental Information (NOOA) over the corresponding period. The period of 2012–16 was chosen because (i) 2016 is the most updated data we have so far and (ii) the rotated features of VHLSS allows the construction of panel data with a maximum of three rounds. The master data are then separated into two sub-datasets: (i) panel data, including households that appear in all three years; and (ii) pooled data, including households that appear only once in all three years. The two sub-datasets are employed to estimate short-run and long-run demand functions.

The thesis is structured as follows. Chapter 2 provides a policy context and some background information on the Vietnam electricity market. Chapter 3 provides a general literature review which details the theory models and empirical strategy for short-run and long-run functions of electricity demand. Chapter 4 provides detail about the procedure of constructing data sets.

Chapter 5 focuses on increasing block tariffs and examines the impact of increasing the block tariff schedule on residential demand. The two aims of the chapter are: (i) to estimate the price elasticity of demand in short-run and long-run; and (ii) to identify whether households respond to marginal prices or average prices.

Chapter 6 focuses on income. It investigates the non-linear relationship between income and electricity demand and its implication for identifying the electricity poverty threshold. The hypothesis is the existence of an income threshold whereby electricity consumption starts to increase with an increase in income. The consumed kWh per capita of households at that income threshold is the electricity poverty threshold.

Chapter 7 focuses on the role of two demographic factors: household size and household composition. The chapter answers two questions: (i) whether the increasing block tariff schedule cancels out the economies of scale in electricity use in Vietnam; and (ii) whether there is a difference in electricity demand across a child, an adult and an elder in Vietnam.

Chapter 8 focuses on the impacts of heatwaves on household electricity demand. The chapter aims to demonstrate that cooling degree days (CDD) – a popular way to represent temperature in electricity demand function – is insufficient to capture the full impact of temperature since it neglects the extreme distribution of temperature which are heatwaves in tropical countries.

Chapter 2. Background information and policy context

2.1 Overview of the residential electricity market

❖ *The surging demand*

Due to the rapid pace of economic growth, electricity demand in Vietnam has surged in recent years. During the period from 2006 to 2015, national electricity consumption grew at an average rate higher than 10 per cent (MOIT and DEA, 2017). The demand for electricity by 2035 is predicted to grow at an average annual rate of eight percent (MOIT and DEA, 2017).

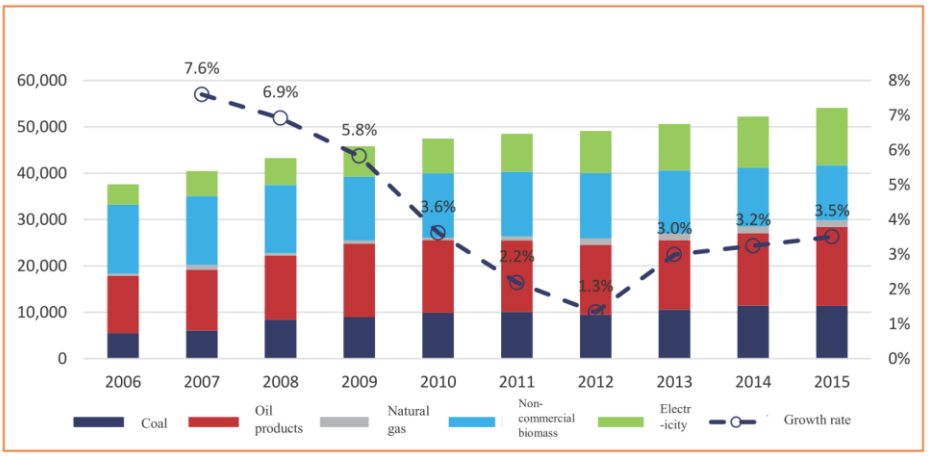


Figure 2-1. Increase in electricity demand, 2006–2015 (Unit. KTOE)
Source. MOIT and DEA (2017, p. 16).

Residential consumption plays a vital role in the sharp increase in electricity demand. First, residential demand accounts for a large proportion of total electricity demand. For example, the total electricity consumption of households accounted for 54 per cent of total electricity consumption in Ha Noi in 2018 (Tiến Hiệp, 2018). Second, the rate of access to electricity of households is 98 per cent (authors calculated from VHLSS 2012–16). With economic growth, households have increased wealth and more electrical appliances resulting in a higher demand for electricity.

❖ *Monopoly in residential electricity market*

Since 2004, several legal documents have been aimed at removing the monopoly in electricity markets, including the residential electricity market. The electricity law (2004) and Decision 26 (2006) regulated that: (i) government is the monopoly in transmission; (ii) generation would be competitive by 2014; (iii) wholesale distribution would be competitive by 2022; and (iv) retail distribution would be competitive from 2022. In 2013, Decision 63 (2013) adjusted the plan to have pilot competitive wholesale in 2015 and officially competitive wholesale in 2021.

In reality, the implementation of the roadmap lags behind the plan. The generation market started to operate in July 2012 with a short suspension in 2017. The competitive wholesale distribution market started to operate in January 2019. Basically, the residential electricity market up to 2019 was a monopoly with Vietnam Electricity (EVN) as a single seller.

2.2 Residential electricity prices

Residential electricity prices in Vietnam are set by the government under Decision 69 of 2013 (Nguyễn Tấn Dũng, 2013). According to the decision, the price schedule is set based on an average electricity selling price. The average electricity selling price is calculated from the production costs and a reasonable level of profit at four stages, including generation, transmission, distribution and supporting services. When the basic input parameters change,

EVN recalculates the average electricity selling price and submits it to the government for approval. The basic input parameters are factors that have a direct impact on the cost of generating electricity that are out of the control of the generating units, including fuel prices, foreign exchange rates, and structure of the actual electricity generation, as well as prices in competitive electricity generation markets.

Residential electricity prices in Vietnam have been in the form of increasing block tariffs (IBTs) since 1994. The IBTs mean that the higher kWh a household consumes, the higher the price per kWh. Households pay a low price for the first block, then pay a higher price for a second block, and so on. According to EVN (2015), IBTs aim to encourage a savings attitude and ensure a low price for low-income households.

Block	*Lower bound*	*Upper bound*	*Apr 2015 – Nov 2017*	*Dec 2017 – Feb 2019*	*Mar 2019 – present*
1	1	50	1,484	1,549	1,678
2	51	100	1,533	1,600	1,734
3	101	200	1,786	1,858	2,014
4	201	300	2,242	2,340	2,536
5	301	400	2,503	2,615	2,834
6	401		2,587	2,701	2,927
ASP			1,622.01	1,720.65	1,864.44

Note. Unit '000 Vietnamdong per kWh; ASP: average selling price; VAT excluded.
Unit for lower bound and upper bound is kWh per month.

Table 2-1. The three most recent retail electricity prices for residential

Source. Author compiled from various legal documents.

There are two price schedules of prices for residential electricity in Vietnam. Both schedules are in IBTs form. The first is the retail price schedule which applies to households that can buy electricity directly from EVN. The second is for wholesale prices. Wholesale prices are applied to rural areas which are remote, have low population density and unorganized infrastructure. In these areas, EVN sells electricity to rural electricity distribution organizations with the wholesale price schedule. These organizations sell electricity to households with their own price schedule based on the wholesale prices. There is heterogeneity in price setting of these organizations. Some apply a single price, while others may apply a three-block design, and so on.

However, the wholesale price schedule is applied only to a small fraction of households. In 2014, EVN provided electricity directly to 84.57 per cent of communes, and 82.59 per cent of rural households (Thục Quyên, 2014). In 2015, EVN provided electricity directly to 87.88 per cent of rural households in Southern provinces (Mai Phương, 2015).

2.3 Subsidy in residential electricity

2.3.1 Subsidy in electricity prices

Poor or low-income households which (i) have monthly electricity consumption less than 50 kWh and (ii) register with EVN can have a special price for the first 50 kWh. The prices for the special block are about 80 per cent of the approved average selling price in 2011, then gradually decreased to 65 per cent in the period from August 2013 to May 2014. Meanwhile, the prices for the normal first block (0–100 kWh) is normally equivalent to 95 per cent of average selling prices. However, since June 2014, the subsidy block has been canceled.

For other households, during the period from March 2011 to May 2014, the first block was from 0 to 100 kWh. The price of the first block is about 95% of average selling price. However, since June 2014, the first block has been divided into 2 blocks. Block 1 is from 0 to 50 kWh. Block 2 is from 51 to 100 kWh. The price of Block 2 is three percentage points higher than the price of Block 1 in term of percentage to average selling price.

2.3.2 Subsidy in cash transfer for poor households

Since 2011, in parallel to the subsidy in electricity prices, the Vietnamese government has also implemented an electricity subsidy program of cash transfers for poor households. According to the program, every household under the national poverty line can receive 30,000 VND (about USD 1.5) each month (Nguyễn Tấn Dũng, 2011). The subsidy is in cash every quarter (Nguyễn Công Nghiệp, 2012).

Periods	*Subsidy*		*Beneficial*
	kWh	Cash	
Dec 2011 to May 2014		30,000	- Poor households • Rural: 400,000 VND per cap per month • Urban: 500,000 VND per cap per month
Jun 2014 to May 2015	30	46,000	- Poor households: national poverty line, if provinces have their own line higher than national line, then apply the province line. • Rural: 700,000 VND per cap per month • Urban: 900,000 VND per cap per month - Households under preferential treatment policy
Jun 2015 to Nov 2017	30	49,000	
Dec 2017 to present	30	51,000	

Note. Unit: Vietnamdong.

Table 2-2. Evolution of electricity subsidy in cash transfer

Source. Author compiled from various legal documents.

In 2014, the subsidy amount was set to 30 kWh at prevailing prices (Đỗ Hoàng Anh Tuấn, 2015). In 2014 prices, the 30 kWh costs 46,000 VND (about USD 2.0). In addition, the benefit is extended to include households under the preferential treatment policy (Nguyễn Tấn Dũng, 2014b). The extended beneficiaries are: (i) non-poor households with members receiving monthly social allowance and consume less than 50 kWh per month from the national grid for living purposes; (ii) households with members receiving the monthly social allowance living in a non-grid area; and (iii) ethnic minority households living in a non-grid area.

In June 2015, due to an increase in electricity price, the subsidy increased to 49,000 VND (Nguyễn Tấn Dũng, 2015). At the new price in December 2017, the subsidy increased to 51,000 VND. In July 2018, the Ministry of Finance (MoF) submitted a proposal cancelling the cash transfer program (H.Anh, 2018). The proposal faced a backlash and so in October 2018 the MoF submitted a draft circular that kept the cash transfer subsidy program at the threshold of 30 kWh (Bộ Tài Chính, 2018).

Chapter 3. Model specification

3.1 Short-run versus long-run demand functions

This study focuses on the residential electricity demand function. Unlike other commodities consumed by households, electricity by itself does not generate utility for consumers. The use of electricity needs to go with appliances such as fans, air conditioners (AC) and televisions (TV). Thus electricity demand can be considered as a derived demand which depends on demand of appliances in households (Taylor, 1975).

Since electricity demand depends on the availability of appliances, there are two types of demand function: short-run and long-run. A short-run demand function is defined by a condition that the appliance stock is fixed (Taylor, 1975, p. 80), while in the long-run function the appliance stock can vary (Taylor, 1975, p. 80).

3.1.1 Theory models of electricity demand function

Fisher and Kaysen (1962) was the first study to explicitly distinguish between short-run and long-run electricity demand functions. Houthakker and Taylor (1970), based on the idea of capital stock in Fisher and Kaysen, derived a complete model for residential electricity demand in both short-run and long-run. So far, most models in electricity demand at the household level are constructed based on the ideas of these two models. The following parts provide a brief explanation of the two models.

3.1.1.1 Fisher and Kaysen (1962)

Fisher and Kaysen (1962) is the pioneer research which distinguished between short-run and long-run demand functions. Fisher and Kaysen (1962) assert that short-run demand is the utilization of appliance which they describe using the term “white goods”.

❖ *Short-run function*

The total electricity consumption of a household (D) during period *t* is the additive function of electricity consumption of *n* appliances. The electricity consumption of an appliance, in turn, is the product of its intensity use (K) with the average stock of the appliance (W). The average stock is measured by the kWh that the appliance consumes in an hour of normal use.

$$D_t = \sum_{i=1}^{n} K_{it} W_{it} \qquad (3\text{-}1)$$

The intensity of use of the ith appliance is a function of electricity price (P) and income (Y)

$$K_{it} = B_i P_t^{\alpha_i} Y_t^{\beta_i} \qquad (3\text{-}2)$$

Substitute (3-2) to (3-1), we have

$$D_t = \sum_{i=1}^{n} B_i P_t^{\alpha_i} Y_t^{\beta_i} W_{it} \qquad (3\text{-}3)$$

They then define

$$C_i = B_i \bar{P}_t^{\alpha_i} \bar{Y}_t^{\beta_i} \qquad (3\text{-}4)$$

$\bar{P}$ and $\bar{Y}$ are the mean of P_t and Y_t over T times period. So the (3-4) becomes

$$D_t = \sum_{i=1}^{n} C_i (P_t/\bar{P})^{\alpha_i} (Y_t/\bar{Y})^{\beta_i} W_{it} \qquad (3\text{-}5)$$

Fisher and Kaysen assume that (3-5) can be approximated by (3-6) with C, α, β being constant.

$$D_t = C(P_t/\bar{P})^{\alpha}(Y_t/\bar{Y})^{\beta} \sum_{i=1}^{n} W_{it} \quad (3\text{-}6)$$

Taking logarithms of both sides of (3-6)

$$lnD_t = C_0 + \alpha lnP_t + \beta lnY_t + lnW_t^* \quad (3\text{-}7)$$

Where

$$C_0 = lnC - \alpha ln\bar{P} - \beta ln\bar{Y} \quad (3\text{-}8)$$

$$W_t^* = \sum_{i=1}^{n} W_{it} \quad (3\text{-}9)$$

W_t^* is the stock of all appliances.

Now the short-run demand function would be

$$lnD_t - lnW_t^* = C_0 + \alpha lnP_t + \beta lnY_t \quad (3\text{-}10)$$

Where α and β are the price and income elasticities at the period *t*.

❖ *Long-run function*

Once white goods are measured by the amount electricity they consume, the long-run demand of electricity is equivalent to the demand for white goods at each household. Thus Fisher and Kaysen (1962) construct a model to explain the changes of household demand for white goods. The key assumption of the model is that the changes are not proportional to the difference between the actual and desired stock of appliances. Fisher and Kaysen argue that for heavy electricity load appliances, each household normally own just one unit. The change of appliance stock at aggregate level is due to the new consumption of households who have not owned these appliances before. Thus the proportion of actual stock is meaningless since actual current stock is zero. In addition, the proportion at the aggregate level includes irrelevant households who have already owned appliances before.

Thus Fisher and Kaysen (1962) propose a model in which the change of white good stocks from time *t-1* to time *t* depends on the changes of permanent income, current income, price of electricity and gas, price of both electricity appliances and gas-using competitors, and other demographic variables. The model is as follows:

$$\Delta lnW_{it} = A_i + \gamma_{i1}\Delta lnY_t^E + \gamma_{i2}lnY_t + \gamma_{i3}E_{it} + (\gamma_{i4}lnG_{it}) + \gamma_{i5}\Delta lnH_t + \gamma_{i6}\Delta lnF_t + \gamma_{i7}lnM_t + \gamma_{i8}lnP_t^E + (\gamma_{i9}lnV_i^E) + u_{it} \quad (3\text{-}11)$$

where

W_i = stock of appliance *i*
Y^E = Friedman permanent income
Y = current per capita income
E_i = price of appliance
G_i = price of gas-using competitor
H = number of urban residential over population
F = number of marriages
P^E = 3-year moving average of electricity price
V^E = 3-year moving average of gas price

3.1.1.2 Houthakker and Taylor (1970)

Houthakker and Taylor (1970) also use the definition of Fisher and Kaysen (1962) on white good stocks which are measured by the amount of Watt that the appliance can draw. Houthakker and Taylor (1970) develop two separate models for short-run and long-run based on utility theory as follows.

❖ *Short-run function*

Houthakker and Taylor consider demand as the utilization of capital stock. The utility is a function of price, income and other factors as set out below.

$$q = u(x, \pi, z)s \tag{3-12}$$

where
u(.) = utility function
x = income
π = price
z = other factors
s = the stock of appliance is measured by the amount of Watt that the appliances can draw as defined in Fisher and Kaysen (1962).

Let $u = \alpha_0 + \alpha_1 x + \alpha_2 \pi + \alpha_3 z$ (3-13)

The short-run demand (3-12) turn to

$$q = (\alpha_0 + \alpha_1 x + \alpha_2 \pi + \alpha_3 z)s \tag{3-14}$$

In short-run, the impact of price and income on electricity demand is

$$\frac{\partial q}{\partial x} = \alpha_1 s \tag{3-15}$$

$$\frac{\partial q}{\partial \pi} = \alpha_2 s \tag{3-16}$$

❖ *Long-run function*

The long-run model is derived from a state adjustment model developed by Houthakker and Taylor (1970). The name "state" adjustment model comes from the assumption that flows respond to the difference between actual and desired states.

In long-run, the expenditure flow on appliance stock is a function of the level of appliance stocks (state variables), income flows and the level of prices which can vary over time.

$$E(t) = \alpha + \beta s(t) + \gamma x(t) + \theta p(t) \tag{3-17}$$

where
E(t) = expenditure flows on new appliance during a very short time interval around *t*
s(t) = a state variable stands for appliance stock at time *t*
x(t) = income flows at the interval
p(t) = the level of price at time *t*

Estimating equation (3-17) faces two difficulties relating to the calculation of variables s(t). First, there is heterogeneity in appliance type. Second, the s(t) faces the problem of depreciation,

the rate of which is not known. Thus a reformulation is conducted to remove s(t). The model is as follows.

$$\dot{s}(t) = E(t) - w(t) \tag{3-18}$$

Equation (3-18) is an accounting identity. The left-hand side is the rate of change of appliance stock in the interval around time *t*. The right-hand side is the new acquired appliance in the interval with a deduction of the depreciation of appliance at the interval. If the depreciation is exponential at constant rate of δ then

$$w(t) = \delta s(t) \tag{3-19}$$

Substitute (3-19) to (3-18) we have

$$\dot{s}(t) = E(t) - \delta s(t) \tag{3-20}$$

Rearrange (3-17), we have

$$s(t) = \frac{1}{\beta}[E(t) - \alpha - \gamma x(t) - \theta p(t)] \tag{3-21}$$

Substitute (3-21) to (3-20)

$$\dot{s}(t) = E(t) - \frac{\delta}{\beta}[E(t) - \alpha - \gamma x(t) - \theta p(t)] \tag{3-22}$$

In long-run (at steady state), $\dot{s}(t) = 0$ then

$$\hat{E} = \frac{\alpha\delta}{\delta-\beta} + \frac{\gamma\delta}{\delta-\beta}x + \frac{\theta\delta}{\delta-\beta}p \tag{3-23}$$

The long-run impacts of income and price to electricity demand are $\frac{\gamma\delta}{\delta-\beta}$ and $\frac{\theta\delta}{\delta-\beta}$ respectively. It is noted that at the long-run $E(t) = \hat{E}$ and $s(t) = \hat{s}$, substitue to (3-17) we have

$$\hat{E} = \alpha + \beta\hat{s} + \gamma x + \theta p \tag{3-24}$$

Subtract the (3-20) from (3-17) we have

$$E(t) - \hat{E} = \beta[s(t) - \hat{s}] \tag{3-25}$$

Equation (3-25) is the explanation for the name "state adjustment model". The expenditure is adjusted to drive the state variable to its steady state value.

3.1.2 Empirical models for electricity demand estimation

So far, researchers have applied various ways to estimate empirically electricity demand. The ways can be different in term of data types, model structures, estimation techniques and so on (see Espey and Espey, 2004 for review). Some researchers utilize the structural form in which they jointly estimate the demand function for electricity and for appliance at the same time (Holtedahl and Joutz, 2004; Reiss and White, 2005). For example, Holtedahl and Joutz (2004) use a proxy variable of urbanization to represent the changes in appliance stocks not explained by income. Holtedahl and Joutz then use a vector autoregressive (VAR) system of four variables, including residential kWh per capita, price of electricity, disposal income per capita

and urbanization (as defined above). The VAR system is to find the long-run relationship between these variables via cointegration analysis. Besides the system, they employ a short-run error correction model (ECM) to estimate short-run elasticities.

However, many researchers employ the reduced form to estimate the electricity demand function. These researchers have developed a wide range of empirical strategies, the most important of which are described in the following section.

3.1.2.1 Empirical strategy for short-run function

❖ *Short-run function without appliance stock*

Though the theoretical model specifies that the short-run function should go with the level of appliance stock, some researchers have attempted to remove appliance stock from the short-run function (Fisher and Kaysen, 1962; Henson, 1984). For example, Fisher and Kaysen (1962) claim that measurement of the stock is tricky due to data quality. Thus they modify the theoretical model slightly to remove appliance stock. Let's go back to equation (3-10).

$$lnD_t - lnW_t^* = C_0 + \alpha lnP_t + \beta lnY_t$$

Fisher and Kaysen (1962) assume that the appliance stock grows exponentially at a constant rate of δ, so

$$W_t^* = W_0^* e^{\delta t} \qquad (3\text{-}26)$$

Substitute (3-26) to (3-10)

$$lnD_t = C_0 + \alpha lnP_t + \beta lnY_t + lnW_0^* + \delta t \qquad (3\text{-}27)$$

Take the first differences, (3-27) becomes

$$\Delta lnD_t = \delta + \alpha \Delta lnP_t + \beta \Delta lnY_t \qquad (3\text{-}28)$$

They then estimate equation (3-28) with data. The estimated $\hat{\alpha}$ and $\hat{\beta}$ are the responses of electricity consumption on changes in price and income. The influence of white stock is included in the intercept.

❖ *Short-run function with appliance stock*

There is considerable consensus among researcher that appliance level should be included in the short-run function. Table 3-1 summarizes some studies incorporating appliance on short-run demand function, the two most distinctive studies being Houthakker (1951) and Parti and Parti (1980). Houthakker (1951) is distinctive because it is the first one estimating short-run function with appliance stock. His function is as follows.

$$lnx = \alpha lnm + \beta lnp + \theta lng + \delta lnh + \varepsilon \qquad (3\text{-}29)$$

Where *x* is electricity consumption, *m* is income, *p* and *g* are price of electricity and gas, *h* is "the average holdings of heavy domestic equipment per customer". Houthakker (1951) justifies the presence of *h* as a representation of past and present prices for complemetary goods. Taylor (1975) comments that the interpretation supports the short-run form of the model. Taylor (1975, p. 84) also gives another justification for the interpretation of (3-29) as a short-run model.

> *Houthakker is silent as to whether the elasticities he has estimated refer to the short-run or the long-run. However, in view of the presence of the holdings of heavy electrical equipment as a predictor, they can be interpreted as representing the effect on consumption of changes in income and prices holding the stock of electricity-consuming capital goods fixed. This being the case, they should thus be interpreted [...] as short-run elasticities.* Taylor (1975, p. 84)

Parti and Parti (1980) is the second distinctive study since it provides a methods to estimate the average energy usage levels of individual appliances even when there is no direct observations for usage of the appliances.

$$E_i = f_i(V)A_i \qquad \text{i=1,..,N} \tag{3-30}$$

where
E_i = electricity consumed from appliance *i*
E_0 = electricity consumed from unobserved appliance
A_i = dummy var = 1 if own appliance *i*; =0 otherwise

If f_i is linear then

$$E_i = \sum_{j=0}^{M} b_{ij}\, V_j A_i \tag{3-31}$$

Where
V_0 = 1
E = total electricity consumption of a household

$$E = E_0 + \sum_{i=1}^{N} E_i = \sum_{i=0}^{N}\sum_{j=0}^{M} b_{ij} V_j A_i \tag{3-32}$$

Let $\bar{E}_i = b_{i0} + \sum_{j=1}^{M} b_{ij}\bar{V}_{ij}$ (3-33)

where
$\bar{E}_i$ = average energy usage of appliance i
$\bar{V}_{ij}$ = average value of the exogenous vars in the household that own the ith appliance

Then

$$E = \sum_{i=0}^{N} b_{i0}A_i + \sum_{i=0}^{N}\sum_{j=1}^{M} b_{ij}\left(V_j - \bar{V}_{ij}\right)A_i + \sum_{i=0}^{N}\sum_{j=1}^{M} b_{ij}\bar{V}_{ij}A_i \tag{3-34}$$

$$E = \sum_{i=0}^{N} \bar{E}_i[(A_i)] + \sum_{i=0}^{N}\sum_{j=1}^{M} b_{ij}[\left(V_j - \bar{V}_{ij}\right)A_i] \tag{3-35}$$

For the first term: $\bar{E}_i$ is the estimated coefficient in empirical of the appliance dummy var A_i — that is, the average energy use of the appliance i^{th}. For the second term: since the vector V is common for all appliance *i* (for example: price and income), then $\bar{V}_{ij} = \bar{V}_j$, then the estimated coefficient of the second term reveals the impact of V on E.

	Appliances variables
Houthakker (1951)	*h* is "the average holdings of heavy domestic equipment per customer".
Barnes *et al.* (1981)	• A vector of dummy variables for appliance, such as range, refrigerators, freezer, dishwasher, color TV, water heater, dryer. • Interaction between number of portable air conditioning units and cooling degree days (CDD). • Interaction between number of rooms, CDD and a dummy for the presence of central AC.
Branch (1993)	• A set of dummy variables for electric water heater, electric oven, microwave oven, freezer, clothes dryer, built-in dishwasher, portable dishwasher, and garbage disposal. • Variables for climate condition: - interaction between electric space heating and heating degree days (HDD) - interaction between AC and CDD - interaction between central AC and CDD.
Hsiao and Mountain (1985)	A set of dummy variables, including electric range, a second refrigerator, a dishwasher, a deep freezer, a washer, a color TV, a swimming pool filter, an air purifier, AC, electric heating, and water heating.
Parti and Parti (1980)	A vector of dummy variables for 16 specified appliance groups and an additional unspecified group.
Zhou and Teng (2013)	Dummy variables for AC, refrigerator and computer.

Table 3-1. Electric appliances in different research papers
Source. Author synthesized.

Table 3-1 shows the consensus on incorporating appliances in short-run demand function. However, there is no common rule for choosing appliances in the function. It seems that the appliance choices are normally based on electricity heavy-use appliance and on data availability.

Nevertheless, the popular form of short-run function for empirical strategy analysis is as follows.

$$lnE = \alpha_0 + \alpha_1 lnY + \alpha_2 lnP + \alpha_3 Z + \alpha_{4j} A_j D_j + \varepsilon \qquad (3\text{-}36)$$

where
E = household electricity consumption
Y = income vector
P = price
Z = related variables such as CDD/HDD, price of gas
A_j = vector of appliance
D_j = dummy vars = 1 if the household owns the asset j, = 0 otherwise

3.1.2.2 Empirical strategy for long-run function

❖ *A discrete time model for the state adjustment model*

Based on the state adjustment model for long-run electricity demand, Houthakker and Taylor (1970) propose a discrete time model for empirical estimation (see Taylor and Houthakker, 2009, pp. 19–22 for detail). The model below is a brief from Houthakker and Taylor (1970).

Recall equation (3-17) and (3-20), we have

$$E(t) = \alpha + \beta s(t) + \gamma x(t) + \theta p(t)$$

$\dot{s}(t) = E(t) - \delta s(t)$

Integrating the two equations over an interval from *t* to *t*+*h*, we have

$$E_t = \alpha h + \beta s_t + \gamma x_t + \theta p_t \quad (3\text{-}37)$$
$$\Delta^* s(t) = q_t - \delta s_t \quad (3\text{-}38)$$

Where $f_t = \int_t^{t+h} f(\tau)\, d\tau$ (3-39)
with f are E, s, x, p accordingly.

Apply the same procedure for the interval *t-h* to *t*, we have

$$E_{t-h} = \alpha h + \beta s_{t-h} + \gamma x_{t-h} + \theta p_{t-h} \quad (3\text{-}40)$$

Subtract (3-40) from (3-37) we have

$$E_t - E_{t-h} = \beta(s_t - s_{t-h}) + \gamma(x_t - x_{t-h}) + \theta(p_t - p_{t-h}) \quad (3\text{-}41)$$

Apply approximation by the mean value theorem to the interval from *t-h* to *t+h*, we have

$$s_t - s_{t-h} \cong \frac{h}{2}\left(\Delta^* s(t+h) + \Delta^* s(t)\right) \quad (3\text{-}42)$$

Substitute (3-38) to (3-42) we have

$$s_t - s_{t-h} \cong \frac{h}{2}[E_t + E_{t-h} - \delta(s_t + s_{t-h})] \quad (3\text{-}43)$$

Substitute (3-43) to (3-41)

$$E_t - E_{t-h} = \frac{h\beta}{2}[E_t + E_{t-h} - \delta(s_t + s_{t-h})] + \gamma(x_t - x_{t-h}) + \theta(p_t - p_{t-h}) \quad (3\text{-}44)$$

From equation (3-37), we can derive

$$s_t = \frac{1}{\beta}(q_t - \alpha h - \gamma x_t - \theta p_t) \quad (3\text{-}45)$$
$$s_{t-h} = \frac{1}{\beta}(q_{t-h} - \alpha h - \gamma x_{t-h} - \theta p_{t-h}) \quad (3\text{-}46)$$

Substitute (3-45) and (3-46) to (3-44), solve for E_t

$$E_t = B_0 + B_1 q_{t-h} + B_2 x_t + B_3 x_{t-h} + B_4 p_t + B_5 p_{t-h} \quad (3\text{-}47)$$

Where

$$B_0 = \frac{\alpha\delta h^2}{B} \quad (3\text{-}48)$$

$$B_1 = \frac{1 + \frac{h(\beta-\delta)}{2}}{B} \quad (3\text{-}49)$$

$$B_2 = \frac{\gamma(1 + \frac{\delta h}{2})}{B} \quad (3\text{-}50)$$

$$B_3 = -\frac{\gamma(1-\frac{\delta h}{2})}{B} \quad (3\text{-}51)$$

$$B_4 = \frac{\theta(1+\frac{\delta h}{2})}{B} \quad (3\text{-}52)$$

$$B_5 = -\frac{\theta(1-\frac{\delta h}{2})}{B} \quad (3\text{-}53)$$

$$B = 1 - \frac{h(\beta-\delta)}{2} \quad (3\text{-}54)$$

Estimating equation (3-47) with h=1 (1 lag) to get estimated coefficients B_1,.., B_5. We then can solve for five structural coefficients α, β, γ, δ, θ. Recall the equation (3-23) as below, we substitute the five structural coefficients to equation (3-23). Then we can calculate the long-run elasticity of demand on price and income.

$$\hat{E} = \frac{\alpha\delta}{\delta-\beta} + \frac{\gamma\delta}{\delta-\beta}x + \frac{\theta\delta}{\delta-\beta}p$$

In practice, Houthakker and Taylor (1970) get the empirical function as follows

$$q_t = \varphi_1 + \varphi_2 q_{t-1} + \varphi_3 x_t + \varphi_4 p_t + \varepsilon \quad (3\text{-}55)$$

❖ *A discrete time model for the flow adjustment model*

In addition to the state adjustment model, Houthakker and Taylor (1970) also derived a flow adjustment model. A flow adjustment model assumes that flows respond to the differences between actual and desired flows. The model is constructed in a similar way to the state adjustment model (see Taylor and Houthakker, 2009, pp. 18–22 for details).

This model assumes that there is a desired demand E_{it}^* of household *i* at time *t*. Assume that

$$E_{it}^* = \alpha P_{it}^{\gamma} y_{it}^{\beta} \quad (3\text{-}56)$$

The adjustment process is that the ratio of the demand this period to last period is proportional to the ratio of desired demand this period to actual demand for the last period. The mathematic expression is as follows.

$$\frac{E_{it}}{E_{i,t-1}} = \left(\frac{E_{it}^*}{E_{i,t-1}}\right)^{\theta} \quad (3\text{-}57)$$

Substitute (3-56) to (3-57) then take the logarithms on both sides, rearrange to solve for E_{it}, we have

$$lnE_{it} = \theta ln\alpha + \theta\gamma lnp_{it} + \theta\beta lny_{it} + (1-\theta)lnE_{i,t-1} \quad (3\text{-}58)$$

Short-run elasticities of price and income are $\theta\gamma$ and $\theta\beta$. (3-59)

Long-run elasticities of price and income are $\theta\gamma/(1-\theta)$ and $\theta\beta/(1-\theta)$. (3-60)

In comparison to the short-run specification which requires only cross-sectional data, the long-run model requires panel data which is more complex to obtain. However, the long-run

specification has two advantages. First, it allows an estimate to be made of the respond behavior in both short-run and long-run scenarios. Second, it allows an estimate to be made of the short-run elasticity without a data requirement for appliances.

Due to these advantages, a series of researchers have applied long-run models with various techniques for different purposes. Hsing (1994) specifies a model that current level of electricity consumption (QE_{it}) depends on the past consumption ($QE_{i,t-1}$), price of electricity, disposable personal income, price of gas, cooling/heating degree days. The model is fitted to five southern states (US) during 1981–90 with different techniques, including ordinary least square (OLS), cross-sectionally correlated and time-wise autoregressive model (CSCTWA) and cross-sectionally heteroscedastic and time-wise autoregressive model (CSHTWA).

Garcia and Cerrutti (2000) employs a log-linear demand function on one-year lag of electricity demand, personal income, price of electricity, price of gas and CDD/HDD. The model is estimated with panel county data from California for 1983–97. The technic is to use dynamic random variables model.

Bernstein and Griffin (2006) estimates a function of electricity demand with the independent variable including not only the lag of demand, but also the lag of other independent variables such as energy price, population, income, and climate conditions. The model is fitted with electricity data from 1977–99 and gas data from 1977–2004. The aim was to determine whether the impact of prices on demand of energy differed at regional, state or sub-state level.

The same long-run model can be seen in Alberini and Filippini (2011) and Alberini *et al.* (2011). Both studies state the long-run model is a dynamic model with a lag of electricity demand on the right-hand side. The former uses the model to compare the efficiency of different estimators, including Kiviet corrected least square dummy variables and the Blundell-Bond estimators. The latter is to analyze the role of prices and income for residential consumption of gas and electricity in the US.

Frondel *et al.* (2019) is the most recent research employing the long-run model to estimate the price elasticity of residential electricity demand in Germany. This research is based on a panel data from the German Residential Energy Consumption Survey (GRECS) from 2006–14. Frondel *et al.* (2019) apply the dynamic Blundell-Bond estimator on the data. The estimated short-run and long-run elasticity are -0.44 and -0.66 respectively.

3.2 Model specification

The detailed model specification for this study is based on the above literature review and the data availability provided in Chapter 4. In general, the study estimates electricity demand in reduced form in both short-run and long-run.

The short-run demand function has a similar form as in equation (3-36).

$$lnE = \alpha_0 + \alpha_1 lny + \alpha_2 lnP + \alpha_3 Z + \alpha_{4j} A_j D_j + \varepsilon \qquad (3\text{-}61)$$

Where

E = household electricity consumption
y = household (per capita) income
P = price
Z = related variables such as CDD/HDD, price of gas
A_j = vector of appliance
D_j = dummy vars = 1 if the household owns the asset j, = 0 otherwise

The long-run demand function has a similar form as in equation (3-58).

$$lnE_{it} = \alpha + \beta lnp_{it} + \gamma lny_{it} + \theta Z_{it} + \delta lnE_{i,t-1} + \varepsilon_t \quad (3\text{-}62)$$

E_{it} = electricity consumption at period *t*
$E_{i,t-1}$ = electricity consumption at period *t-1*
p_{it} = price of electricity at period *t*
y_{it} = household (per capita) income at period *t*
Z_{it} = a vector of other factor such as price of gas, CDD/HDD

Chapter 4. Data

This study utilizes data at the household level. The estimation needs four types of data:

(i) electricity consumption and appliance availability at household level
(ii) economic condition (e.g. income and expenditure) and demographics (such as household size, household composition)
(iii) prices of electricity
(iv) climate condition (e.g. temperature related to the billing period).

Data types (i) and (ii) can be obtained from the Vietnam Household Living Standard Surveys (VHLSS). Data type (iii) can be obtained from various legal document regarding electricity prices. Data (iv) can be obtained from the US National Oceanic and Atmospheric Administration (NOAA).

These data sources have been merged and then divided into two separate data sets to estimate short-run and long-run demand function. The following section provides detailed information about each data source, as well as the procedure used to process the data.

4.1 Temperature data

The temperature data was obtained from the Global Historical Climatology Network (GHCN) of NOAA's National Centers for Environmental Information. GHCN provides daily temperature data for 14 weather stations in Vietnam (Table 4-1). While there are 173 weather stations across Vietnam (Nguyen Tan Dung, 2007), data from them are hard to access. The best up-to-date, free data source is GHCN.

Station id	*Station name*	*Latitude*	*Longitude*
48840	THANH HOA, VM	19.750	105.783
48830	LANG SON, VM	21.833	106.767
48820	NOIBAI INTERNATIONAL, VM	21.221	105.807
48887	PHAN THIET, VM	10.933	108.100
48826	PHU LIEN, VM	20.800	106.633
48848	DONG HOI, VM	17.483	106.600
48914	CA MAU, VM	9.1830	105.150
48806	SON LA, VM	21.333	103.900
48917	PHU QUOC, VM	10.217	103.967
48855	DANANG INTERNATIONAL, VM	16.044	108.199
48845	VINH, VM	18.737	105.671
48900	TAN SON HOA, VM	10.820	106.670
48825	HA DONG, VM	20.967	105.767
48808	CAO BANG, VM	22.667	106.250

Table 4-1. List of weather stations in GHCN data
Source. Author compiled from GHCN data.

The data downloaded from GHCN includes both detail information on the stations and the temperature data. The station information includes station codes, names, position in longitude and latitude, and establishment date. The temperature data includes minimum, maximum and average daily temperature. Unless specified explicitly, all the temperatures in this study are average temperature.

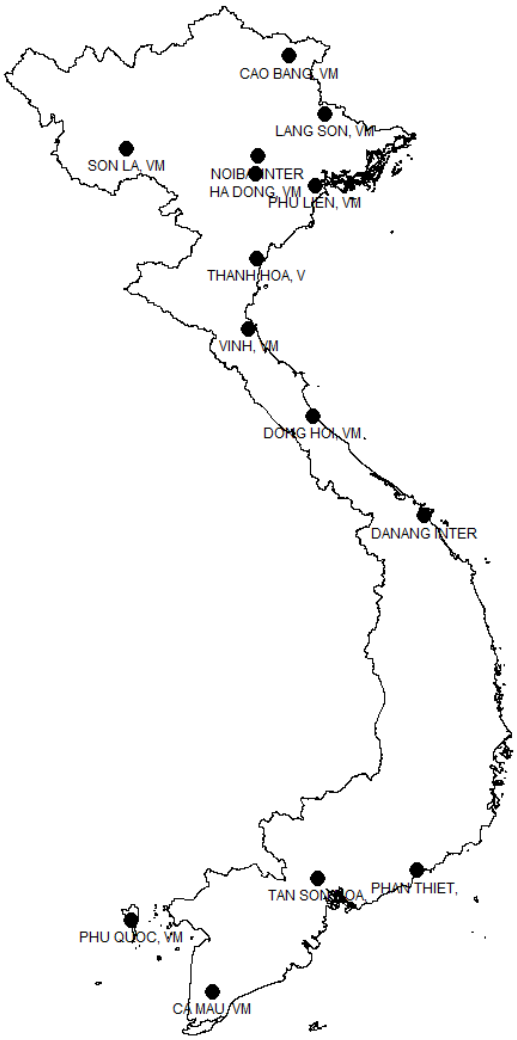


Figure 4-1. Position of the 14 weather stations
Source. Author illustrated.

As shown in Figure 4-1, the 14 weather stations are distributed in a pattern that covers the whole Vietnam area. Thus it is reasonable to use data from the 14 stations as representative of temperature data across Vietnam.

The temperature data is converted to cooling degree days (CDD) and heating degree days (HDD). CDD is the amount of temperature that needs to be cooled down to reach a certain base temperature. The certain base θ is the level of thermal comfort. CDD for a day *i* is max(T_i – θ, 0). CDD for a month is the sum of CDD for all days of a month. In contrast, HDD is the amount of temperature that needs to be increased to ensure a certain level of thermal comfort (δ). HDD for day *i* is the max(δ – T_i, 0). HDD for a month is the sum of HDD for all days of a month.

4.2 Vietnam Household Living Standard Survey data

4.2.1 Overview of VHLSS

The Vietnam Household Living Standard Survey (VHLSS) is a major micro data source on household welfare in Vietnam. It provides all information related to income and expenditure of households. The precursor of VHLSS is the two surveys Multi-Purpose Household Survey (MPHS) and Vietnam Living Standard Survey (VLSS) (Phung and Nguyen, no date).

The MPHS was conducted every one to two years from 1994, with a sample of 25,000-47,000 households. It collected data on household income and expenditure. However, expenditure on healthcare, education and employment were collected for different years.

The VLSS was conducted in two rounds: the first in 1992–93 with 4,800 households; and the second in 1997–98 with 6,000 households. The two surveys overlapped in some indicators (Phung and Nguyen, no date). MPHS was larger, but not standardized, while VLSS was smaller, though more standardized (Phung and Nguyen, no date). There was a need to merge the two surveys in this study.

In 2002, the General Statistics Office of Vietnam (GSO) started to conduct VHLSS to replace both MPHS and VLSS. VHLSS collects data on income and expenditure at household levels. In some specific years (e.g. 2008 and 2014), VHLSS adds questionnaires on weight for the Consumer Price Index (CPI).

So far, VHLSS has been conducted over two periods. In the 2000–2010 period, VHLSS was conducted biennially from 2002 (Phung and Nguyen, no date). In the 2011–2020 period, VHLSS has been conducted annually. Both income and expenditure are collected in even years, while only demographic, employment and income data are collected in odd years (GSOVN, 2011a).

Data type	*2008*	*2010*	*2012*	*2014*	*2016*
Income	x	x	x	x	x
Income and Expenditure	x	x	x	x	x
Weight of CPI	x			x	

Note. Each data type is collected in a separate sample.
Table 4-2. Available data on each round of VHLSS 2008–16
Source. Author compiled.

Each VHLSS round has four related documents:
- Decision to conduct VHLSS issued by the Director of GSO
- Detail implementation plan attached to the Decision
- Handbook for implementation
- Results of Survey.

4.2.2 VHLSS samples

❖ *Target population*

The survey targets the civilian, non-institutionalized population of Vietnam at household level.

❖ *Sampling frame*

- Period of 2000–2010: 16,470 enumeration areas (EA) of Census 1999 (GSOVN, 2010, p. 3)
- Period of 2011–2020: 172,000 EA of Census 2009 (BCĐTW, 2010, p. 10)

Each EA consists of about 100 households in both census 1999 (Phung and Nguyen, no date, p. 16) and census 2009 (BCĐTW, 2010, p. 10).

❖ *Sample sizes*

At each round of VHLSS, separate samples are collected for different data types. Table 4-3 shows that there are three types of data, including income, income and expenditure, and weight of CPI. Thus there are three sample sizes related to each VHLSS round.

Data	*Questionnaire*	*2008*	*2010*	*2012*	*2014*	*2016*
Sampling frame		Census 1999		Census 2009		
Income	1A	36,756	36,756	37,596	37,596	37,596
Income and Expenditure	1B	9,189	9,189	9,399	9,399	9,399
Weight of CPI	1C	24,189			25,090	
Total		70,134	45,945	46,995	72,085	

Note. The sizes are compiled from various detail implementation plans of VHLSS.
The sample sizes in actual data are approximate to the sizes.

Table 4-3. Sample sizes of VHLSS from 2008–2014

Source. GSOVN(2008), GSOVN(2010), GSOVN(2011b) GSOVN(2013) and GSOVN(2015).

❖ *Sampling*

The sampling of VHLSS is stratified random sampling.

❖ *Stratification*

Stratification of VHLSS consists of three stages

- Primary sample unit (PSU): communes/wards
- Secondary sample unit (SSU): Enumeration Areas
- Tertiary sample unit (TSU): households.

(i) First stage for PSUs

First, the list of communes in the sampling frame is allocated to a strata of joining between province and urban/rural. For example, strata 1: HCMC urban, strata 2: HCMC rural, strata 3: Ha Giang urban, strata 4: Ha Giang rural, and so on. In each strata, the number of selected wards is determined by the rule of probability proportionate to size (PPS). PPS means that the number of selected households is proportioned to the square root of number of households in each strata. PPS can balance between the equality across province and the proportion size of each province.

(ii) Second stage for SSUs

In each ward, three EAs are selected. However, in each round of VHLSS, only one EA is selected. The two other EAs are for rotation in the next rounds. Thus Phung and Nguyen (no date, p. 17) argues that though the stratification is a three-stage design, it works technically as a two-stage design.

	2008	*2010*	*2012*	*2014*	*2016*
Number of wards/EA	3,063	3,063	3,133	3,133	3,133

Note. This does not apply for weight of CPI sample.

Table 4-4. Number of wards/EA in each round of VHLSS

Source. Author compiled.

(iii) Third stage for TSUs

In each EA, a fixed number of households is selected for each type of information.

	2008	*2010*	*2012*	*2014*	*2016*
Income (1A)	3	3	3	3	3
Income and Expenditure (1B)	12	12	12	12	12

Note. There are two additional standby households for 1A and three for 1B in each EA.

Table 4-5. Number of households selected in each EA

Source. Author compiled.

The sampling process shows that VHLSS is representative at national and provincial level.

4.2.3 Sample rotation, survey time and weight

❖ *Sample rotation*

From one round of VHLSS to the next round, 50 per cent of samples are rotated. Instead of rotating households, VHLSS rotates EAs. The EAs of half of the commune are retained while the EAs of the other half are rotated (Phung and Nguyen, no date, p. 24). For example, 50 per cent of EAs in VHLSS 2008 participated in 2006 in which 25 per cent participate in both 2004 and 2006 and 25 per cent only show up in 2006.

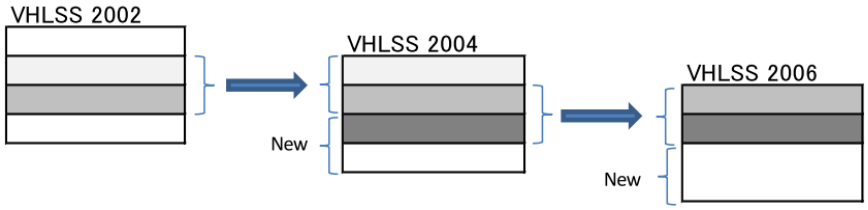


Figure 4-2. Rotation for VHLSS sampling
Source. Furuta (2014).

The rotation process implies that panel data of a 4-round period cannot be constructed.

❖ *Survey time*

VHLSS	*Waves*	*Period (months)*
2008	May/ Sep	2
2010	Jun/ Sep/ Dec	1
2012	Mar/ Jun/ Sep/ Dec	1
2014	Mar/ Jun/ Sep/ Dec	1
2016	Mar/ Jun/ Sep/ Dec	1

Table 4-6. Survey plan for VHLSS 2008–2014
Source. GSOVN (2008, 2010, 2011b, 2013, 2015).

❖ *Weights in VHLSS*

In the data, there are three variables for household weight: wt9; wt36; and wt45. The weight wt9 is the weight for sample 1A (about 9,000 households), the wt36 is the weight for sample 1B (about 36,000 households) and the weight wt45 is the common weight for both samples (about 45,000 households).

The weight included in VHLSS data is the household weight. The individual weight is the product of household weight with household size. This study employs household weight for all estimation.

4.2.4 Data process from VHLSS

Data is extracted and processed to household levels. There are seven main groups of data:

- administration such as household id, location, weight, survey month
- household head data, including gender, age, education
- demographics, including household size, household composition
- economic conditions, including annual/monthly income and expenditure at both total and per capita

- housing condition, including ownership, area in square meter
- appliances, including electric appliances of the household
- energy, including both electricity and other energy expenditure, in which electricity expenditure covers three important areas
  - kWh consumed at the month prior to the survey month
  - electricity bill at the month prior to the survey month
  - annual electricity bill.

4.3 Merging electricity price with VHLSS

❖ *Price schedules*

VHLSS contains data on electricity bills and kWh consumption for the month prior to the survey month. If the price schedule for the billing month is available, both marginal prices and average prices for each household at the billing month can be calculated precisely.

There is an obstacle in determining appropriate price schedule. As discussed in Chapter 2, Vietnam has two price schedules for two different groups. The first is the retail price schedule applied to roughly 85 per cent of households in Vietnam who can buy electricity directly from EVN. The second is the wholesale prices applied to rural areas which are remote, have low population density and unorganized infrastructure. The retail price for the second group is unknown.

In this circumstance, I propose to apply the retail price schedule for the whole sample. Three reasons justify this choice. First, the wholesale price is applied to a small fraction of the population (about 15 per cent) and the fraction is diminishing. Second, though there is heterogeneity in price setting by rural electricity distribution organizations, they go in a sample pattern of flat rate or IBTs. There is no decreasing block tariffs. Thus the prices are still positively correlated with retail prices.

Variable	*Obs*	*Mean*	*Std. Err.*	*Std. Dev.*	*[95% Conf. Interval]*	
kwh	8,305	195.951	1.846	168.211	192.333	199.569
kwh1	8,305	179.791	1.388	126.448	177.071	182.511
diff	8,305	16.160	0.977	89.050	14.245	18.076
t = 16.538						p-value=0.0000

Note. The table is calculated for urban area only and for all three years.
kWh: original kWh; kWh1: derived kWh

Table 4-7. Paired t-test for mean comparison between original and calculated kWh

Source. Author estimated.

Finally, one can argue that we can use average price calculated directly from electricity bills and kWh consumption from questionnaires. However, the original data of electricity bills and kWh consumption are not consistent. The wholesale price only applied to a small part of the rural area implying that all urban areas had the retail price only applied. Thus in urban area there should be no difference between the kWh in questionnaires and the kWh calculated from electricity bills and the corresponding retail price schedule. However, a paired sample t-test (Table 4-7) shows a significant difference between the two kWh at 0.01 level.

Since the data on kWh consumption and electricity bills are inconsistent, I propose to choose electricity bills data and derive the kWh from the bills. This is considered reasonable since kWh is more abstract to people than a bill in monetary terms because people normally remember their last month's bill rather than remember how many kWh they consumed.

❖ *Calculating electricity prices for each household in VHLSS*

The procedure used to calculate prices is as follows.

- For each tariff schedule, calculate the bills if consume up to the lower bound and upper bound of each block. We denote the bills for lower and upper bounds as LB and UB. For each block we have (LkWh, UkWh, LB, UB) where LkWh, UkWh are the upper and lower bound of kWh, and LB, UB are the corresponding bill if consumed up to the LkWh and UkWh.
- For each household, find the survey month. The month prior to the survey month is the billing month.
- From the list of tariff schedules, find the appropriate tariff schedule for the billing month.
- Choose the block in the tariff schedule that the household's bill falls within the range [LB, UB] of the block.
- Assign the price of the block to the household as the marginal price (MP)
- Calculate the derived kWh as the following:

  kWh1 = (household's bill-LB)/MP + LkWh
- Calculate the average price AP = Electricity bill/ kWh1.

4.4 Merging temperature data with VHLSS

VHLSS has data on the address at the enumerator level of households and the survey month. These data can help to determine the temperature that households face in the corresponding billing periods. The temperature for each household is determined by the following rule.

- For each household, identifying the nearest available weather station at the household's billing period. It should be noted that the billing period is always one month prior to the survey month. Each round in VHLSS is conducted in four waves of March, June, September and December. Thus we have billing periods in the four months of February, May, August and November.
- Assigning the temperature of the weather station at that billing period to the household.

The nearest station is determined by calculating and comparing the distance between the location of a household and the location of all weather stations. A household location is represented by the center coordinates of the ward where the household locates because wards are the finest available geographic location of households. Thus for each household, the distances from its ward center coordinates to each available weather station are calculated. The station that has minimum distance is chosen as the nearest one.

The distance is calculated by the Haversine formula (Sinnott, 1984) as follows.

$$dlon = lon_2 - lon_1 \tag{4-1}$$

$$dlat = lat_2 - lat_1 \tag{4-2}$$

$$a = \left(\sin(^{dlat}/_{2})\right)^2 + \cos(lat_1) * \cos(lat_2) * \left(\sin(^{dlon}/_{2})\right)^2 \tag{4-3}$$

$$c = 2 * arcsin\left(min\left(1, sqrt(a)\right)\right) \tag{4-4}$$

$$d = R * c \tag{4-5}$$

R is the radius of the Earth which is 6,373 km.

Latitude and longitude in Haversine formula are in radians form. The latitude and longitude in the data are in decimal degrees. Thus we need to convert the latitude and longitude in data by multiplying them by PI/180.

The calculation of the above formula needs Vietnamese shapefile at ward level. The shapefile provides the center coordinates of each ward. Shapefile is a set of files storing geometric shape like points, lines, and polygons of geographic areas. However, the most available updated official shapefile is from 2008. The shapefile does not catch up with the changes in administration units from 2008 to present, thus cannot be used for 2012–16 data. However, updated, free shapefiles are available, the best source of which is Global Administrative Areas (GADM). The most updated shapefile of Vietnam in GADM is April 2018 which is suitable to match with the data.

However, while the free GADM shapefile has ward names, the VHLSS data has ward codes only. Therefore, we need to assign ward codes to corresponding ward names in the shapefile. We perform the task by utilizing the official administration list issued by the General Statistics Office (GSO). The official list contains both ward names and ward codes. Merging the GSO list with the shapefile by ward name results in a new shapefile having ward codes. The new shapefile is then able to merge with our data. The assigning process is detailed in the Annex 1.

4.5 Deflating monetary variables to a base year and cleaning data

All monetary variables are adjusted to the same based price by the CPI index. CPI is the annual average consumer price index, with the previous year being 100. The data is from GSOVN (2018). The base year is denoted by the suffix *1, 2* and *3*. The base years of 2012, 2014 and 2016 correspond to the suffixes of *1, 2, 3*.

The cleaning process removes six cases of negative income which may come from data input errors. The process also removes all households which do not connect to national grid or have zero electricity use last month. There are 715 households in this category accounting for 2.6 per cent of the whole sample. Due to the large size of the sample, and no evidence of extreme outliers, I have not removed any case for outlier reasons.

4.6 Data sets for short-run and long-run demand function

Short-run and long-run electricity demand functions have different requirements for data. The short-run function requires that appliances in households are assumed to be fixed. Thus the data for short-run function should not have the same households at different times, while the data for long-run function requires lags of electricity consumption and thus needs the same households at different points of time.

Therefore, we decide to separate the master data set into two separate data sets for short-run and long-run estimation. The data separation is based on the rotation rule in VHLSS. Due to the rotation rule, each round of VHLSS contains a 50 per cent sample of the previous round and 25 per cent of the round prior to that. I propose to separate the two data sets in following way.

- The pooled data set for short-run includes (i) the whole sample in 2016, (ii) the 2014 sample part which does not overlap with the 2016 sample, and (iii) the 2012 sample part which does not overlap with the 2014 sample.
- The panel data set for long-run includes all households that appear in all three years: 2012, 2014 and 2016.

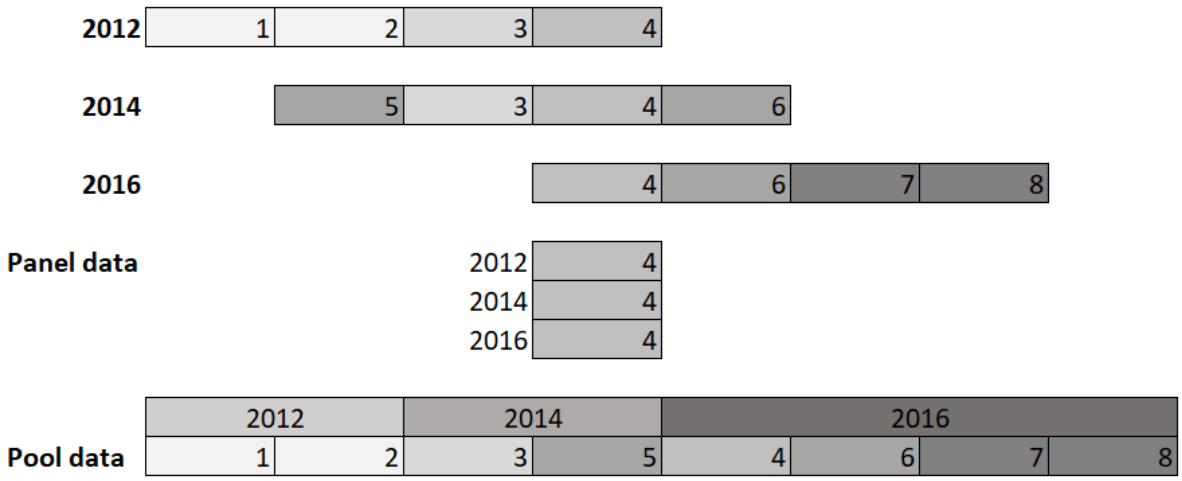


Figure 4-3. The construction of data sets for short-run and long-run functions
Source. Author illustrated.

4.7 Descriptive statistics

The list of variables and description are detailed in Appendix B. The two data sets are similar in summary descriptive for most of major variables.

		Panel data set		*Pooled data set*	
Group	*Variables*	*Mean*	*Sd./ Cum %*	*Mean*	*Sd./ Cum %*
Administration	urban				
	Rural	67.56	67.56	67.04	67.04
	Urban	32.44	100.00	32.96	100.00
	reg6				
	Red River Delta	26.08	26.08	24.58	24.58
	Midlands & Nth. Mountains	12.27	38.35	11.79	36.37
	Nth. and Coastal Central	22.42	60.77	21.61	57.98
	Central Highlands	5.81	66.58	5.69	63.67
	Southeast	14.52	81.10	17.81	81.48
	Mekong Delta	18.90	100.00	18.52	100.00
Head	male				
	Female	25.94	25.94	26.92	26.92
	Male	74.06	100.00	73.08	100.00
	edu				
	None	18.94	18.94	16.97	16.97
	Primary School	25.47	44.40	24.63	41.60
	Secondary School	30.08	74.49	29.77	71.36
	High School	14.48	88.97	15.91	87.27
	A degree	6.83	95.80	7.95	95.23
	.	4.20	100.00	4.77	100.00
	age	52.71	13.85	51.31	14.20
Demographic	hhsize	3.77	1.55	3.77	1.57
	sexratio	0.47	0.21	0.48	0.21
	elderratio	0.19	0.31	0.17	0.30
	childrenratio	0.20	0.21	0.21	0.21
	highestedu				
	None	5.70	5.70	5.86	5.86
	Primary School	18.09	23.79	16.40	22.26

Group	Variables	Panel data set		Pooled data set	
		Mean	*Sd./ Cum %*	*Mean*	*Sd./ Cum %*
	Secondary School	29.19	52.98	28.48	50.74
	High School	28.99	81.97	30.41	81.14
	A degree	16.99	98.96	17.74	98.88
	.	1.04	100.00	1.12	100.00
Economics	incpc	2,739.24	2,723.58	2,838.26	2,853.47
	incpc1	2,506.84	2,492.19	2,571.80	2,602.97
	exppc	2,745.99	3,159.33	2,818.19	2,734.67
	exppc1	2,510.81	2,843.06	2,552.27	2,444.89
Housing	sqm	82.13	52.47	80.97	53.70
	pcsqm	25.47	20.63	24.91	20.30
	ownership				
	No	3.04	3.04	5.34	5.34
	Yes	96.95	99.99	94.56	99.91
	.	0.01	100.00	0.09	100.00
Appliances	aircon	0.13	0.34	0.15	0.36
	fan	0.91	0.28	0.91	0.29
	fridge	0.63	0.48	0.63	0.48
	washing	0.31	0.46	0.31	0.46
	oven	0.07	0.25	0.08	0.27
	computer	0.20	0.40	0.22	0.41
	tv	0.94	0.23	0.94	0.24
	waterheater	0.25	0.43	0.24	0.43
	generator	0.02	0.13	0.02	0.13
Energy	lpg	109.45	113.90	107.63	114.26
	lpg1	101.18	107.25	98.35	105.18
Electricity	eleclm	240.76	250.53	241.61	254.67
	eleclm1	219.69	224.92	218.02	226.41
	kwh	148.70	132.01	146.73	145.01
	kwh1	135.72	111.74	134.20	111.71
	elecly	2,840.83	2,877.98	2,809.84	2,798.05
	elecly1	2,592.42	2,589.83	2,536.11	2,495.09
	mp	1.82	0.36	1.86	0.38
	mp1	1.68	0.31	1.68	0.32
	rsp	41.77	61.22	43.17	61.70
	rsp1	38.15	55.48	38.89	55.18
	ap	1.64	0.20	1.66	0.21
	ap1	1.50	0.15	1.51	0.15
	w	0.03	0.02	0.03	0.02

Note. Description of variables is detailed in Table B-1, Appendix B.
Table 4-8. Similarity between data sets for short-run and long-run function
Source. Author compiled.

The differences between the two data sets are mainly in survey year, billing month, temperature, and number of price schedules, as detailed in the Table 4-9. Obviously, the pooled data has a wider span of time and has more tariff schedule variation than the panel data.

Variable group	Panel data set			Pooled data set		
	Variable	*Mean*	*Sd./ Cum %*	*Variable*	*Mean*	*Sd./ Cum %*
Administration	year			year		
	2012	31.59	31.59	2012	25.46	25.46
	2014	33.5	65.09	2014	26.15	51.62
	2016	34.91	100	2016	48.38	100.00
	month			month		
	2012m2	0.03	0.03	2012m2	2.00	2.03
	2012m3	0.1	0.13	2012m3	4.31	6.34
	2012m5	0.16	0.29	2012m5	6.15	12.62
	2012m7	0.04	0.33	2012m8	5.90	18.58
	2012m8	16.35	16.68	2012m11	6.75	25.46
	2012m9	0.33	17.01	2014m2	11.47	36.94
	2012m10	0.29	17.3	2014m5	11.88	49.04
	2012m11	14.29	31.59	2016m2	11.50	63.12
	2014m11	33.5	65.09	2016m5	11.85	74.98
	2016m8	34.91	100	2016m8	12.95	87.92
				2016m11	12.08	100.00
				2016m8	12.95	87.92
				2016m11	12.08	100.00
Temperature	cdd25	78.82795	47.42493	cdd25	75.67	53.45
	hdd15	0.005584	0.096321	hdd15	2.86	9.30
	2012m1 -			2011m3 -		
Price group	2012m6	0.29	0.29	2011m12	0.03	0.03
	2012m7 -			2012m1 -		
	2012m12	31.3	31.59	2012m6	12.64	12.67
	2014m6 -			2012m7 -		
	2015m3	33.5	65.09	2012m12	12.79	25.46
	2015m4 -			2013m8 -		
	2017m11	34.91	100	2014m5	23.58	49.04
				2014m6 -		
				2015m3	2.57	51.62
				2015m4 -		
				2017m11	48.38	100

Note. For month in the pooled data set, only keep categories having percentage over 2%.
Table 4-9. Differences between data sets for short-run and long-run functions
Source. Author compiled.

The two tables above are the brief compared summary statistics of some major variables in the two data sets. The full data descriptive for each data set are in Appendix B.

Chapter 5. Price elasticity of residential electricity demand

5.1 Introduction

Price elasticity of residential electricity demand is the key number for policy makers in designing electricity tariff schedules. Much research has been carried out on this topic, though mostly using data at an aggregate level (Houthakker, Verleger and Sheehan, 1974; Hsing, 1994; Holtedahl and Joutz, 2004; Alberini and Filippini, 2011). A recent new trend has been to employ micro data at the household level, but the micro data is often either missing data in price (Branch, 1993; Alberini, Gans and Velez-Lopez, 2011) or is narrowed to a regional level rather than a national level due to the absence of national data on tariff structures (Reiss and White, 2005; Zhou and Teng, 2013). Thus this chapter aims to estimate price elasticity using household data at national scale with full detailed IBT schedules. In addition, the chapter also investigates whether households respond to marginal price or average price under an IBT schedule.

The context of the chapter is Vietnam from 2012–16. As discussed in Chapter 2, the residential electricity market in Vietnam is a monopoly market. The single seller is Vietnam Electricity (EVN) – a state-owned enterprise. Electricity prices for residential in Vietnam has been in IBT form since 1994. The IBT schedule is set by government. According to EVN (2015), the IBT form aims to encourage a saving attitude and to ensure a low price for low-income households.

In this context, the results of this study provide empirical evidence, not only to broaden our understanding of price elasticity under IBTs, but also for the Vietnamese Government in designing appropriate IBTs schedules. The following four sections of this chapter provide a literature review, model specifications and data, results and discussion, and conclusions.

5.2 Literature review

IBTs are non-linear prices. The higher kWh a household consumes, the higher the price a household pays. For example, a simple two-block IBTs has the following structure:

$$p = \begin{cases} p_1, 0 \leq q < z_1 \\ p_2,\ z_1 \leq q \end{cases}$$

IBTs pose two major problems for measuring the impact of price on electricity demand. First, it poses a question of what kind of prices should be used: the marginal price of the last block or the average price of all blocks. Second, it poses an empirical problem of endogeneity since both prices and quantity are jointly determined.

5.2.1 Controversy in the type of prices

The classic treatment with block tariffs is from neo-classical ordinal utility theory with an assumption of perfect information (Taylor, 1975; Nordin, 1976). The theory indicates that the tariff structure has two effects on demand. The first one is the effect from changes of marginal price, which is the price of the last consumed block. Households maximize their utility subject to budget constraints. The budget constraint is constructed from IBTs. Thus whenever the marginal price changes, households change their behaviors to cope with new marginal prices, as shown in the Figure 5-1.

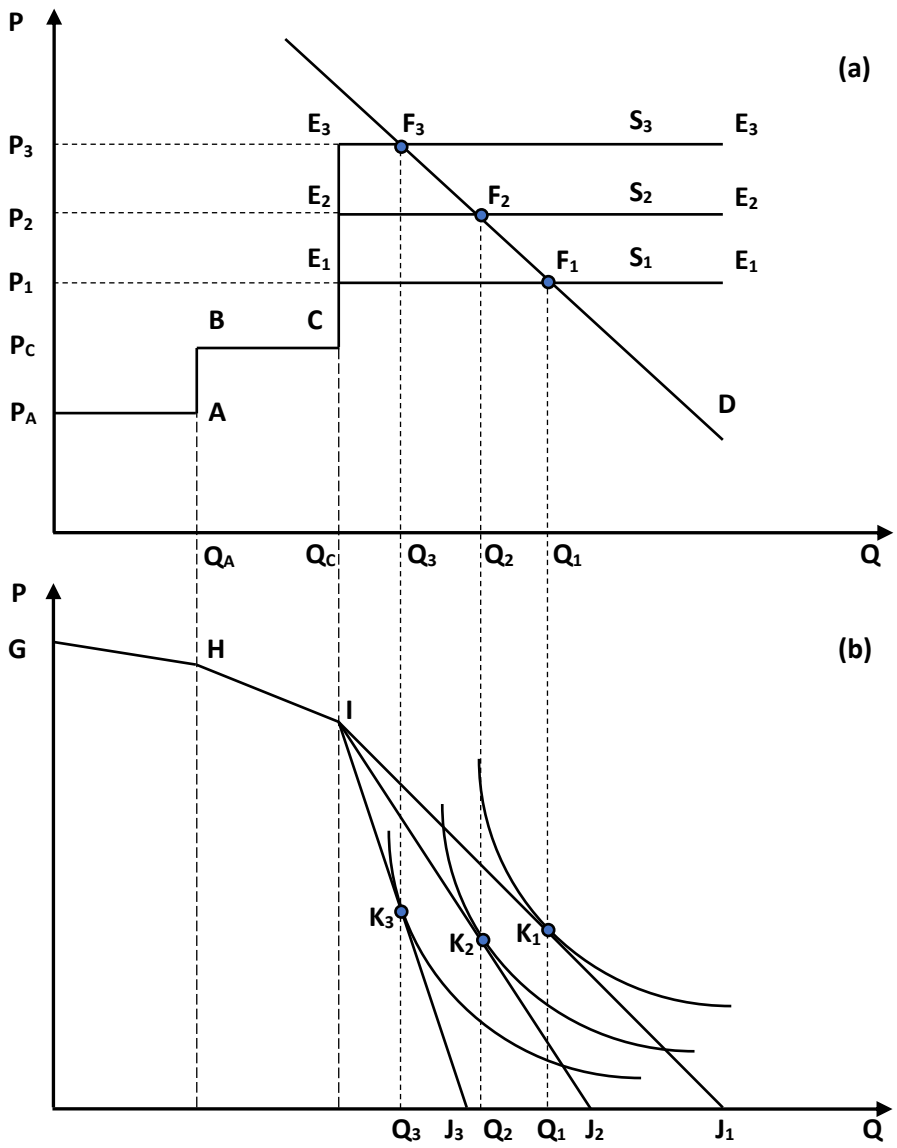


Figure 5-1. Changes of marginal price and demand
Source. Billings and Agthe (1980).

Figure 5-1 shows the relationship between marginal price and demand. Graph (a) illustrates a three-block IBTs. From zero to Q_A , price is P_A. From Q_A to Q_C , price is P_C. Consumption above Q_C has the price of P_1. P_1 is marginal price. P_2 and P_3 represent for increases in marginal price. The price line of $P_A ABCE_1$ works technically as an upward supply curve. Meanwhile, graph (b) illustrates the corresponding budget lines and consumer choices under the assumption of maximizing utility. Due to the IBT form of price, the budget lines are not linear. When marginal price is P_1, the budget line is $GHIJ_1$. Households maximize their utility at K_1 and consume at the corresponding Q_1. When marginal price increases to P_2, the budget line rotates inward to $GHIJ_2$. Households choose K_2 and Q_2. When marginal price increases to P_3, households choose K_3 and Q_3. Taking the quantity of Q_1, Q_2, Q_3 to graph (a) with corresponding prices, household choices would be F_1, F_2, F_3. The line that connects the three points of F_1, F_2, F_3 is the demand curve.

The second effect of IBTs on demand comes from changes to intra-marginal prices which is the price of any block prior to the last block. Taylor (1975) shows that as long as changes in intra-marginal prices do not cause block switching, the changes would create an impact which is

equivalent to an income effect (Figure 5-2). If the price of the first block (P_1) decreases while the price of the second block (P_2) stays constant, the budget line would change from ABC to AB'C'. The parts BC and B'C' are parallel since there is no change in P_2 and the price of good 2. Customers react by increasing the kWh demanded from q2 to q3. Due to the parallel feature of BC and B'C', the change from q2 to q3 is purely an income effect. It is analogous to an increase in customer income.

Based on Taylor (1975), Nordin (1976) presents a way to take into account the effect of intra-marginal price. The effect is measured by the actual total payment for all previous blocks minus the payment for all previous blocks if the price for all previous blocks is at marginal price. Nordin calls this element x2c to separate it from x2a and x2b of Taylor (1975) which are the average price of all previous blocks and total payment for all previous blocks. Researchers later call x2c the Rate Structure Premium or sometimes, "the difference". The term Rate Structure Premium (RSP) is adopted from now to denote the effect.

It should be noted that in Nordin's specification, RSP = x2b – (x2b if the price for all previous blocks is marginal price), is applied for decreasing block tariff. In the case of IBTs, RSP should be vice versa, RSP = (x2b if the price for all previous blocks is at marginal price) – x2b. The underlying idea is to artificially linearize the non-linear budget line. The real budget line is AB'C'. However, households can consume as they face an artificial linear line DC'. Households are technically given an additional income AB'D which is the RSP. In empirical tests, the estimated coefficient of RSP should be equal and the same sign as the beta coefficient of income.

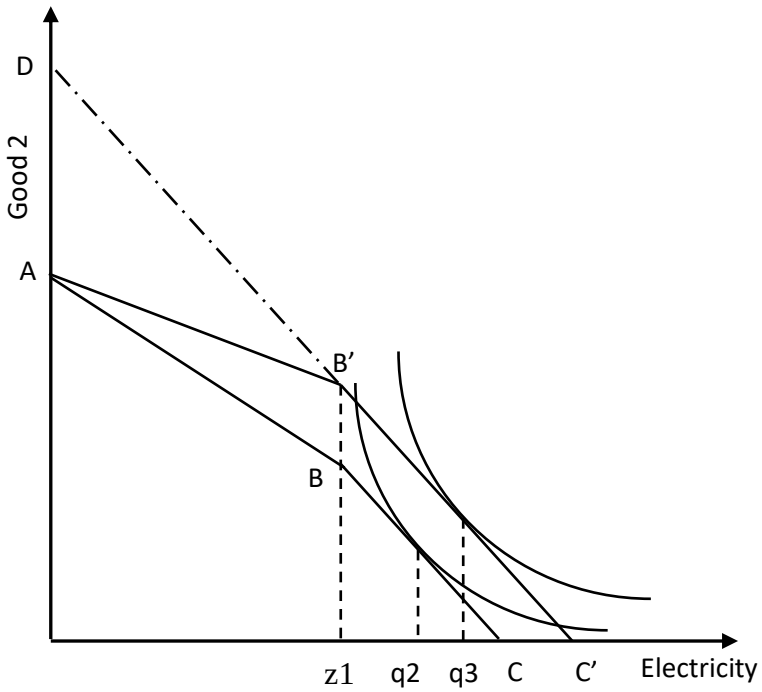


Figure 5-2. Changes of intra-marginal price and demand
Source. Author illustrated.

This approach paves the way for a series of researches applying marginal price to analyzing price elasticity in block pricing tariffs (Hausmann, Kinnucan and McFaddden, 1979; Billings and Agthe, 1980; Barnes, Gillingham and Hagemann, 1981; Henson, 1984; Agthe and Billings, 1987; Nieswiadomy and Molina, 1989).

Meanwhile, many researchers employ average price for their empirical analysis (Parti and Parti, 1980; Dubin and McFadden, 1984; Hsing, 1994; Bernard, Bolduc and Belanger, 1996; Garcia-Cerrutti, 2000; Zhou and Teng, 2013). This comes from the fact that some researches use data at aggregate level, rather than household level, and thus cannot utilize marginal prices. Some researchers justify their choice by attacking the assumption of perfect information. They claim that the structure of IBTs are so complicated that customers cannot be fully aware of changes in the structure as well as in intra-marginal prices and, thus customers will react to average prices (Foster and Beattie, 1981; Fell, Li and Paul, 2014). This fact is similar to a theory from behavioral economics developed by Liebman and Zeckhauser (2004). The key assumption is that customers have limited knowledge of the actual price, so they perceive the price in a crude and simple way. Liebman and Zeckhauser (2004) define the term "shmedule" to represent the situation. "Shmedule" means an inaccurate perceived schedule. One type of "schmedule" is ironing which "arises when an individual facing a multipart schedule perceives and responds to the average price at the point where he consumes" (Liebman and Zeckhauser, 2004).

5.2.2 Models for testing price types

The choice of price type is greatly controversial. Some researchers develop models to test the kind of prices to which customers respond (Opaluch, 1982; Shin, 1985; Chicoine and Ramamurthy, 1986; Ito, 2014). However, the results are also controversial. Shin (1985) and Ito (2014) found evidence supporting the hypothesis that households react to average prices. Chicoine and Ramamurthy (1986) developed the model of Opaluch (1982) to carry out empirical tests with water consumption of households in Illinois 1983. They find that neither marginal price nor average price are adequate to explain demand. In other words, households react to both prices. Nieswiadomy and Molina (1991) apply the Shin (1985) model and find empirical evidence that consumers react to marginal prices under increasing block tariffs, while they react to average prices under decreasing block tariffs.

5.2.2.1 Opaluch (1982) model

Opaluch starts a model with two-block pricing. The first block price is P_1 if consumers consume up to Q_1. The price of the second block is P_2 for all kWh consumed beyond Q_1. As Nordin (1976) specifies, the premium rate structure $RSP = (P_1 - P_2)Q_1$ and $MP = P_2$. Opaluch then decomposes the average price (AP) as follows.

$$AP = \frac{P_1Q_1+P_2Q_2}{Q} = \frac{P_1Q_1+P_2(Q-Q_1)}{Q} = P_2 + \frac{(P_1-P_2)Q_1}{Q} = P_2 + \frac{RSP}{Q} = MP + \frac{RSP}{Q} \qquad (5\text{-}1)$$

His basic demand function is

$$Q = \alpha_0 + \alpha_1 X + \alpha_2(Y - RSP) + \alpha_3 AP \qquad (5\text{-}2)$$

Y = household income
X = factors other than income and prices

If we substitute (5-2) to (5-1) and assume that coefficients of each components of AP can be estimated independently, we have

$$Q = \alpha_0 + \alpha_1 X + \alpha_2(Y - RSP) + \alpha_3 MP + \alpha_4 \frac{RSP}{Q} \qquad (5\text{-}3)$$

In (5-3) α_2 represents the impact of RSP as an income effect in Nordin (1976). The interesting coefficient is α_4. If $\alpha_4 = 0$ then customers fully react to marginal prices. If $\alpha_4 = \alpha_3$ then customer react to average prices.

5.2.2.2 Shin (1985) model

Shin argues that information is costly, thus customer may not react fully to marginal price. Instead, customers react to perceived price. He defines perceived price as the following.

$$P^* = MP\left({}^{AP}/_{MP}\right)^k \tag{5-4}$$

In (5-4), *k* represents price perception. (AP/MP) captures the impact of RSP on price perception. Empirically, the perceived price can be entered in a demand equation in logarithmic form as follows.

$$lnQ = \alpha_0 + \alpha_1 lnX + \alpha_2 lnY + \alpha_3 lnP^* \tag{5-5}$$

Substitute (5-4) to (5-5) we have

$$lnQ = \alpha_0 + \alpha_1 lnX + \alpha_2 lnY + \alpha_3 lnMP + (\alpha_3 k)\ln({}^{AP}/_{MP}) \tag{5-6}$$

From (5-6), we can estimate the parameter *k*. Going back to (5-4), the estimated value of *k* reveals the true price where customers react. If $k = 0$ then customers fully react to MP. If $k = 1$ then customers react to AP. If $0 < k < 1$ then customers react to a perceived price which lies between MP and AP. If $k > 1$ then the perceived price goes beyond AP. Shin (1985, p. 594) gives his justification for $k > 1$ as "It is possible, though unlikely, that confusion caused by a rapidly increasing monthly bill due to a large price increase or fuel adjustment costs may result in a perceived price greater than AP".

5.2.2.3 Ito (2014) model

Ito (2014) develops a model to find the shape of perceived prices with the aim to find out whether the perceived prices are MP, AP or expected MP. His model is as follows.

Suppose that

q_{it} = consumer *i*'s average daily electricity use during period *t*

$P(q_{it})$ = marginal price at q_{it}

q_{kit} = the kWh consumption *k* per cent away from q_{it}

$$q_{kit} = (1 + k/100)q_{it} \tag{5-7}$$

$P_{k,it}$ = marginal price at q_{kit}

$$P_{k,it} = p(q_{kit}) \tag{5-8}$$

The key assumption for the model is that customer behaviors are based on the surrounding marginal prices of q_{it}. Ito (2014, p. 554) justifies the assumption by arguing that customers may not be fully aware of the structure of block pricing or that there is uncertainty in *ex-post* consumption. Thus the perceived price is determined by weighted $P_{k,it}$. The weight is denoted by w_k. Ito (2014) assumes that *k* is 100 per cent and models the density function of w_k as the following function.

$$w_k(\delta, \pi) = \begin{cases} \delta \frac{\exp(-k.\pi_l)}{\sum_{k>0} \exp(-k.\pi_l)} \; if\ k \leq 0 \\ (1-\delta) \frac{\exp(-k.\pi_r)}{\sum_{k>0} \exp(-k.\pi_r)} \; if k > 0 \end{cases} \tag{5-9}$$

δ: the relative weight between both sides of q_{it}. For example, if $\delta = 0$, consumers do not care about the price on the left-hand side of q_{it}

π_l & π_r: the slope of density function

$w_k(\delta, \pi)$ can be estimate by the following function

$$\Delta lnq_{it} = \beta \sum_{k=-100}^{100} w_k(\delta, \pi)\, \Delta lnP_{k,it} + \theta X \qquad (5\text{-}10)$$

From the estimate $w_k(\delta, \pi)$, three weighting parameters $\delta, \pi_l\ and\ \pi_r$ can be estimated. Ito (2014, p. 556) shows that the three parameters can reveal the true shape of perceived prices.

- If $\delta = 0.5$ and $\pi_l = \pi_r$, customers react to expected marginal prices.
- If the slopes $\pi_l\ and\ \pi_r$ are steep, customers react to marginal prices.
- If $\delta = 1$, customers react to average prices.

Among the three models, the Shin (1985) model has some advantages compared to Opaluch (1982) and Ito (2014). First, the Shin model allows the incorporation of both marginal prices and average prices which the Opaluch model cannot do (Nieswiadomy and Molina, 1991). Second, the Shin model is easier to implement and interpret than the Ito model. Since this thesis is the first one to explore this problem in Vietnam, Shin's model is the most appropriate one.

5.2.3 Endogeneity in block pricing

The empirical problem of IBTs is endogeneity. Endogeneity was discovered very early in research on electricity demand (Halvorsen, 1973; Taylor, 1975; Foster and Beattie, 1981). Electricity prices are different to prices of other goods which are determined jointly by supply and demand. Electricity prices in Vietnam are set by government, thus electricity price structures are known prior to consumption. However, the final price that a household will pay depends on the quantity the household consumes. As a result, both price and quantity are still jointly determined, though price structure is known prior. Taylor (1975, p. 79) comments that a decreasing block tariff means consumers face a downward supply curve. Billing and Agthe (1980) makes the same comment that IBTs work, in effect, as an upward supply curve. Nevertheless, prices and quantities are determined at the same time.

$$\begin{cases} q = f(p) \\ p = f(q) : the\ IBTs \end{cases}$$

Therefore, the estimates of an electricity demand curve in reduced form q = f(p) would lead to biased estimated parameters (Henson, 1984; Nieswiadomy and Molina, 1989). Henson (1984, p. 149) claims that under IBTs, beta coefficients of MP and RSP would be biased toward and away from zero. The greater the number of blocks and the more steeply across blocks, the greater in absolute value the bias tends to be. In addition, the bias will not disappear asymptotically, thus OLS estimates are not only biased but also inconsistent (Henson, 1984, p. 149).

Empirical researches have employed several techniques to deal with the endogeneity. The first way is to estimate simultaneously a system equation of p and q (Agthe and Billings, 1987; Fell, Li and Paul, 2014). The second way is to use instrument variables (IVs). In the second way, some researchers have use lagged prices as IVs (Parti and Parti, 1980), some use a predetermined prices such as price at sample mean of kWh (Agthe and Billings, 1987). The lagged price IV is simple but inappropriate for Vietnam since the lag in our data is two years. The prefixed prices faces the problem of measurement errors since such prices do not reflect the marginal price from the actual schedule (Taylor, 1975; Henson, 1984).

The popular form of IV in this case is the price of predicted quantity. This form is proposed by McFadden, Puig and Kirschner (1977). They show that the IV can produce a consistent parameter. This form has been applied widely by other researches (Barnes, Gillingham and

Hagemann, 1981; Henson, 1984; Terza, 1986; Nieswiadomy and Molina, 1989, 1991). The basic idea is as follows.

- Step 1: Calculate a predetermined price vector W
- Step 2: Estimate Q=f(W, X) where X are other exogenous variables
- Step 3: Use the above model to predict $\hat{Q}$
- Step 4: Apply the IBTs on $\hat{Q}$ to find the corresponding price as the IV $\hat{P}$
- Step 5: Estimate demand function again with $\hat{P}$ as the instrument for the original prices

It should be noted that there are some alternative ways to calculate the predetermined price W. McFadden, Puig and Kirschner (1977) uses prices calculated at sample mean of kWh. Hausmann *et al.* (1979) and Terza (1986) use actual prices.

Recently, Le Viet Phu (2017) proposes a new IVs for Vietnam, using registration status, the grid connection types and payment methods as the instruments for prices. Estimated results show that the IVs are quite good with many data samples. The downside of this approach is that it requires some specific data which are not publicly available, such as grid connection types and payment methods.

In our case, based on the nature of our data, I choose to apply the approach of McFadden, Puig and Kirschner (1977).

5.3 Model specification and data

5.3.1 Short-run model

Recalling the short-run demand function (3-61) in Chapter 3, we have the general form

$$lnE = \alpha_0 + \alpha_1 lny + \alpha_2 lnP + \alpha_3 Z + \alpha_{4j} A_j D_j + \varepsilon$$

where

E = household electricity consumption
y = household (per capita) income
P = price
Z = related variables such as cdd/hdd, price of gas
A_j = vector of appliance
D_j = dummy vars = 1 if the household owns the asset *j*, = 0 other wise

We detail the general from as follows.

Income

The income form is monthly income per capita.

Prices

Two separate models are estimated, one for AP and one for MP and RSP. The beta coefficients of MP and AP are expected to be negative. The beta coefficient of RSP should be positive.

Related products

Related products include expenditure on liquefied petroleum gas (LPG) and generators. The price of LPG should be used as a proxy for prices of substitution goods. However, in Vietnam, the price of gas fluctuates frequently and there is heterogeneity across suppliers. Thus I use gas expenditure instead of gas prices with an underline assumption that LPG expenditure increases with gas prices. It is worth noting that LPG in Vietnam is only for cooking. It is not used for heating as in Western countries.

The role of generators is quite complicated. It is hard to say a generator is a complement to or a substitute for electricity. In theory, it should be a substitute. However, in Vietnam, households buy generators mainly for using in cut-off times, with buyers normally rich households. They have many electric appliances and thus need generators to keep those appliances running smoothly, even in black-out periods. So, the more electricity a household consumes, the higher the likelihood that the household owns a generator.

Appliances (Aj)

Based on data availability, we choose a set of eight electric appliances including air conditioner (AC), fan, fridge, washing machine, oven, computer, TV and water heater. Ideally, we should include induction cookers since they are also a heavy electricity consuming appliance and is an increasing trend in Vietnam. Unfortunately, in the questionnaire (section 6), LPG cookers and induction cookers are grouped in a same category (code 27). Electric rice cookers, electric cookers and gas pressure cookers are grouped in a same group too (code 28). Currently, there is no way to separate electric and non-electric appliances in these categories.

Housing condition

Type	*Dwelling type*	*N*	*mean*	*sd*
[1]	Villa/ houses with a private kitchen and bathroom	5,331	3,889.114	3,789.406
[2]	Houses with a shared kitchen or bathroom	3,338	2,304.046	2,377.297
[3]	Semi-permanent houses	16,606	2,137.282	1,698.521
[4]	Temporary and other types of house	2,176	1,331.269	1,021.172

Note. This table is based on master data before dividing into panel and pooled data sets.
Table 5-1. Income per capita at 2012 price over dwelling type
Source. Author compiled.

Housing condition variables include square meter per capita and ownership. The model has not included dwelling types, as well as other dwelling conditions, such as roofs, pillars, walls. First, dwelling type and income are highly correlated. Indeed, they measure the same thing, economic condition of a household. Pairwise comparison with Bonferroni correction shows that the differences in income per capita across dwelling types are significant at the level of 0.01. The model already has income per capita, so it is not necessary to include dwelling condition.

Dwelling type	*Contrast*
[2] vs [1]	-1397.106***
	[62.8387]
[3] vs [1]	-1661.875***
	[50.09773]
[4] vs [1]	-2391.324***
	[53.31721]
[3] vs [2]	-264.7686***
	[42.21565]
[4] vs [2]	-994.2178***
	[45.99025]
[4] vs [3]	-729.4492***
	[26.00353]

Note. Delta-method Std. Err. in brackets; * p<0.10, ** p<0.05, *** p<0.01
Table 5-2. Pairwise comparison of income per capita at 2012 price over dwelling type
Source. Author compiled.

Pairwise comparison is calculated with Bonferroni correction, the calculation is conducted via modelling income per capita on dwelling types since one-way ANOVA cannot be done due to the rejection of Bartlett's test for equal variances.

Second, one can argue that dwelling is also an asset so should be included in the model. However, dwellings are different from appliances because appliances consume electricity directly, while dwellings do not. We keep only the square meter per capita since it relates to the demand for lighting and cooling.

Demographic and education

The group includes head characteristics, household highest education, and household demographic. The household head characteristics include gender and age. Education is characterized by dummies of primary, secondary, high school and a degree. The base line for education is no education.

The demographic includes both household size and household composition as prescribed in Deaton and Paxson (1998). The household size is the total number of household members. The household composition includes the ratio of male, the ratio of children aged 15 and younger, the ratio of elders aged 60 and above. The ratio of children and elders represent the difference in the member preference with adult preference in electricity consumption. Meanwhile, the variable household size corresponds to the concept of doubling the number of household members while keeping family composition and resources constant, considering that the income and square meter area are in per capita. The coefficient of household size variable facilitates seeing whether there are "economies of scale"[1] in terms of kWh in electricity consumption.

$$\partial\left(\frac{kWh}{n}\right)\Big/_{\partial n} = \left(\frac{\partial kWh}{\partial n}n - kWh\right)\Big/_{n^2} \tag{5-11}$$

From the estimate equation we have

$$\beta_n = {}^{\partial lnkWh}/_{\partial lnn} = {}^{\partial kWh}/_{\partial n}\,{}^{n}/_{kWh} \tag{5-12}$$

Or

$$\partial kWh/_{\partial n} = \beta_n({}^{kWh}/_{n}) \tag{5-13}$$

Substitute (5-13) to (5-11), we have

$$\partial\left(\frac{kWh}{n}\right)\Big/_{\partial n} = {}^{kWh(\beta_n - 1)}/_{n^2} \tag{5-14}$$

The equation (5-14) implies that if $\beta_n > 1$, the ${}^{kWh}/_{n}$ increases if n increases, there are "diseconomies of scale". Vice versa, if $\beta_n < 1$, the ${}^{kWh}/_{n}$ decreases if n increases, there are "economies of scale".

Climate condition

[1] The term "economies of scale" is in quotes since it is originally for expenditure, but we have applied it to quantity only.

The climate condition is represented by the cooling degree days (CDD) at 25°C (cdd25). The heating degree days (HDD) is not included in the model because Vietnam is a tropical country with high level of average temperature. In fact, the heating degree days (HDD) at the base of 15°C (hdd15) is nearly zero.

Dummy variables for time and provinces

The model also includes dummy variables for time and provinces to control unobserved factors which vary across years and geographic locations.

5.3.2 Long-run model

Recalling the long-run demand function (3-62) in Chapter 3, we have the general from

$$lnE_{it} = \alpha + \beta lnp_{it} + \gamma lny_{it} + \theta Z_{it} + \delta lnE_{i,t-1} + \varepsilon_t$$

where

E_{it} = electricity consumption at period *t*

$E_{i,t-1}$ = electricity consumption at period *t-1*

p_{it} = price of electricity at period *t*

y_{it} = household (per capita) income at period *t*

Z_{it} = a vector of other factor such as price of gas, cdd

The long-run model employs the same variables as in the short-run model above with two adjustments. First, the model includes the new variable of lagged electricity consumption. Second, all asset variables are removed from the model.

5.3.3 Perceived price model

I choose to apply the Shin (1985) model (5-6) as introduced in the literature review section.

$$lnQ = \alpha_0 + \alpha_1 lnX + \alpha_2 lny + \alpha_3 lnMP + \alpha_4 \ln(^{AP}/_{MP})$$

Where *y* is income per capita, and X are a vector of all other variables correspond to the short-run and long-run model, except for income and prices.

The parameter *k* is calculated directly from α_3 and α_4. Variance of *k* is calculated by the method of large sample variance approximation in Kmenta (1997, p. 486).

$$k = {}^{\alpha_4}/_{\alpha_3} \qquad (5\text{-}15)$$

$$Var(k) = \left[\frac{\partial k}{\partial \alpha_3}\right]^2 Var(\widehat{\alpha_3}) + 2\left[\frac{\partial k}{\partial \alpha_3}\right]\left[\frac{\partial k}{\partial \alpha_4}\right] Cov(\widehat{\alpha_3}, \widehat{\alpha_4}) + \left[\frac{\partial k}{\partial \alpha_4}\right]^2 Var(\widehat{\alpha_4}) \qquad (5\text{-}16)$$

or

$$Var(k) = \left[\frac{-\alpha_4}{\alpha_3^2}\right]^2 Var(\widehat{\alpha_3}) + 2\left[\frac{-\alpha_4}{\alpha_3^2}\right]\left[\frac{1}{\alpha_3}\right] Cov(\widehat{\alpha_3}, \widehat{\alpha_4}) + \left[\frac{1}{\alpha_3}\right]^2 Var(\widehat{\alpha_4}) \qquad (5\text{-}17)$$

5.3.4 Econometric techniques

Due to the nature of the data, this chapter applies the approach developed by McFadden, Puig and Kirschner (1977) as discussed in the literature review section. For all models I estimate with both OLS and 2SLS for IVs of prices. The predetermined price for calculating price IVs is the price at sample mean kWh of a province per month. The validity of IVs is tested by Durbin-Wu-Hausman test and a weak instrument test. All models are run with household weight.

The long-run model incorporates the lag of dependent variable. Thus we have a dynamic panel which may result in a dynamic panel bias. In dynamic panel, there may be a correlation between the lag and unobserved demand shocks in error terms that results in endogeneity. Let's consider a simple model as the following.

$$y_{i,t} = \alpha_0 + \alpha_1 y_{i,t-1} + \epsilon_{i,t} \qquad (5\text{-}18)$$

Where $\epsilon_{i,t} = v_i + u_{i,t}$
v_i is the unobserved fixed effect factor.

In this case, both OLS estimators and within estimators are biased. OLS estimates of α_1 is overestimated since it includes the predictive power of v_i (Roodman, 2009, p. 103). Within estimator such as least-squares dummy variable (LSDV) or mean differencing does not remove dynamic panel bias (Bond, 2002). Taking two examples with mean differencing and first differencing estimator, the mean differencing estimator estimates the following function:

$$y_{i,t} - \frac{1}{T-1}\sum_{t=2}^{T} y_{i,t} = \alpha_0 + \alpha_1\left(y_{i,t-1} - \frac{1}{T-1}\sum_{t=2}^{T} y_{i,t}\right) + \left(\epsilon_{i,t} - \frac{1}{T-1}\sum_{t=2}^{T} \epsilon_{i,t}\right) \qquad (5\text{-}19)$$

Obviously, $\epsilon_{i,t} - \frac{1}{T-1}\sum_{t=2}^{T} \epsilon_{i,t} = u_{i,t} - \frac{1}{T-1}\sum_{t=2}^{T} u_{i,t}$ and the fixed effect term v_i is canceled out. However, the term $\frac{1}{T-1}\sum_{t=2}^{T} y_{i,t}$ contains $y_{i,t}$ which correlates with $\epsilon_{i,t}$ (or $u_{i,t}$). At the same time, the term $\frac{1}{T-1}\sum_{t=2}^{T} \epsilon_{i,t}$ contains $\epsilon_{i,t-1}$ which correlates with $y_{i,t-1}$ (Roodman, 2009, p. 103).

The first differencing estimator estimates the following function:

$$\Delta y_{i,t} = \alpha_0 + \alpha_1 \Delta y_{i,t-1} + \Delta\epsilon_{i,t} \qquad (5\text{-}20)$$

The error term $\Delta\epsilon_{i,t} = u_{i,t} - u_{i,t-1}$ no longer contains the fixed effect factor. However, the $y_{i,t-1}$ in the term $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ correlates with $u_{i,t-1}$ in $\Delta\epsilon_{i,t}$ leading to bias as the case of mean differencing.

In these cases, the estimated coefficient α_1 is underestimated. Thus a good estimate of α_1 should fall in the value range between OLS and within estimators. In addition, the absolute value of the estimates should be less than unity to ensure the condition of converging to equilibrium (Roodman, 2009, p. 103).

A series of econometric techniques have been developed to overcome the above problem (see Roodman, 2009; Cameron and Trivedi, 2010, pp. 287–298 for reviews). Anderson and Hsiao (1982) use the first difference equation and $y_{i,t-2}$ as the instrument for $\Delta y_{i,t-1}$. Kiviet (1995) uses LSDV with bias correction. Arellano and Bond (1991) uses difference generalized method of moments (GMM) estimator. Arellano and Bover (1995) uses forward orthogonal deviations transformation. Blundell and Bond (1998) uses the system of GMM.

In this study, I choose to apply the Blundell and Bond (1998) technic. First, Blundell and Bond (1998) argue that their system GMM is more efficient and stable than the difference GMM of Arellano and Bond (1991). Second, the Blundell and Bond (1998) technic allows to include time invariance regressors since it employs level equation. However, it should be noted that fixed effect dummy variables should not be included explicitly since they can cause bias especially when T is small (Roodman, 2009, p. 115). Third, there are at least two researchers

applying this technic to estimate long-run electricity demand (Alberini and Filippini, 2011; Alberini, Gans and Velez-Lopez, 2011).

Blundell and Bond (1998) system GMM estimators based on an addition assumption that differences of IVs are orthogonal to the fixed effects in error terms. They design a system of two equations. The first is the original level equation and the second is the transformed difference equation. The two equations are identical in terms of relationship and coefficients. The two equations are estimated with a self-stacked data set constructed from the original data set with two observations for each individual.

In this study, the Blundell and Bond (1998) system GMM estimators is conducted with an user-written command xtabond2 developed by Roodman (2009).

5.3.5 Data

The short-run demand function is estimated with the pooled data set. The long-run demand function is estimated with the panel data set. The pooled data set is used for short-run demand function for two reasons. First, the pooled data set ensures the condition that appliances are fixed in short-run since there is no household appears twice in the pooled data set. Second, the pooled data set has the advantage of price variation since the data spans over three years with six different price schedules.

> *It is critical that all customers do not face an identical rate schedule. A typical example might be 100 customers from each of 10 schedules, where ideally the differing rate schedules result from customers residing in similar towns within a single geographic region, but serviced by different utility companies. This is clearly different from having variation in rate schedules only because some customers own all electric homes, or for some other reason which will likely result in a shift in the individual's demand function. This sort of variation in rate schedules is critical for estimation, and in particular, for addressing the simultaneity issue.* (Opaluch, 1984, p. 417)

For the short-run function estimation, all monetary terms are adjusted to 2012 price. Meanwhile, monetary terms in the long-run estimation are adjusted to 2014 price since we lose the data of 2012 in calculating the difference of instrument variables.

5.4 Results and discussion

5.4.1 Short-run demand

	MP model		*AP model*	
	OLS	*IV*	*OLS*	*IV*
lnmp1	1.9324***	-1.5309***		
	[0.0449]	[0.2086]		
lnrsp1	0.0100***	0.0041***		
	[0.0004]	[0.0008]		
lnap1			4.7069***	-2.0321***
			[0.2174]	[0.2293]
Income				
lnincpc1	0.1363***	0.2409***	0.1630***	0.2414***
	[0.0095]	[0.0145]	[0.0112]	[0.0144]
Household head				
male	-0.006	-0.0126	0.0049	-0.0148
	[0.0069]	[0.0120]	[0.0076]	[0.0119]

	MP model		*AP model*	
	OLS	*IV*	*OLS*	*IV*
age	0.0003	0.0016***	0.0002	0.0015***
	[0.0003]	[0.0004]	[0.0003]	[0.0004]
Demographics				
lnhhsize	0.2163***	0.5498***	0.2337***	0.5428***
	[0.0097]	[0.0331]	[0.0180]	[0.0300]
lnsexratio	0.0044***	0.0044***	0.0064***	0.0048***
	[0.0009]	[0.0011]	[0.0009]	[0.0011]
lnchildrenratio	0.0003	0.001	0.0005*	0.001
	[0.0003]	[0.0007]	[0.0003]	[0.0006]
lnelderratio	-0.0001	0	-0.0003	0.0001
	[0.0003]	[0.0006]	[0.0004]	[0.0006]
Household highest degree (base is no education)				
Primary	0.0752***	0.0653***	0.1143***	0.0678***
	[0.0177]	[0.0171]	[0.0203]	[0.0176]
Secondary	0.1165***	0.1210***	0.1707***	0.1221***
	[0.0189]	[0.0208]	[0.0240]	[0.0222]
High school	0.1333***	0.1439***	0.2019***	0.1451***
	[0.0191]	[0.0211]	[0.0268]	[0.0222]
A degree	0.1059***	0.1222***	0.1640***	0.1197***
	[0.0203]	[0.0251]	[0.0245]	[0.0265]
Housing condition				
lnpcsqm	0.0517***	0.1359***	0.0516***	0.1362***
	[0.0051]	[0.0090]	[0.0078]	[0.0095]
ownership	-0.0366***	-0.0732**	-0.0605***	-0.0728**
	[0.0123]	[0.0313]	[0.0155]	[0.0290]
Appliances				
aircon	0.0494***	0.3167***	-0.1041***	0.3172***
	[0.0115]	[0.0241]	[0.0110]	[0.0216]
fan	0.1141***	0.1014***	0.1441***	0.1103***
	[0.0194]	[0.0269]	[0.0217]	[0.0277]
fridge	0.2056***	0.5485***	0.4364***	0.5377***
	[0.0131]	[0.0270]	[0.0130]	[0.0234]
washing	-0.0110*	0.1596***	0.0169*	0.1376***
	[0.0065]	[0.0164]	[0.0088]	[0.0124]
oven	0.0011	0.0806***	-0.1069***	0.0931***
	[0.0102]	[0.0123]	[0.0104]	[0.0128]
computer	-0.0135	0.1153***	-0.0208***	0.1048***
	[0.0087]	[0.0138]	[0.0062]	[0.0132]
tv	0.2465***	0.2629***	0.2911***	0.2734***
	[0.0200]	[0.0233]	[0.0220]	[0.0228]
waterheater	0.0147	0.1085***	0.0485***	0.1011***
	[0.0092]	[0.0232]	[0.0119]	[0.0198]
Related goods				
generator	-0.0393**	0.0806***	-0.0287	0.0651**
	[0.0154]	[0.0305]	[0.0226]	[0.0308]
lpg1	0.0002***	0.0007***	0.0003***	0.0007***
	[0.0000]	[0.0002]	[0.0001]	[0.0001]
Temperature				

	MP model		*AP model*	
	OLS	*IV*	*OLS*	*IV*
cdd25	0.0004***	0.0012***	0.0004***	0.0011***
	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Admin				
urban	0.0556***	0.1590***	0.0752***	0.1534***
	[0.0072]	[0.0155]	[0.0089]	[0.0136]
year_2014	-0.0379***	0.0881***	-0.0440***	0.0851***
	[0.0092]	[0.0152]	[0.0089]	[0.0131]
year_2016	-0.1004***	0.1701***	-0.0752***	0.2337***
	[0.0135]	[0.0208]	[0.0096]	[0.0176]
Cons	1.8187***	1.5204***	0.2655	1.4948***
	[0.0791]	[0.1102]	[0.1790]	[0.1080]
N	18971	18971	18971	18971
R-sq	0.856	0.604	0.816	0.621
adj. R-sq	0.855	0.602	0.815	0.62

Note. Std. Err. in brackets; * p<0.10, ** p<0.05, *** p<0.01
Estimates of province dummy variables are excluded.

Table 5-3. Estimated results of the short-run demand function

Source. Author estimated.

5.4.1.1 Endogeneity tests

At first glance we can see that the estimated parameters in OLS models are biased and have an opposite sign to the theory prediction. First, in both OLS models, the beta coefficients of MP and AP are positive, meaning that when prices increase, the consumed quantity increases. Second, in both models, the beta coefficient signs of appliances are inverted too. For example, in the AP model, the presence of ACs, ovens or computers in a household would reduce the electricity that the household consumes. Similar problems appear in the MP model with the beta coefficient of washing machines. Also, in the MP model, the coefficients of ovens and computers are insignificant, meaning that owning an oven or a computer has no effect on kWh consumption.

	MP model		*AP model*	
Durbin-Wu-Hausman test	F	p-value	F	p-value
Model with weight and cluster for provinces	593.265	0.0000	250.27	0.0000
Model without weight and cluster for provinces	686.215	0.0000	3233.97	0.0000

Table 5-4. Endogeneity tests for the short-run demand function

Source. Author estimated.

The IV models substantially improve the estimate results. All beta coefficients are in expected signs. The coefficients of prices are negative. The coefficients of appliances are all positive and significance at 0.01 level.

The Durbin-Wu-Hausman test is conducted to test for endogeneity. The Durbin-Wu-Hausman test runs an OLS regression of the short-run model with an additional variable v_i. v_i is the fitted residual from the first-stage regression. Under the null hypothesis that price is exogenous, the coefficient of v_i would be zero (Cameron and Trivedi, 2010, p. 184). Table 5-4 shows the robustified Durbin-Wu-Hausman test which confirms that there is endogeneity in both model MP and model AP.

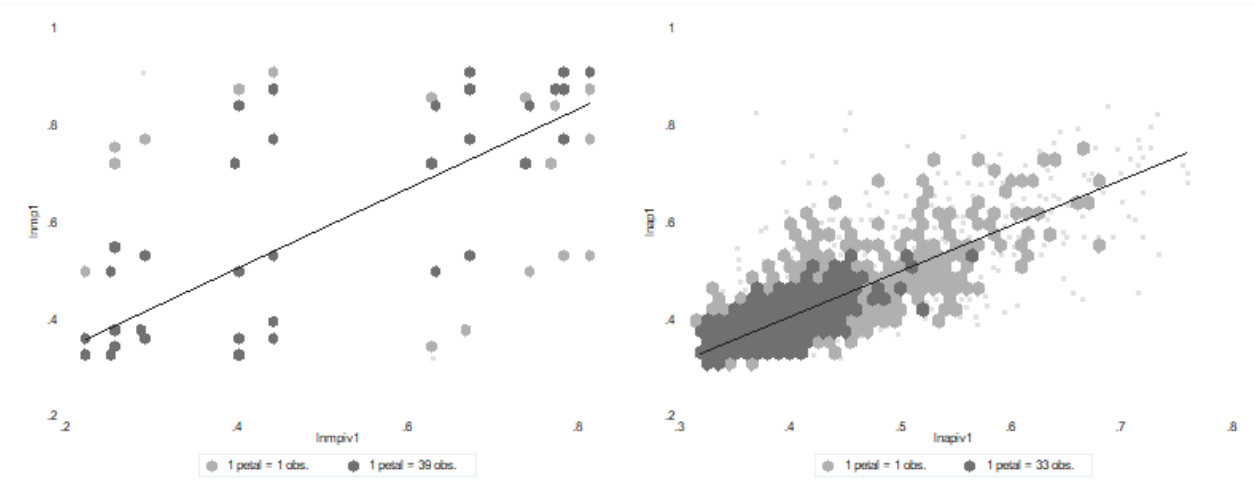
5.4.1.2 Weak instrument tests

The formal test for weak instrument was developed by Stock and Yogo (2005). The test statistic is the Cragg and Donald minimum eigenvalue (Cragg and Donald, 1993). In our case with just-identified model, Stock and Yogo (2005) determined the critical value based on the criterion of Walt test distortion size. The Walt test is a joint significance test of endogenous variables in structural model at the level of 0.05 (Cameron and Trivedi, 2010). Addressing the concern that weak instruments can cause the distortion of Wald tests in a finite sample, Stock and Yogo (2005) calculate the critical value for weak instrument test based on the tolerance for the distortion. Below is the Cragg and Donald eigenvalue and corresponding Stock and Yogo critical value.

	Cragg and Donald eigenvalue	*Stock and Yogo critical value at the distortion size of*			
		5%	10%	20%	30%
MP model	358.793	7.03	4.58	3.95	3.63
AP model	2653.67	16.38	8.96	6.66	5.53

Table 5-5. Weak instrument tests for the short-run demand function
Source. Author estimated.

In Table 5-5, the eigenvalues are much greater than the critical values (7.03 and 16.38) even if we do not tolerance a distortion greater than 0.05. Thus the null hypothesis that instruments are weak is rejected.



Note. The graph is generated without weight.
Figure 5-3. Scatter density plot with fitted line for prices and price IVs
Source. Author compiled.

Unfortunately, Stata can only conduct the Stock and Yogo test with the model without weight and cluster. Thus other tests should be considered to test the weak instruments with weight and cluster. The core idea here is to test whether instrument variables are highly correlated with endogenous variables. A weak correlation between instrument variables and endogenous variables can be an evidence of weak instrument variables.

The first attempt is to look at the density scatter plot of each endogenous variable and their corresponding IV as well as the correlation between them. The density scatter plots show strong correlations between MP, AP and their instrument variables. The correlation coefficients in

Table 5-6 confirm the close relationship. All correlation coefficients are greater than 0.7 and significant at 0.01. The initial result suggests that the instrument variables are not weak.

In the second step, we take a closer look at the relationship between the endogenous regressors and their instrumental variables by considering the role of exogenous variables. If the instrument variables, i.e lnrspiv and lnmpiv, has small explanation power on endogenous variables after controlling for exogenous variables, the instruments are weak.

	lnmp1	*lnmpiv1*	*lnrsp1*	*lnrspiv1*	*lnap1*	*lnapiv1*
lnmp1	1					
lnmpiv1	0.7835*	1				
lnrsp1	0.7168*	0.5978*	1			
lnrspiv1	0.5946*	0.6964*	0.7231*	1		
lnap1	0.9143*	0.7742*	0.5773*	0.5263*	1	
lnapiv1	0.7617*	0.9183*	0.5257*	0.5745*	0.8238*	1

Note. With weight, *: significant at the level of 0.01,
The description of the variables is in the Table B-1, Appendix B.
Table 5-6. Pairwise correlation between prices and price IVs
Source. Author compiled.

Shea's adjusted partial R-square and F statistic from the first stage regression are employed for the diagnostic. The Shea's adjusted partial R-square shows the explanatory power of instrument variables in explaining the endogenous variables after controlling for exogenous variables. That is the R2 from OLS of (i) residual from OLS of endogenous variables on exogenous variables on (ii) residual from OLS of exogenous variables on instrument variables (Cameron and Trivedi, 2010, p. 190). The F statistic is the test for the joint significance of the instruments in the first-stage regression of endogenous variables on instruments and exogenous variables. The critical value for the F-statistic is 10 as suggested in Staiger and Stock (1997). If F-statistic is smaller than 10, the instruments are weak.

		Shea's Adj. Partial R-sq.	*F statistics*	*Prob>F*
Model with weight and cluster for provinces				
Model MP	lnmp1	0.0504	233.408	0.0000
	lnrsp1	0.0989	257.606	0.0000
Model AP	lnap1	0.1271	2188.81	0.0000
Model without weight and cluster for provinces				
Model MP	lnmp1	0.0472	395.199	0.0000
	lnrsp1	0.0932	772.858	0.0000
Model AP	lnap1	0.1191	2653.67	0.0000

Table 5-7. Weak instrument tests from first stage regression for the short-run function
Source. Author compiled.

Table 5-7 shows the Shea's adjusted partial R-square and F statistic from the first stage regression. In the table, the Shea's R-square are not too low and the F statistic is much larger than the critical value of 10 as suggested in Staiger and Stock (1997). Thus we can conclude that lnmpiv1, lnrspiv1 and lnapiv1 are not weak instruments.

5.4.1.3 Sensitivity analysis

The sensitivity analysis for the short-run model are conducted in various sub-samples. The list of sub-samples is as follows.

(i) Five largest municipalities including: Ha Noi, Hai Phong, Da Nang, Can Tho and Ho Chi Minh City
(ii) Urban/ Rural areas
(iii) Six social-economics regions
(iv) The top and bottom 20 per cent of income quantile

Sample	*AP*	*MP*	*RSP*	*incpc1*	*w*
Five largest municipalities	-1.1866***	-0.8016**	0.0033***	3,764.232	0.0317
	[0.1858]	[0.3172]	[0.0010]		
Urban/Rural					
Urban	-0.3285	-0.2768	0.0041***	3,603.511	0.0302
	[0.2739]	[0.3063]	[0.0013]		
Rural	-2.8221***	-2.0319***	0.0042***	2,064.453	0.0264
	[0.4123]	[0.3129]	[0.0011]		
Six social-economics regions					
Red River delta	-1.3433***	-0.8823***	0.0045***	2,848.494	0.0309
	[0.1963]	[0.2420]	[0.0010]		
Midlands and Nth. Mountains	-3.4461***	-3.0068***	0.0042*	1,867.602	0.0253
	[0.5754]	[1.0994]	[0.0024]		
Nth. and Coastal Central	-2.9032***	-1.8798***	0.0049***	2,005.663	0.0267
	[0.3443]	[0.4862]	[0.0015]		
Central Highlands	-1.7595*	-1.4702**	0.0070***	2,246.280	0.0224
	[0.9332]	[0.7259]	[0.0018]		
Southeast	-0.8031***	-2.0774**	-0.0025	3,749.346	0.0293
	[0.2906]	[1.0097]	[0.0023]		
Mekong delta	-2.0823***	-2.0596***	0.0043	2,280.295	0.0262
	[0.4845]	[0.4644]	[0.0032]		
Income quantile					
Top 20%	0.2923	0.0826	0.0037**	5,502.895	0.0225
	[0.3907]	[0.3691]	[0.0017]		
Bottom 20%	-1.8291**	-1.8665***	0.0078***	671.3481	0.0349
	[0.8772]	[0.6076]	[0.0024]		

Note. * p<0.10, ** p<0.05, *** p<0.01; Standard errors in brackets
w: the share of electricity bills in total household income (at mean)
incpc1: monthly income per cap at 2012 price (at mean)

Table 5-8. Sensitivity analysis for different sub-samples

Source. Author estimated.

The estimates from the sub-samples confirm the stable of the model with price IVs to data. For all sub-samples, though the size of beta coefficients is different due to the different data, there is no major contradiction in signs and significances of beta coefficients with the national model. The price elasticity of all sub-sample models is listed in Table 5-8.

First, the signs and significance of all beta coefficients are in line with theory prediction. The estimated coefficients of AP and MP are all negative. The estimated parameter of RSP are positive except for the case of Southeast region. However, the estimates parameter of RSP in Southeast is insignificant at the level of 0.1, thus the results are still valid.

Second, richer samples have inelastic demand and poorer samples have elastic demand which are also in line with theory prediction. All the poorer samples including the rural area, Midlands and Northern Mountains, Central Highlands and Northern and Coastal Central have inelastic demand. Meanwhile, the five largest municipalities, the Red River delta, have the absolute values of the elasticity less than 1.0. The elasticity in urban area and top 20 per cent of income quantile are quite small and insignificant at the level of 0.1. This means that electricity demand does not respond to prices in the two samples. This is reasonable if we consider that the top 20 per cent income quantile is too rich to consider the prices when electricity bills account only for three per cent of their income. The urban area also has rich households. In addition, due to a lifestyle of night life, their demand for electricity is higher than in rural areas.

The interesting case here is the Southeast region. The Southeast region is a rich region with a monthly income per capita in 2012 prices of 3,749 thousand VND. However, its demand elasticity is in opposite pattern with other rich regions. While in the Red River delta, demand is inelastic in MP model and elastic in AP model, the Southeast region is the only region which has elastic demand in MP model and inelastic in AP model. This flags an idea on the type of price that the two areas react to. We can form a hypothesis that the Red River delta will react to MP while the Southeast will react to AP.

5.4.1.4 Estimated coefficients

❖ *Price elasticity*

In the model, the estimated parameters of MP and AP are -1.53 and -2.03 respectively. Both the parameters are significant at the 0.01 level. This means that if MP increases 1.0 per cent, the kWh consumed by households decreases 1.53 per cent, *ceteris paribus*. If AP increases 1.0 per cent, the kWh consumed by households decreases 2.03 per cent, *ceteris paribus*.

The above result means that the AP elasticity is larger than MP elasticity [2]. This is reasonable if we consider the following argument from Billing and Agthe (1980).

> *The use of average price (AP) as the only price variable in the estimating equation tends to produce excessively large estimates of the price elasticity of demand when block rates occur in the rate schedule. This problem is particularly severe when marginal price (MP) increases while the intramarginal rates remain constant. In this case the change in average price (AP) is smaller than the change in MP since AP includes intramarginal units of water which have not changed in price. Since fewer units of water will be purchased at the new, higher MP (assuming elasticity of demand is not zero), it is possible for AP to remain constant or even decrease in size when MP increases. For example, suppose a consumer purchases 20 units of water at an intramarginal rate of 30¢ per unit and 10 units at a marginal price of 60¢, and decreases water use to a total of 25 units when MP increases to 80¢. Before the change the consumer was paying (20 x 30¢) + (10 x 60¢) = 1200¢ for 30 units of water, giving AP = 40¢. After the change, the water bill is (20 x 30¢) + (5 x 80¢) = 1000¢ for 25 units of water, giving AP = 40¢. In this case, the use of AP would predict no change in water use, when in fact a price change caused sales to change. If Q falls by more than 5 units, AP would actually fall in response to a rise in MP! If instead, the consumer's purchases decrease to 26 units when MP increases to 80¢, the new water bill would be (20 x 30¢) + (6 x 80¢) = 1080¢, giving AP = 41.54¢. The calculated arc elasticity using AP is (AQ/AP).(P/Q) = (-4/1.54)x(40.77/28) = -3.71. In contrast, the elasticity calculated using MP is (-4/20) x (70/28) = -.50. It is thus clear that the use of average price in the estimating equation*

[2] The result is stable in both estimation with weight and without weight. The estimation without weight shows that the estimated parameters of MP and AP are -1.83 and -2.33.

as the single price variable will tend to produce incorrect results. (Billings and Agthe, 1980, pp. 73–74)

Interestingly, though this study employs different methods and data, the estimated parameter of MP is similar to the parameter estimated in Le Viet Phu (2017). Le Viet Phu (2017) finds that the elasticity of MP is -1.408. However, there is a large difference between the estimated coefficients of AP in my study and in Le Viet Phu (2017). In my study, the elasticity of AP is 2.024 larger than the elasticity of MP. In Le Viet Phu (2017), the elasticity of AP is 0.972 and smaller than the elasticity of MP. In his paper, Le Viet Phu (2017) does not justify explicitly the reasons why his AP elasticity is smaller than the MP elasticity. From our point of view, the explanation may come from the fact that his data includes both households with increasing block tariffs and household with flat tariffs in which the latter accounts for a significant share.

The estimated parameters of AP and MP also show that electricity demand in Vietnam is elastic. This result contradicts the findings of other researchers in which electricity demand in other countries is inelastic in terms of both AP and MP (see Zhou and Teng, 2013, p. 396 for detail). However, the contradictions are reasonable if we consider that Vietnam is a low middle-income country. Meanwhile, most of the researches which are reviewed in Zhou and Teng (2013) are for high-income countries such as Greece, Singapore, Israel, Taiwan, Australia and Korea. The table from sensitivity analysis (Table 5-8) partly shows that richer areas in Vietnam also have inelastic demand, but the population size of the rich areas is quite small compared to the whole national sample. In addition, the results are also in line with the conclusion from Frondel and Kussel (2018) that "only those households that are informed about prices are sensitive to price changes, whereas the electricity demand of uninformed households is entirely price-inelastic". Considering the fact of "perfect information" on electricity prices in Vietnam (refer to Section 5.4.3 for the detailed discussion), it is reasonable for the residential electricity demand to be elastic.

Nevertheless, we can conclude that electricity demand in Vietnam is elastic. However, it is worth noting that when we interpret price elasticity, it is only valid locally because if the price change is large enough, consumers will switch blocks causing discontinuities in demand (Taylor, Gail and Philip, 1977).

Except for prices, the estimated parameters of other variables are stable in both models. In the next part we consider the estimated parameters of the remaining variables

❖ *Rate Premium Structure*

The estimated parameter of RSP is 0.0041 and significant at the level of 0.01. This means if intra-marginal prices change, the marginal prices do not change, customers do not switch block, for any 1.0 per cent increase from RSP caused by intra-marginal price change, households increase their kWh consumption 0.0041 per cent.

In term of theory prediction, the result only matches half the prediction. The estimated parameters of RSP have the same sign with the coefficient of income. However, the magnitude of the estimated coefficient of RSP is much smaller than the estimated coefficient of income which is not in line with prediction.

In term of empirical tests, the above result is not something unexpected. In fact, a lot of researchers find the same situation where the sign of the RSP's beta coefficient is correct as predicted but the magnitude is not (see Arbués, Villanúa and Barberán, 2010, p. 87 for reviews). In explaining this result, previous researchers have argued that the size of RSP is quite small relative to income and thus has a small impact on customer perceptions of income (Billings and

Agthe, 1980; Nieswiadomy and Molina, 1989). The reason appears to be true in our sample. The RSP accounts for only 0.34 per cent of household income which is too small to be noted.

❖ *Income elasticity*

The estimated coefficient of income variable is 0.24. This means that if a household's income per capita in 2012 prices increases 1.0 per cent, the electricity quantity demanded increases 0.24 per cent. This confirms the fact that electricity is a type of necessary good for households. This result is a bit larger than Le Viet Phu (2017), at about 0.07, and similar to the income elasticity in China during the period 2007–2009 at the range of 0.14–0.34 (Zhou and Teng, 2013).

❖ *Demographic*

For the household size, the estimated coefficients of variable lnhhsize is 0.54 and significant at the level of 0.01 at both AP and MP models. This means that if the size of a household doubles, its electricity demand increases 54 per cent, *ceteris paribus*. The estimated parameter is less than unity implying that households do enjoy economies of scale in electricity use. Intuitively, if household size doubles while the kWh consumption increases less than double, the kWh consumption per capita should decrease. In general, the existence of economies of scale in electricity use is similar to other researches such as Ironmonger *et al.* (1995), Filippini and Pachauri (2004), Brounen *et al.* (2012), Salari and Javid (2017).

For the case of Vietnam, this result is similar to the estimates of Le Viet Phu (2017). Le Viet Phu (2017) estimates the beta coefficient of hhsize is 0.1462. However, the beta coefficient is semi-elasticity since his function form is lnkWh = f(hhsize). My estimates can be transformed to semi-elasticity as the following.

$$\partial \ln kWh1 \,/_{\partial \ln hhsize} = \left(\partial \ln kWh1 \,/_{\partial hhsize}\right) hhsize \qquad (5\text{-}21)$$

Or

$$\partial \ln kWh1 \,/_{\partial hhsize} = \left(\partial \ln kWh1 \,/_{\partial \ln hhsize}\right)\left({}^{1}/_{hhsize}\right) \qquad (5\text{-}22)$$

Where

$\partial \ln kWh1 \,/_{\partial \ln hhsize}$ = elasticity of kWh1 with respect to household size

hhsize = household size calculated at sample mean.

My elasticity is 0.54. The household size at sample mean is 3.7. Therefore, the semi elasticity is 0.54/3.7 = 0.1459 which is approximately the same Le Viet Phu (2017) estimates.

For the household composition, interestingly, the coefficients of children ratio and elder ratio is nearly zero and insignificant at the level of 0.1. There are no significant differences in kWh consumption between an adult and a child or an elder. This result is somewhat similar to Poulsen and Forrest (1988) which uses data at district level to regress electricity use in Sydney on variables including family structure which is the fraction of nuclear family in a district. They find that family structure is insignificant in determining the level of electricity use.

However, this result contradicts many other researches such as Tonn and Eisenberg (2007), Brounen *et al.* (2012). Tonn and Eisenberg (2007) find that elders use higher level of energy than youngers. Brounen *et al.* (2012) analyze data of 300,000 Dutch households in 2008–09 and find that elder households consume two to four per cent less electricity than two married

adults. They explain that elders use more time at home but use less energy-consuming appliances. In addition, they also find that in comparison to electricity consumption of middle-aged married couples, families with children have lower per capita electricity consumption in kWh. Families with children less than four years old have higher per capita consumption while families with children from five to 12 and above 12-year-olds have increasingly higher per capita consumption than middle-aged married couples. They explained this as the "Nintendo-effect" where older children intensively use TV, gaming devices and personal computers.

In our case, the no significant difference in demand between an adult and an elderly person may come from the fact that elderly people have a stronger saving attitude. In this case, the saving attitude obviates any increase in electricity consumption that incurs from their longer time spent in their home. The saving attitude may come from two sources. First, elderly people who were over 60 in 2014 had passed two wars in Vietnam when living standards were extremely low. Thus a saving attitude is built into their daily activities. Second, elderly people retire at 60 and their retired salary is considerably lower that their income at work. They need to save money to cope with unexpected events in their retired lives.

The no significant difference in demand between an adult and a child may come from two facts. First, the "Nintendo-effect" does not work in Vietnam. Children in Vietnam also play electronic games intensively. However, instead of playing at home as in European countries, they go to gaming centers which are popular in Vietnam. Their electricity expenditure for gaming is thus not included in household electricity bills. Sanquist *et al.* (2012) investigated lifestyle factors in US residential electricity consumption. They identified five lifestyle factors associated with AC, laundry usage, personal computer usage, climate zone of residence and TV use. Among the five factors, it seems that personal computer usage is the factor that differ significantly between a child and an adult in Vietnam. If children go out to play electronic games, there should be no difference in electricity demand between a child and an adult. Second, among the five lifestyle factors, AC is also another key factor causing differences in electricity demand between a child and an adult. However, in Vietnam during hot days, many households apply a saving strategy by gathering all members in a main room and sharing an air conditioner. Thus there should be no difference between a child and an adult in electricity consumption.

❖ *Housing condition*

The estimated parameters of housing condition are in line with theory prediction. The estimated coefficient of square meter per capita is 0.136 and significant at the 0.01 level meaning that if the living area per capita double, households increase their kWh consumption 13.6 per cent. This result is similar to Kavousian *et al.* (2013). Kavousian *et al.* (2013) argues that a larger dwelling requires higher energy consumption for cooling/heating due to the larger volume to condition and higher loss to the outside.

The estimates coefficient of ownership is -0.073 and significant at the level of 0.01. If householders own the dwelling, their kWh consumption is 7.3 per cent lower than for householders who do not own the dwelling. This may appear strange at first glance, since people who rent houses normally are not wealthy people. However, in Vietnam, people who rent a house normally live with landlords thus, must pay highest block prices. In that case, the beta coefficient of the rent variable is reasonable.

❖ *Related goods*

The estimates confirm that LPG are a substitute good for electricity. The estimated coefficient of LPG expenditure is positive (0.0007) and significant at the 0.05 level. If expenditure for LPGs increases 100 thousand VND, the kWh consumption increases seven per cent. Indeed, the positive sign of the coefficient is in line with prediction. Households in Vietnam use LPG

mainly for cooking purpose. In short-run, the demand for LPG should be inelastic due to the fixed assets of gas cookers. Thus if price of LPG increases, quantity consumed may decrease a bit, but LPGs expenditure must increase. At the same time, due to the decrease in LPG quantity consumed, households may switch to using electricity. In this case, LPGs expenditure and kWh consumption move in the same direction.

On the other hand, the estimated parameter of generators shows that generators are not a substitute for electricity in Vietnam. The estimate coefficient is 0.08 in MP model and 0.065 in AP model. In both models, the estimates coefficient is positive and significant at the level of 0.05. This means a household with a generator uses electricity 6.5–8.0 per cent higher than a household without a generator. As discussed before in model specification, in Vietnam, households buy generator mainly for using in cut-off periods. Buyers are normally rich households. They have many electric appliances and thus need generators to keep these appliances running smoothly, especially in black-out times. In this way, generators seem to be complement to electricity.

❖ *Appliances*

The estimated coefficients of all appliances are positive and significant at 0.01 level. In addition, the estimates show that all appliances have strong impacts on kWh consumption of a household. A household equipped with any appliances has kWh consumption at least 10 per cent higher than a household without appliances. A household with the full set of appliances consumes 166 per cent higher kWh than a household without all the appliances.

Among those appliances, fridges are the most electricity consuming appliance. A household with a fridge consumes 54 per cent higher kWh than a household without a fridge. Following fridges, appliances in decreasing order of kWh consumption are ACs, TVs, washing machines, computers, fans, ovens and water heaters.

These results are in line with other researches such as Zhou and Teng (2013) or Le Viet Phu (2017). The fact that a fridge is the most consuming appliance is understandable since fridges are run continuously, regardless of time of day, weather or seasons.

❖ *Temperature and urban*

The estimated coefficient of cdd25 is 0.0012 and significant at the level of 0.01. If CDD at the base of 25^0C increases 30 degree-days (1 degree-day for each day), household kWh consumption increases 3.6 per cent, *ceteris paribus*.

The standardized coefficient of urban is 0.15 and significant at the level of 0.01. A household in an urban area consumes electricity 15 per cent higher than the same household living in a rural area.

5.4.2 Long-run demand

Before discussing results, let's recall two major points. First, the consistent estimates for the lag in electricity consumption should be in the range of OLS and LSDV estimates. Second, the sign of the price coefficients should be negative.

The *xtivreg* model simply ignores the dynamic bias. The model uses price IVs as in short-run case and takes difference to remove fixed effect. In this model, beta coefficients of prices are negative though insignificant at the 0.1 level. However, the estimated coefficient of the lag electricity consumption is far away from consistent. In both models of MP and AP, the

estimated coefficients of lag electricity consumption are out of the range defined by OLS and LSDV estimates. Standard errors of those coefficients are also large.

	OLS	*LSDV*	*Panel IV*	*BLUNDELL & BOND*
L.lnkwh1	0.1575***	-0.1103***	-0.2949***	0.0616**
	[0.0110]	[0.0219]	[0.0634]	[0.0291]
lnincpc2	0.0776***	0.0656***	0.1106***	0.0598***
	[0.0089]	[0.0221]	[0.0364]	[0.0107]
lnmp2	2.3253***	2.1986***	-0.7698	2.6250***
	[0.0438]	[0.0862]	[0.8099]	[0.0907]
lnrsp2	0.0159***	0.0132***	-0.0011	0.0202***
	[0.0004]	[0.0008]	[0.0066]	[0.0009]
lnhhsize	0.1192***	0.1754***	0.5928***	0.0925***
	[0.0178]	[0.0521]	[0.1489]	[0.0205]
lnsexratio	0.0027	0.0025	0.0045	0.0030*
	[0.0017]	[0.0037]	[0.0052]	[0.0016]
lnchildrenratio	-0.0001	-0.0014	-0.0024	-0.0007
	[0.0006]	[0.0016]	[0.0025]	[0.0006]
lnelderratio	-0.0005	-0.0008	-0.0007	-0.0006
	[0.0006]	[0.0019]	[0.0026]	[0.0007]
lnpcsqm	0.0242***	0.0533*	0.1774***	0.0203*
	[0.0091]	[0.0277]	[0.0537]	[0.0108]
ownership	0.0248	0.0709	0.1317	0.0324
	[0.0345]	[0.0772]	[0.0861]	[0.0367]
lpg2	0.0002***	0.0001	0.0004**	0.0001
	[0.0001]	[0.0001]	[0.0002]	[0.0001]
cdd25	0.0004***	0.0004**	0.0011***	0.0003**
	[0.0001]	[0.0002]	[0.0003]	[0.0001]
urban	-0.0083	-0.222	-0.0241	-0.0008
	[0.0098]	[0.2447]	[0.3024]	[0.0129]
N	3646	3646	3626	3646
R-sq	0.916	0.971	0.048	

Note. * p<0.10, ** p<0.05, *** p<0.01; Standard errors in brackets
L.lnkwh1: the lag of dependent variable

Table 5-9. Estimated long-run function of MP model

Source. Author estimated.

	OLS	*LSDV*	*Panel IV*	*BLUNDELL & BOND*
L.lnkwh1	0.3138***	-0.1563***	-0.2624***	0.2488***
	[0.0158]	[0.0267]	[0.0271]	[0.0349]
lnincpc2	0.1561***	0.1107***	0.1057***	0.1842***
	[0.0142]	[0.0316]	[0.0275]	[0.0183]
lnap2	4.3761***	5.3752***	-0.1339	4.2118***
	[0.1353]	[0.2513]	[0.5511]	[0.1793]
lnhhsize	0.1674***	0.1950***	0.5101***	0.2242***
	[0.0266]	[0.0665]	[0.0719]	[0.0283]
lnsexratio	0.0056**	0	0.0039	0.0073***
	[0.0028]	[0.0051]	[0.0045]	[0.0023]
lnchildrenratio	0.0008	-0.0014	-0.0022	0.0003
	[0.0009]	[0.0020]	[0.0021]	[0.0009]
lnelderratio	-0.001	-0.0024	-0.0009	-0.0011
	[0.0009]	[0.0019]	[0.0022]	[0.0008]
lnpcsqm	0.0230*	0.0351	0.1519***	0.0485***
	[0.0131]	[0.0317]	[0.0352]	[0.0142]

	OLS	*LSDV*	*Panel IV*	*BLUNDELL & BOND*
ownership	0.0214	0.0275	0.1209	0.0192
	[0.0428]	[0.0802]	[0.0761]	[0.0445]
lpg2	0.0003***	0.0003**	0.0004***	0.0004***
	[0.0001]	[0.0001]	[0.0001]	[0.0001]
cdd25	0.0006***	0.0004**	0.0010***	0.0007***
	[0.0002]	[0.0002]	[0.0002]	[0.0002]
urban	-0.0333**	-0.01	-0.0458	0.0018
	[0.0142]	[0.3529]	[0.2690]	[0.0179]
N	3646	3646	3626	3646
R-sq	0.831	0.958	0.28	

Note. * p<0.10, ** p<0.05, *** p<0.01; Standard errors in brackets
L.lnkwh1: the lag of dependent variable

Table 5-10. Estimated long-run function of MP model

Source. Author estimated.

Meanwhile, the system GMM estimator (xtabond2) incorporates the problem of bias in dynamic panel. In both models, the standardized parameters of the lag dependent variable fall in the range defined by OLS and LSDV estimates. Unfortunately, the signs of price variables are positive and significant at the level of 0.01 in both models. Robust checks are performed in two ways. First, I try other estimators such as Anderson and Hsiao (1982), difference GMM[3]. Second, I change the predetermined prices used to calculate price IVs. However, regardless of the changes in estimators or in IVs calculation, there is no major change in results. Although the magnitude of estimated parameters changes a bit, the signs and significances of all variables do not change (Table 5-11).

Model	*L.lnkwh1*	*lnmp*	*lnrsp*	*Lnap*
Anderson and Hsiao estimator				
MP	0.1535***	1.9343***	0.0170***	
	[0.0397]	[0.2865]	[0.0027]	
AP	0.4352***			4.8577***
	[0.0698]			[0.5158]
Single predetermined price at sample mean				
MP	0.0639**	2.6108***	0.0198***	
	[0.0288]	[0.0911]	[0.0009]	
AP	0.2491***			4.2174***
	[0.0349]			[0.1814]
Actual prices as predetermined prices				
MP	0.0892***	2.3877***	0.0157***	
	[0.0271]	[0.0530]	[0.0005]	
AP	0.2463***			4.5091***
	[0.0353]			[0.1223]

Note. * p<0.10, ** p<0.05, *** p<0.01; Standard errors in brackets

Table 5-11. Robustness tests for the long-run model with different IVs and estimators

Source. Author estimated.

[3] Indeed, with T=3, the Anderson and Hsiao (1982) estimators and difference GMM are identical.

Standard tests including endogeneity test and weak instrument variable test confirm that the above problem does not come from IV. First, the endogeneity tests confirm that there is endogeneity in both model MP and model AP (Table 5-12).

MP model		*AP model*	
F	p-value	F	p-value
33.1819	0.0000	138.895	0.0000

Note. Tests from models with weight and cluster adjusted at household level.
Endogeneity test is Durbin-Wu-Hausman test; H_0: variables are exogenous.
Test is based on Anderson and Hsiao estimator.

Table 5-12. Endogeneity tests for long-run models

Source. Author estimated.

Second, the weak instrument test rejects the hypothesis that instrument variables are weak. As discussed in the short-run estimation, the F statistic is much larger than the critical value of 10 as suggested in Staiger and Stock (1997). L.lnkwh1, lnmpiv2, lnrspiv2 and lnapiv2 are not weak instruments.

The only explanation here is the variation in price schedules. In fact, the two data sets are nearly identical in term of statistical summary (see the comparison in Chapter 4 for reference). If we treat the panel data as pooled data to estimates short-run electricity function, we get nearly the same result as the short-run function estimated from pooled data (Table 5-14). However, due to the correction for dynamic panel bias, the variation in price reduces sharply.

		Shea's Adj. Partial R-sq.	*F stats**	Cri. val
Model MP	L.lnkwh1	0.2154	113.25	10
	lnmp2	0.0409	22.76	10
	lnrsp2	0.0500	11.92	10
Model AP	L.lnkwh1	0.1280	134.64	10
	lnap2	0.1813	90.63	10

Note. Suffix 2 – adjusted to 2014 price.
* F test for join significance of instruments in first stage models
Test is based on Anderson and Hsiao estimator with weight.

Table 5-13. Weak instrument tests for long-run models

Source. Author estimated.

In the panel data set, coincidently, each year has a corresponding price schedule. The three years have three different price schedules. However, the correction for dynamic panel bias reduces the variation in price schedule. With the system GMM estimator, we employ $\Delta lnkWh1_{t-1} = lnkWh1_{t-1} - lnkWh1_{t-2}$ as the IVs for $lnkWh1_{t-1}$ in estimated equation. Obviously, except for the IVs, data for other variables collapse to 2016 data. Thus we have only unique price schedule in 2016 for estimation. If we use Anderson and Hsiao estimator, the model included difference of price except for the 1st period. We lose the price schedule of 2012.

The reduction in variation of price schedule can have a significant impact on estimators. Let's recall what Opaluch (1984, p. 417) claims.

> *It is critical that all customers do not face an identical rate schedule. A typical example might be 100 customers from each of 10 schedules, where ideally the differing rate schedules result from customers residing in similar towns within a single geographic*

region, but serviced by different utility companies. This is clearly different from having variation in rate schedules only because some customers own all electric homes, or for some other reason which will likely result in a shift in the individual's demand function. This sort of variation in rate schedules is critical for estimation, and in particular, for addressing the simultaneity issue. (Opaluch, 1984, p. 417)

A simple test is performed to test the role of price schedule variation. The short-run function estimated from the pooled data is used as the base for comparison. First, the panel data is treated as a pooled data to estimate short-run function. Second, data of 2012 is removed from the panel, the panel data of 2014 and 2016 is treated as a pooled data to estimate short-run function again. The results show that the estimated parameters of the pool-like panel data for all three years is close to the short-run function estimates with the pooled data set. However, the estimated parameters of MP for the pool-like panel data of 2014 and 2016 are still negative though insignificant at 0.1 level. The estimated parameter of AP is still negative and significant at 0.01 level, but the absolute value is approaching zero. The test result supports the explanation that price schedule variation has strong impact on estimation.

	MP		*AP*	
	OLS	*IV*	*OLS*	*IV*
Short-run estimates with the panel data as a pooled data				
lnmp1	2.0210***	-1.7207***		
	[0.0542]	[0.4587]		
lnrsp1	0.0114***	0.0083***		
	[0.0005]	[0.0018]		
Lnap1			4.9691***	-3.1948***
			[0.2716]	[0.5012]
Short-run estimates with the panel data of 2014 and 2016 only				
Lnmp2	2.3825***	-0.3938		
	[0.0642]	[0.6006]		
Lnrsp2	0.0158***	0.0101***		
	[0.0006]	[0.0028]		
Lnap2			5.2555***	-2.6803***
			[0.3138]	[0.6629]

Note. * p<0.10, ** p<0.05, *** p<0.01; Standard errors in brackets.
Table 5-14. Short-run estimates with the panel as a pooled data
Source. Author estimated.

Nevertheless, the conclusion is that with the current data set, it is hard to estimate the long-run demand function. The main obstacle is the nature of data which contain maximum of three years with single price schedule each year. If we ignore the dynamic panel bias, the estimated parameters are inconsistent. If we address the dynamic panel bias, the price schedule variation reduces sharply resulting in the wrong sign for price elasticity.

5.4.3 Perceived price model

Since we have an obstacle in estimating the long-run model, we have concentrated on the short-run model. First, the estimated parameters of the perceived price model are in line with the short-run model in term of signs and significance of parameters. However, it should be noted that the beta coefficient of lnmp1 is not the impact of marginal price on kWh. It is the impact of perceived price $P^* = MP\left({}^{AP}/_{MP}\right)^k$.

Second, the estimated coefficient of variable *k* is 0.153 and insignificant at the level of 0.1. This means that households fully react to MP. This result is similar to Nieswiadomy and Molina (1991). Nieswiadomy and Molina (1991) find evidence showing that customers react to AP in decreasing block tariff, but react to MP in increasing block tariff.

	OLS	*IV*
lnincpc1	0.1249***	0.2417***
	[0.0092]	[0.0144]
lnrsp1	0.0129***	0.0037***
	[0.0003]	[0.0013]
lnmp1	3.8332***	-1.6497***
	[0.0842]	[0.2277]
lnapmp1	3.5368***	-0.2525
	[0.1043]	[0.5527]
k		0.153
		[0.3219]
N	18971	18971
R-sq	0.873	0.602
adj. R-sq	0.872	0.600

Note. * p<0.10, ** p<0.05, *** p<0.01; Standard errors in brackets.
The table presents estimates of selected variables. The full result is in Appendix C.

Table 5-15. Estimates for perceived price model

Source. Author estimated.

In this case, some facts can provide an explanation for the result. In Vietnam, up to 2017, electricity invoices were usually delivered monthly to customers, mostly in person by EVN staffs. The staffs go to each household, meet household heads, hand over the bills and get the payments in cash directly. The electricity bill details all information, including the size of each block up to the household's consumption, the corresponding prices of each block, the payment for each block and the total value of the bill. The fact implies that customers have perfect information on electricity price. Thus it is reasonable that customers react to MP.

From 2018, EVN widely implemented online payments. Customers receive a text message on their mobile phone informing the total kWh and total value of electricity bill. Customers can use their online banking service to pay the bills. If customers want to see the bill detail, they must access the EVN website. The change in payment method could change the satisfaction of perfect information assumption and thus could change the type of prices that customers react to. We can test the perceived price again later when we have updated data. But for now, with the current availability of data and background information, we can conclude that households in Vietnam react to MP.

5.5 Conclusion

This chapter investigates the price elasticity of residential electricity demand under an IBTs schedule in the context of Vietnam from 2012 to 2016. The highlight is the master data set which is formed by merging national micro survey data at a household level (VHLSS) with other data such as electricity prices and climate condition in Vietnam, during a long period from 2012 to 2016. Due to the rotation rule of the VHLSS, the master data are divided into a pooled data set and a panel data set. The pooled data set includes all households in (i) the whole 2016 sample, (ii) the 2014 sample part which does not overlap with 2016 sample, and (iii) the 2012 sample part which does not overlap with the 2014 sample. The panel data set run includes all households that appear in all three years of 2012, 2014 and 2016.

The short-run demand function is estimated with the pooled data set. The long-run demand function is estimated with the panel data set. The pooled data set ensures the condition that electric appliances are fixed in short-run since there is no household that appears twice in the pooled dataset. Besides, the pooled data set has the advantage of price variation since the data span over three years with six different price schedules. Meanwhile, the panel data set contains data of the same households in three different years thus allowing observations about their changes in electricity consumption behaviors when electricity appliance stocks vary.

The estimate results show that (i) household electricity demand is elastic in the short-run. This contradicts the findings of studies made in other countries but agrees with Phu (2017) results in Vietnam, with different data and method. (ii) The long-run elasticity appeared positive, a surprise which may be caused by the lack of price variation in the long-run data set. (iii) Households in Vietnam respond fully to marginal price rather than to average price.

Chapter 6. Income and electricity poverty in Vietnam 2012–16

6.1 Introduction

Electricity plays a vital role in improving household welfare. Electricity provides a highly efficient lighting source at low unit cost. With lighting from electricity, households get increased time for working and receiving education which can improve productivity. Besides, electricity can also be used for cooling to protect householders from physical exposure during heatwaves.

The Vietnamese Government has put effort into ensuring the access of households to this kind of modern energy. So far, 98 per cent of households in Vietnam can access the national grid. The remaining two per cent are households in remote rural areas or island areas. In addition, the Government provides electricity subsidies to poor households in the form of both price reductions and cash transfers. First, electricity prices are designed in increasing block tariffs (IBTs) form so that poor households who normally consume electricity within the first block pay prices lower than the average selling price. Second, poor households are subsidized in cash for money equivalent to 30kWh at prevalent prices.

However, two issues arise from the Government's subsidy policies. First, the current first block is set at 50 kWh and the cash transfer subsidy is set at 30kWh. It seems that these thresholds are set arbitrarily. Second, the thresholds are set uniformly nation-wide, even though electricity demand differs between urban and rural areas due to differences in economic conditions, climate conditions, and lifestyle.

These two issues can be addressed by considering the concept of "electricity poverty". The electricity poverty line is the amount of kWh that household requires to meet its basic needs. The line can be used as the reference level for the first block threshold as well as the level of cash transfer subsidy.

In this study, we identify the electricity poverty threshold by investigating the non-linear relationship between income and electricity demand. The approach is adopted from Barnes *et al.* (2011). In practice, the study tests the hypothesis that there exists an income threshold from which household electricity consumption starts to increase as income increases. The average kWh consumption of households belonging to that income group is the electricity poverty threshold. In addition to the whole-nation sample, we also estimate the thresholds for other samples, including urban and rural areas, to see whether there is a difference in the electricity poverty threshold between those areas.

The following section provides a literature review which aims to find an appropriate method of estimating the thresholds, the third section provides model specification and data preparation, the fourth section considers results and discussion, and the last section provides a conclusion.

6.2 Literature review

In literature, there is no specific definition of "electricity poverty". However, there is a vast amount of research on the broader concept of energy poverty. The concept of electricity poverty is derived analogously to energy poverty. There is a consensus on energy poverty that it is the level of energy enough to fulfill the "basic needs". However, the way to measure the "basic needs" is still controversial. So far, there are two approaches to measure "basic needs": direct and indirect.

6.2.1 Direct measurement approach

The direct measurement approach is also known as the engineering approach. This approach measures energy poverty by estimating the direct energy required for basic needs. In other words, this approach calculates the minimum amount of energy required for a basket of goods and services.

Pachauri *et al.* (2004, pp. 2084–2085) categorizes energy consumption into four types according to the four levels of energy supplied, as follows.

> *— Primary energy is the energy embodied in natural resources (e.g., coal, crude oil, sun-light, uranium) that have been mined, collected or extracted but have not undergone any anthropogenic conversion or transformation.*
> *—End-use energy is the energy content of energy supplied to the consumer at the point of end-use (e.g., electricity at the electricity meter).*
> *—Useful energy is the energy that has been transformed in the form required for actual use, e.g., the heat generated from a hot plate or the mechanical energy applied to air for air circulation.*
> *—But finally, the direct demand by households is on energy services: a cooked meal, a well-lit room, being transported from A to B, a hot shower, etc.*

The direct measurement will measure energy poverty at the fourth level – energy service. One example is Bravo *et al.* (1983), cited in He and Reiner (2016). Bravo *et al.* (1983) suggests estimating energy needs for: preparation and preservation of food and supply of water; space conditioning; personal cleanliness; and recreation and social communication. Pachauri *et al.* (2004) did a survey from various research sources to form a list of power requirements of energy services of a five-member household such as "scooter per 5km/day", "Lighting, one electric bulb (5h/day, 40W)" "Lighting, 1–2 kerosene lamps".

The advantage of the approach is intuitive. However, the approach faces two issues. First, there is no consensus on definition of basic needs. Basic needs is a subjective concept which varies with climate, region, demographic (Pachauri *et al.*, 2004). Second, the calculation must be based on many assumptions, such as type of equipment, and energy consumption habits. Thus it is only appropriate on a case-by-case basis.

6.2.2 Indirect measurement approach

The indirect measurement approach measures energy poverty analogous to the income poverty concept. Instead of seeing the basic needs as a set of energy services, the indirect approach sets up a threshold for the total amount of energy consumption regardless of the components of the consumption. There are three ways to define the thresholds. The first one is to define an absolute threshold. The second one is a relative threshold. The last one is a demand-based threshold.

(i) Absolute thresholds are set to the average kWh of households that live under a poverty line. Foster *et al.* (2000) defines the threshold as the average energy consumption of households having expenditure per capita which falls within 10% of the official expenditure poverty line. This method is simple since the poverty line is readily available in every country. However, the underlying assumption of this method is that income-poor households are also energy poor households which is not necessarily correct.

(ii) Relative thresholds set the threshold as a percentage of household expenditure. DTI (2007) define a household is in energy poverty if it spends more than 10% of its income on fuel. This approach is based on the idea that poor household spends a higher fraction of income on energy than richer households. This approach is also simple and easy to apply. However,

it neglects the role of other factors, such as energy efficiency or price changes. In addition, the ratio is quite arbitrary.

(iii) Demand-based approach estimates demand functions to identify the fixed quantities of goods that households must consume and consider the quantities as subsistence levels. The approach is sensitive to model selection but has some advantages in comparison to other methods. First, it is simple since it depends on micro survey data which are available in every country. Second, it can avoid the problem of arbitrary choices as relative threshold methods. In addition, it is better than the direct approach since it does not impose any requirement of basic-needs baskets. Households choose their own "basic needs" baskets. What we do is to observe household behaviors.

Stone-Geary model is the most well-known representative of the approach. The Stone-Geary model is based on neo-classical ordinal utility theory. There are two goods of *w* and *z*. A household has the utility function of $U = (Q_w - \gamma_w)^{\beta_w}(Q_z - \gamma_z)^{\beta_z}$ with $\beta_w + \beta_Z = 1$. The household will maximize its utility under a budget constraint of $I = P_w Q_w + Q_z$ (the price of z is normalized to 1). The household consumes a minimum acceptable level of each good first (γ_i), then spends residual income in fixed proportion (β_i) to each good to get utility. γ_i is called subsistence level and β_i is marginal budget share. Solving the equations results in the following demand function.

$$Q_w = (1 - \beta_w)\gamma_w + \beta_w . \frac{I - \gamma_z}{P_w} \qquad (6\text{-}1)$$

Assuming that $\gamma_z = 0$, equation (6-1) can be estimated empirically with micro data to get $\hat{\gamma}_w$. The Stone-Geary model has been frequently used to estimate residential water demand (Martínez-Espiñeira* and Nauges, 2004; Dharmaratna and Harris, 2012; Hung and Chie, 2013). However, the Stone-Geary model is only appropriate for inelastic demand (Gaudin, Griffin and Sickles, 2001, p. 404). The restriction of inelastic demand can be seen from equation (6-2).

$$E_P^D = \frac{\partial Q_w}{\partial P_w}\frac{P_w}{Q_w} = -\frac{\beta_w(I - \gamma_z)}{\beta_w(I - \gamma_z) + P_w(1 - \beta_w)\gamma_w} > -1 \qquad (6\text{-}2)$$

Recently, Barnes *et al.* (2011) proposes a new way to identify the subsistence level without the restriction of inelastic demand. They define the concept of energy poverty as "the threshold point at which energy consumption begins to rise with increases in household income". The core idea is to find a portion of electricity use that may not responsive to income rather than not responsive to price as in Stone-Geary model.

Several researchers have applied the concept to implement in different databases. Khandker *et al* (2012) uses this approach to justify the absolute threshold method. They found that "in rural areas some 57 per cent of households are energy poor, versus 22 per cent that are income poor. For urban areas the energy poverty rate is 28 per cent compared with 20 per cent that are income poor". He and Reiner (2016) applied the concept for electricity consumption in China and found that "whereas the first IBT block was set at 240 kWh per household for Beijing, we estimate basic needs to be roughly 90 kWh per month for rural households and 150 kWh for urban households".

However, Barnes et al. (2011) approach is based on an assumption of non-linear relationship between income and kWh consumption. The approach asserts that the relationship between household income and kWh consumption has an S-shape (Douglas F.

Barnes, Khandker and Samad, 2011). Households with income less than I_1 are households that do not have enough income for the basic needs such as food or clothes. Though the households can't fulfill their basic needs, they still need to use a certain amount of electricity (kWh_1), the reason being that electricity for the households work as a capital good. For example, poor households use electricity for lighting to work at night or charge their mobile phone. In Vietnam, mobile phones in these households play a role of a capital good since they are used as a communication channel to get jobs. The role of mobile phones is getting more important for part-time or seasonal workers. In this case, no matter how poor the household is, it needs to use that amount of kWh. For these households, kWh as a function of income are roughly horizontal.

When household income passes the level of I_1, households have enough income to fulfill their basic needs, thus can allocate a part of income in other goods that can increase their welfare. The other goods include electricity as well as electric appliance. For those households, kWh as a function of income are increasing. The richer a household, the higher kWh the household consumes. However, when household income reaches the level of I_2 and higher, the households do not care about electricity price anymore. They also own a lot of electric appliances which surpass their needs. In this case, kWh_2 is the saturation state.

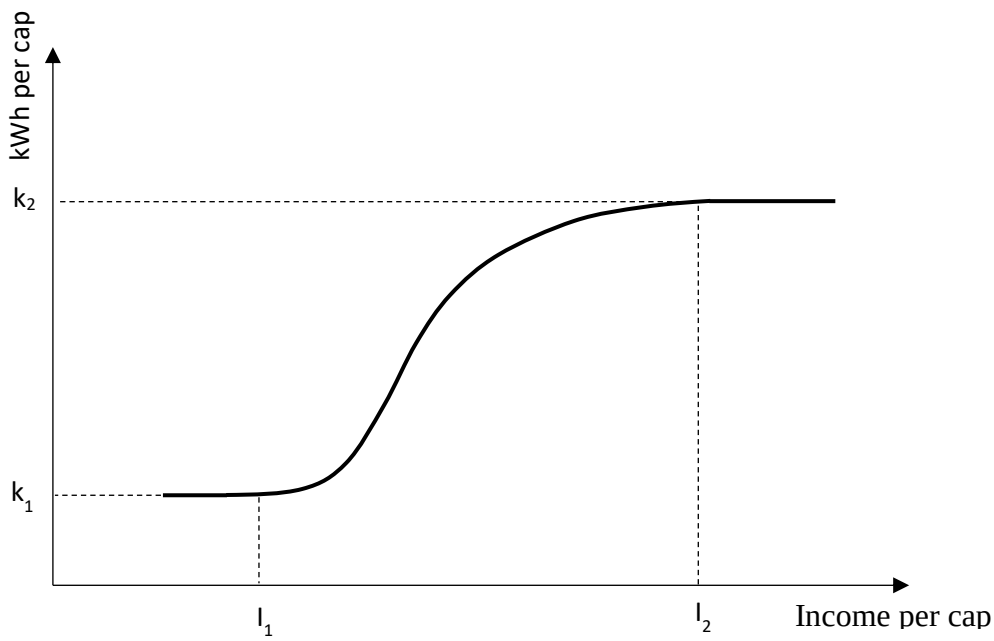


Figure 6-1. The S-shaped relationship between household income and kWh consumption
Source. Author illustrated.

The chapter emphasize on testing empirically the existence of I_1. In other words, the chapter estimates the electricity poverty thresholds (kWh_1) for national, rural and urban areas in Vietnam.

6.3 Methods and data

❖ *Methods*

The aim of the study is to test empirically the existence of threshold I_1 in Figure 6-1. The threshold is tested in two ways. First, the relationship between household income and electricity consumption is visualized with density scatter plots with a series of fitted lines for different income quantiles. The visualization would reveal whether the thresholds exists.

In the second way, the threshold is formalized test by the econometric approach pioneered by Barnes *et al.* (2011):

- divide income per cap to *n* quantiles, take the poorest quantiles as the base
- estimate the equation

$$\ln kWhpc_i = \alpha_0 + \sum_{j=2}^{n} \alpha_{ij} I_{ij} + \beta Z + \varepsilon_i \quad (6\text{-}3)$$

Where

$\ln kWhpc_i$ = ln of kWh per cap of household *i*
i = household index
j = quantile index, the higher j, the higher quantile
I_{ij} = 1 if the income per cap of household *i* fall in j^{th}-quantile, =0 otherwise
Z = all other variables in short-run demand function in Chapter 5 except for income

The coefficients α_j show the difference in kWh usage per capita between the base group (poorest group) and the group *j*. If $\alpha_j = 0$, there is no difference in electricity use between the base and the group *j*. If $\alpha_j > 0$, the group *j* consumes higher level of electricity per capita than the base group does. If $\alpha_j < 0$, the j group consume lower level of electricity per capita than the base group does.

If the theory prediction on the existence of I_1 is correct, we expect that the estimated coefficient of α_j has the following pattern. The first few α_j would be zero, the rest of α_j would be greater than zero. The first group (group *k*) that has $\alpha_k > 0$ is the group that has I_1. Let's denote $\alpha_j = 1$ for any $\alpha_j > 0$, we expect to see a series of α as follows.

$$\alpha_2 = 0, \alpha_3 = 0, \ldots, \alpha_{k-1} = 0, \alpha_k = 1, \alpha_{k+1} = 1, \ldots, \alpha_n = 1 \quad (6\text{-}4)$$

We can identify the electricity poverty threshold k_1 in Figure 6-1 by taking the mean kWh per cap of the group *k*.

The number of quantiles is chosen to ensure that there is at least $\alpha_2 = 0$. This is to avoid the situation of all $\alpha_j = 1$. In that case, there is an overestimate of k_1 since the base group can be divided further to a group having $\alpha = 0$ and a group having $\alpha = 1$. The choosing procedure is trial and error, with increasing number of quantiles until satisfying the above rule.

For econometric techniques, we apply price IVs to estimate the equation (6-3) as in Chapter 5. Second, all models are estimated with household weight. Third, as discussed in Chapter 5, it is not reliable to estimate the long-run function with panel data due to the lack of price variation. Thus this chapter only estimates the short-run function.

❖ *Data*

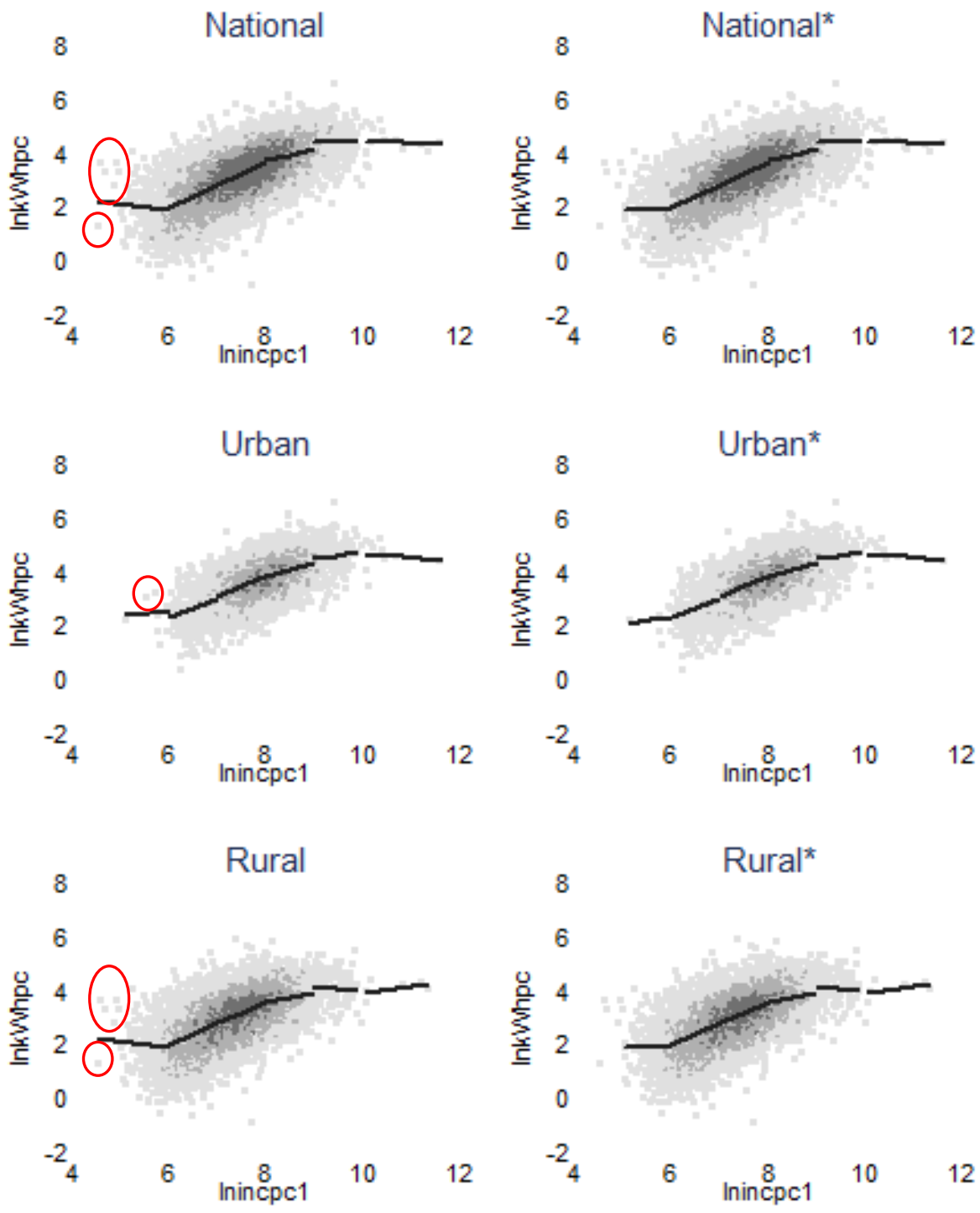
The data for both visualization and econometric estimation are the pooled data set. All variables in monetary terms are adjusted to the 2012 price base. The data set is separated into two sub-data sets of rural and urban so that we can create a national model as well as separate models for urban and rural areas.

6.4 Results

6.4.1 Visualization approach

At first glance, it seems that all fitted lines in the left-hand side graphs in Figure 6-2 are S-shaped as predicted. These graphs are for full samples without removing any outliers. The S-shaped pattern supports the hypothesis that kWh per capita starts to increase after a certain level of income. The pattern happens to all samples, including national, urban and rural samples.

However, these results are a bit unexpected since the first part of the fitted lines have downward slopes meaning that the higher the income per capita, the lower the kWh per capita consumption. A closer look at the graph shows that the national sample and the rural sample have six outliers while the urban sample has two outliers. If these outliers are removed, the downward-slope issue no longer exists. The graphs for samples without outliers are in the right-hand side of Figure 6-2 (with the * notation). As we can observe from the graph, the first parts of the fitted lines have horizontal or upward slopes.



Note. The graphs are drawn without weight. * Samples without outliers; outliers are in circles. Fitted lines are estimated by OLS estimators for different income range

Figure 6-2. Density scatter plots with a series of fitted lines for different income quantiles

Source. Author estimated.

6.4.2 Econometric approach

The quantile choosing process ends up with the quantiles of 100 for all three samples. The detail of estimates is provided in Table 6-1.

	Original data			*Data without outliers*		
lnkwh1pc	*National*	*Rural*	*Urban*	*National*	*Rural*	*Urban*
Income quantiles (selected first 10 quantiles out of 100 quantiles)						
_Iqincpc1_2	-0.0199	-0.1177	-0.0164	0.0737	0.033	-0.0503
	[0.0886]	[0.1034]	[0.0923]	[0.0835]	[0.0938]	[0.0986]
_Iqincpc1_3	0.1047	0.0218	0.0878	0.1589**	0.1259	0.0664
	[0.0748]	[0.0904]	[0.0784]	[0.0679]	[0.1039]	[0.0840]
_Iqincpc1_4	0.1631**	0.0077	0.1600**	0.2087***	0.0941	0.1479*
	[0.0780]	[0.0733]	[0.0798]	[0.0806]	[0.0846]	[0.0758]
_Iqincpc1_5	0.1681**	0.1457*	0.1125	0.2195***	0.2552***	0.128
	[0.0702]	[0.0811]	[0.0879]	[0.0638]	[0.0861]	[0.0853]
_Iqincpc1_6	0.1315*	0.108	0.2153**	0.1639**	0.1927**	0.1978*
	[0.0731]	[0.0820]	[0.0975]	[0.0726]	[0.0883]	[0.1042]
_Iqincpc1_7	0.2007***	0.1121	0.1206	0.2517***	0.1931**	0.0987
	[0.0731]	[0.0816]	[0.0846]	[0.0739]	[0.0781]	[0.0895]
_Iqincpc1_8	0.1615**	0.0746	0.1605*	0.2188***	0.1498**	0.1679*
	[0.0758]	[0.0709]	[0.0853]	[0.0701]	[0.0757]	[0.0902]
_Iqincpc1_9	0.1903**	0.1317*	0.2841***	0.2478***	0.2244**	0.2724***
	[0.0774]	[0.0776]	[0.0867]	[0.0768]	[0.0881]	[0.0945]
_Iqincpc1_10	0.2030**	0.1214	0.3180***	0.2562***	0.2271***	0.3517***
	[0.0804]	[0.0822]	[0.0770]	[0.0787]	[0.0872]	[0.0863]
Prices						
lnmp1	-1.1447***	-1.5356***	-0.1579	-1.1723***	-1.5815***	-0.2229
	[0.1966]	[0.2485]	[0.2854]	[0.2039]	[0.2638]	[0.3170]
lnrsp1	0.0028***	0.0026**	0.0034**	0.0025***	0.0023*	0.0025
	[0.0009]	[0.0013]	[0.0016]	[0.0009]	[0.0012]	[0.0018]
N	18971	13171	5800	18965	13165	5798
R-sq	0.633	0.557	0.665	0.627	0.547	0.645
adj. R-sq	0.629	0.55	0.654	0.623	0.541	0.634

Note. All quantiles above the 10th quantiles are positive and significant at the level of 0.05. Full estimated results are in Table D-1 and Table D-2, Appendix D.

Table 6-1. Estimates of kWh per cap on the quantiles of income per cap

Source. Author estimated.

First, there is no unexpected result – all the beta coefficients have signs as expected. The endogeneity test fails to reject the existence of endogeneity (refer to Table D-3 Appendix D). The weak IV test for price rejects the hypothesis of weak IVs (refer to Table D-3 Appendix D).

For income quantile variables, if the outliers that are detected in visualization are not removed, the pattern of α_j is not in the form of 0..0, 1..1 as predicted. For example, if we take the significant level at 0.05, the pattern of α_j for national sample is 0011011… and for the urban sample is 00101001… .

However, if the outliers are removed, the estimate $\hat{\alpha}_j$ for all three samples are in line with the theory predicted pattern. The income quantiles start to be significant at the level of 0.05 at the third (3rd) quantile for national sample, at fifth (5th) quantile for rural sample and at ninth (9th) quantile for urban sample. All the quantiles below the thresholds are insignificant at the level of 0.05. All the quantiles above the thresholds are positive and significant at the level of 0.05.

The quantile threshold of national is smaller than the thresholds of urban and rural. The reason is that although all three models divide income to percentiles, they are based on different income ranges: the national sample is based on the income of all households while the urban/rural samples are based only on urban/rural households.

Table 6-2 shows the mean kWh that households in the threshold quantile groups consume for a month. On average, the monthly electricity poverty threshold for national is 51.36 kWh, for urban area is 118.61 kWh and for rural area is 54.15 kWh.

		National (group 3)		*Urban (group 9)*		*Rural (group 5)*	
		N	*mean*	*N*	*mean*	*N*	*mean*
(1)	kwh1pc	193	12.39	58	28.72	136	13.42
(2)	incpc1	193	416.63	58	1,161.30	136	449.96
(3)	hhsize	193	4.14	58	4.13	136	4.04
(4)	kwh1*		51.36		118.61		54.15

Note. Unit of monthly income per cap at 2012 price (incpc1): '000 VND; * (4) = (1) *x* (3).
Table 6-2. Average kWh consumption of households having corresponding income quantiles
Source. Author estimated.

This result is similar to the estimates of He and Reiner (2016) for Beijing. He and Reiner (2016) found that the electricity poverty threshold for urban area is higher than for rural area. They are 150 kWh and 90 kWh corresponding. In this research, we find lower thresholds but in the same pattern. The thresholds are about 118 kWh for urban households and 54 kWh for rural households.

The result is quite reasonable for two reasons. First, households in urban areas have a higher demand for lighting. It is due to the difference in lifestyle between urban and rural areas. In rural areas, people tend to go to sleep earlier than people in urban areas. Second, due to urban heat island effect, temperature in urban areas is higher than in rural areas, especially in summer. This fact is confirmed by the estimated coefficient of cdd25 in the estimates result. Thus households in urban area need more electricity for cooling needs.

No.	*Electricity Services*	*kWh Consumption*
1	Charge phone everyday	1
2	Boil 5 teapots everyday	2
3	Lighting, 4 lamps 4 hour per day	7
4	Fan, 8 hours per day	17
5	Rice cooker OR Television	23
	Total	*50*

Table 6-3. Electricity services a household can consume with 50 kWh per month
Source. Ha-Duong (2016).

Nevertheless, Table 6-2 reveals that at national level, the minimum level of electricity for a household of four people is 51 kWh. The result is understandable if we see what a household can do with 50 kWh per month. As shown in Table 6-3, with 50 kWh, a household can barely fulfill its need for lighting, charging phones, cooling and cooking.

6.5 Conclusion

This chapter employs an electricity demand model to identify the electricity poverty threshold for households in Vietnam during the period 2012–16. The electricity poverty threshold is defined as the threshold at which electricity consumption begins to rise with increases in

household income (Douglas F. Barnes, Khandker and Samad, 2011). The electricity demand in use is a short-run demand function. The demand function is estimated for the whole national sample as well as for urban and rural areas.

Estimated results provide evidence that an electricity poverty threshold does exist. For the whole national sample, the electricity poverty threshold for a household of four people is 51 kWh. The threshold is approximately the threshold for the first block in IBTs schedule in Vietnam. However, the empirical threshold is higher than the threshold for the cash transfer subsidy program. Besides, estimated results also disclose that there is a significant difference in the electricity poverty thresholds between urban and rural areas. The thresholds are about 118 kWh for urban households and 54 kWh for rural households. This finding can be used as empirical evidence for policy makers in Vietnam in evaluating current policies, as well as in designing appropriate policy for electricity prices and subsidy programs in the future.

Chapter 7. Economies of scale in residential electricity consumption

7.1 Introduction

Vietnam moved to a market-oriented economy in 1986. However, electricity is one of a few special goods that still have prices set by government. Since 1994, it has set electricity price in increasing block form to support low-income households and provide a disincentive to high consumption due to a mismatch between demand and supply. In the newest proposal for electricity price reform, EVN provides three alternative price schedules, including two increasing block tariffs and one single price (EVN, 2015). However, many experts disagree with the single price structure and are in favor of increasing block tariffs. Debatable topics include the number of blocks; price gaps between blocks; and the impact of increasing block tariffs on low-income households (Châu Anh, 2015; Đình Dũng, 2015).

Yet there is no research or official discussion on the impact of the increasing block tariffs on large-size households. This is a serious gap since large-size households will suffer the high price due to high demand while they usually have low income[4]. In this case, increasing block tariffs may turn out to be a penalty for some low-income households instead of protecting them.

This chapter uses data from the Vietnam Household Living Standard Survey (VHLSS) 2012–16 to investigate whether the current increasing block tariffs have a negative impact on large-size household electricity consumption. In other words, we examine how the increasing block tariffs impacts on economies of scale of household electricity consumption in Vietnam from 2012 to 2016. The result will provide empirical evidence for policy makers to design electricity price schedules in the future.

7.2 Literature review

7.2.1 Economies of scale

Economies of scale in household consumption is the phenomenon in which the cost per capita that maintains a given level of living standard may reduce as household size increases (Nelson, 1988, p. 1301). Economies of scale of household consumption may come from three sources (see Nelson, 1988 for review).

- First, economies of scale come from increasing return in household production such as cooking meals.
- Second, it may come from "bulk buy". When household size increases, demand for goods and services increases. The household may obtain discounts for purchasing goods and services in bulk.
- Third, it may come from the consumption of public goods in which the consumption of one household member does not eliminate, or eliminate completely, the consumption of other members. Since the public goods, such as lighting or air conditioners, can be shared as the size of household increases, the cost of the goods per capita declines. In addition, an increase in household size can also reduce the cost per capita for that public good because of the increase in the utilization rate of the public goods which are indivisible such as water heating, pilot light or a refrigerator.

Economies of scale in household consumption have been found in many goods and services. Nelson (1988) found substantially and statistically significant economies of scale for five classes of goods and services including food, shelter, household furnishing/operation, clothing and transportation in US data during 1960–61 and 1972–73. Deaton and Paxson (1998) found

[4] Correlation between household size and income per capita in VHLSS 2014 is negative and significant at the 0.05 significance level.

that at any given household per capita expenditure, expenditure per head on food falls as the household size increases in seven countries (USA, Great Britain, France, Taiwan, Thailand, Pakistan and South Africa).

A major empirical problem in detecting economies of scale is to separate the impact of household size from the impact of household composition. Nelson (1988, p. 1302) suggests that "Observed household demands may be expected to vary with household size not only because of economies of scale, but also because of the varying preferences or needs of household members, from infants to grandparents."

Three approaches have been employed to handle this problem so far. The first approach requires a strict assumption that preferences are identical among all household members (Nelson, 1988). In empirical section, Nelson (1988) studies only all-adult households with "heads" aged 35-55. Thus he can eliminate the impact of composition factor in observed demand.

The second approach is to use two separate variables for household size and composition (Ironmonger, Aitken and Erbas, 1995; Deaton and Paxson, 1998). The household size variable is the total number of household members. The household composition can be represented by category variables (Ironmonger, Aitken and Erbas, 1995) or continuous variables (Deaton and Paxson, 1998). Ironmonger, Aitken and Erbas (1995) uses this approach for three types of adult-only households including young households with adults from 15 to 45, older households with adults over 45 and mixed households with adults over 15. Deaton and Paxson (1998) use *(k-1)* variables for household composition. Each household is separated to *k* groups defined by age and sex. Each of the *(k-1)* variables above is the ratio to household size of household members who fall in the corresponding group. In this approach, the variable household size corresponds to the concept of doubling the number of household members while keeping family composition constant. Therefore, the approach can eliminate the impact of difference in member preferences in household consumption.

The third approach is to use an adult equivalent scale. The adult equivalent scale shows the proportional increase in household expense caused by each additional member to ensure the household is as well off as a single adult household (Browning, Chiappori and Lewbel, 2013). The expense proportion of each additional member is normally determined by household size and member age. For example, the OECD equivalent scale (also called Oxford scale) takes the first member as 1, each additional adult member as 0.7, each additional child member as 0.5 (OECD, no date). The approach is simple and easy to implement. However, the scale is defined by the increase in a wide range of needs, rather than the increase in a single item such as electricity.

Of all three approaches, Deaton and Paxson (1998)'s approach has an important advantage. In addition to identifying the impact of household size, it allows an investigation into differences in preference between a certain group of the *(k-1)* groups with the base group (the kth group). Therefore, this chapter applies the Deaton and Paxson (1998) approach. Each household will be separated into three groups including children who are 15 and younger, adults from 16 to 59 and elders who are over 60. Two variable children ratio and elder ratio will be employed to represent household composition. The beta coefficients of the variables indicate whether there is a difference in consumption between a child or an elder and an adult.

7.2.2 Economies of scale for household electricity expenditure

Electricity consumption has high potential for economies of scale in household consumption since it is a typical public good. People do not consume electricity directly but indirectly via appliances which can be shared among household members, such as lighting or cooling devices.

When a household's size increases, it can maximize the use of shared goods, including electricity use (Ironmonger, Aitken and Erbas, 1995). Therefore, the household can decrease the amount of electricity consumption per capita.

So far, researchers have found empirical evidences for economies of scale in household electricity (or energy in general) consumption. Brounen *et al.* (2012) use data from Dutch households in 2008–09 to investigate the impacts of dwelling characteristic, as well as demographics, on energy use, including gas and electricity. One of their models has a dependent variable of kWh use per capita matched with household size as an independent variable. The beta coefficient of the household size variable is -0.18 which is significant at the level of 0.05, meaning that when household size increase, the electricity use per capita decreases.

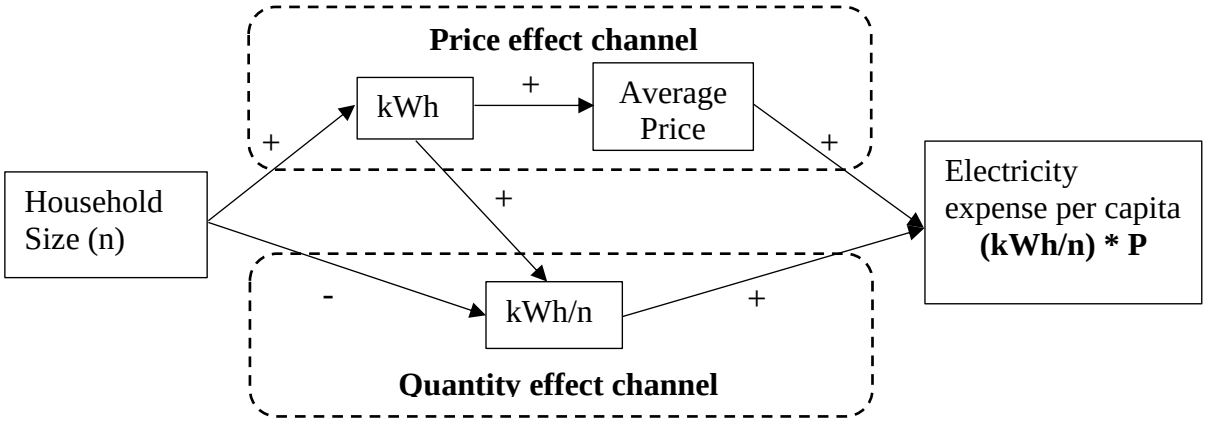
Filippini and Pachauri (2004) use India household data in 1993–94 to estimate India residential electricity demand. Their dependent variable is monthly total kWh consumption of households in logarithmic form. One of independent variable is a dummy variable of household size. The dummy var is 1 if household size surpasses median size of six people and is zero otherwise. They found that the beta coefficient of the dummy is negative and significant at the level of 0.05, meaning that total kWh consumption of a household decreases if the household size increases which is evidence supporting economies of scale.

Zhou and Teng (2013) estimate urban residential electricity demand in China using household survey data for a province during 2007–09. The dependent variable is household annual kWh consumption in logarithmic form. The household size variable is the total number of members. The beta coefficient of household size is about 0.08 which is significant at the level of 0.01, indicating that household size does have a scale effect on electricity consumption.

The above researches are on the economies of scale in electricity use in term of kWh. Only a few researches have studies economies of scale in electricity expense. Ironmonger, Aitken and Erbas (1995) investigate data from Australia in 1987 and 1990 and find that as household size increases, energy efficiency increases and energy expense per capita decreases. They explain that the economies of scale is the result of a fixed base charge per household which is independent of the level of energy use. Salari and Javid (2017) use US household data 2010–12 to model household energy expenditure in the US. They find that if household size doubles, electricity expenditure of a household increases 26–30 per cent.

However, to our best knowledge, none of the researches on the economies of scale of electricity expense are within an IBTs context. This is a gap since IBTs can offset the economies of scale from saving in quantity. When a household becomes larger, its electricity use per cap in kWh decreases due to sharing property. At the same time, the higher level of kWh consumption, the higher price the household has to pay with IBTs leading to higher average prices. The higher average price can cancel the saving from sharing property.

Figure 7-1 shows the two effects of changes in household size on electricity expense per capita. The first effect is a quantity effect due to the sharing characteristic. When household size increases, the household electricity consumption in kWh increases. However, due to sharing characteristic, the electricity consumption per capita in kWh decreases. The second is price effect. When the household size increases, the household electricity consumption in kWh increases. Thus the price each member has to pay increases due to increasing block tariffs. If the quantity effect dominates, households enjoy economies of scale. If the price effect dominates, diseconomies of scale exist.



Note. kWh – Household consumption of electricity in kWh; P – Electricity price.
kWh/n – Electricity consumption per capita in kWh.

Figure 7-1. Channels of economies of scale in household electricity expenditure

Source. Author compiled.

7.3 Methods and data

This chapter utilizes both non-parametric and parametric methods. The non-parametric method controls less variables and applies only to a fraction of the sample, but it is intuitive and allows a comparison directly among households having the same composition. The parametric method allows a test of economies of scale in the whole sample, keeping other variables constant.

7.3.1 Non-parametric method

This method is adopted from Deaton and Paxson (1998). The approach is based on an Engel curve that demonstrates the relationship between the expense for a good and income of a household. This method compares two household *i* and *j*. The two households have the same composition but are different in household size to capture the idea of doubling a household keeping its composition constant.

In this chapter, a household composition is represented by the ratio of children/adults/elders. For example, households with the pattern of (0, 1, 0), (0, 2, 0), (0, 3, 0) are households with the same pattern of adults only. Households with the pattern of (1, 1 ,0), (2, 2, 0), (3, 3, 0) are households without elders and have the number of children being equal the number of adults. This classification allows us to compare whether a larger household has smaller electricity share at any given income per capita level.

If (7-1) exists then economies of scale exist.

$$E\left(\frac{Elec_bill}{n}\middle|i,\frac{hhinc}{n}\right) < E\left(\frac{Elec_bill}{n}\middle|j,\frac{hhinc}{n}\right) \qquad (7\text{-}1)$$

Where

n = household size

$Elec_bill$ = electricity bills

hhinc = household income per capita

i, j = two households with the same household composition but $n_i > n_j$

Since we compare two households with the same economic condition, we can divide both sides of equation (7-1) by income per capita ($\frac{hhinc}{n}$). Testing equation (7-1) is now equivalent to

$$E\left(\frac{Elec_bill}{hhinc}\middle|i,\frac{hhinc}{n}\right) < E\left(\frac{Elec_bill}{hhinc}\middle|j,\frac{hhinc}{n}\right) \quad (7\text{-}2)$$

In practice, we test

$$E(w|i,z) < E(w|j,z) \quad (7\text{-}3)$$

Where

w = $\frac{Elec_bill}{hhinc}$: the ratio of electricity bill over household income (7-4)

z = ln($\frac{hhinc}{n}$) : household income per capita (7-5)

For each household type *i* or *j*, the procedure to calculate $E(w|z)$ is as follows.

- Step 1: Use local polynomial regression to calculate expected w_m at all points of regression grid. In particular, estimate w = f(z) + u_i where f(.) is not specified with local polynomial regression. Local polynomial regression is a non-parametric method which lets the data suggests an appropriate function form of f(.) instead of imposing a structure on data as parametric method. The procedure is detailed in Fan and Gijbels (1991). First, divide z to 50-point (z_m) equal periods. Second, at any point z_m, calculate a local weighted average of w on the neighborhood of z_m. The closer z_i to z_m, the higher weight z_i has. There is no or little weight assigned for z_i which is far from z_m.

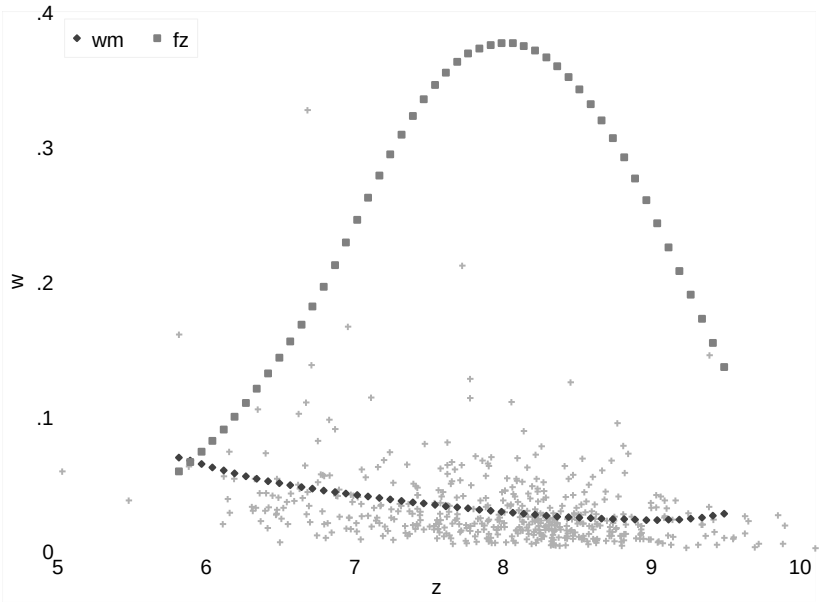
- Step 2: Calculate $E(w|z)$ as the weighted average of the 50 w_m above

  $$E(w|z) = \frac{\sum_{m=1}^{50}(w_m f(z_m))}{\int_{m=1}^{50} f(z_m)} \quad (7\text{-}6)$$

  where $f(z_m)$ = The kernel density of z at point z_m.

  If we choose the same f(z) for both *i* and *j*, it is equivalent to the condition that *i* and *j* have the same income per capita. Therefore, we can directly compare (7-6) across household types.

- Step 3: Calculate standard errors of the $E(w|z)$ by bootstrap method.



Note. Dots other than wm and fz are the density scatter plot of w over z
Figure 7-2. Illustration of w_m and $f(z_m)$ over regression grid
Source. Author illustrated.

Figure 7-2 illustrates the w_m and $f(z_m)$ over regression grid for the household type of single adult only. The horizontal axis is income per cap in logarithmic form (z). The wm is the local polynomial fitted line for w over z. The fz is the kernel density of z. The plot other than wm and fz are the density scatter plot of w over z. The regression grid is defined by dividing a certain range of z to 50-point (z_m) equal periods. For each value of z_m, there are two corresponding values in vertical axis. The first value is w_m which is the fitted value of wm at z_m. The second is $f(z_m)$ which is the value of fz at z_m. Since there are 50 points of z_m, there are also 50 points of w_m as well as 50 points of $f(z_m)$ as in Figure 7-2.

7.3.2 Parametric method

❖ *Econometric function*

This approach employs an econometric model to test economies of scale in electricity consumption. The model is like the short-run demand function. The function is as follows.

$$\begin{aligned} \textit{ln}\ \text{eleclm1}_i = {} & \beta_0 + \beta_1\, \textit{ln}\ \text{hhsize}_i + \beta_2\, \textit{ln}\ \text{childrenratio}_i + \beta_3\, \textit{ln}\ \text{elderratio}_i + \beta_4\, \textit{ln}\ \text{sexratio}_i \\ & + \beta_5\, \textit{ln}\ \text{incpc1}_i \\ & + \sum \beta_{6j}\, \text{H}_{ij} + \beta_7\, \text{ownership}_i + \beta_8\, \textit{ln}\ \text{pcsqm}_i + \sum \beta_{9j}\, \text{A}_{ij} + \beta_{10}\, \text{lpg}_i + \beta_{11}\, \text{cdd25}_i \\ & + \beta_{12}\, \text{y2014}_i + \beta_{13}\, \text{y2016}_i + \sum \beta_{14j}\, \text{tinh}_{ij} + \varepsilon_i \qquad (7\text{-}7) \end{aligned}$$

in which:

eleclm1	=	electricity bill of the last month (of the survey month) at 2012 price
hhsize	=	total number of household members
childrenratio	=	ratio of members who are 15 and younger over hhsize
elderratio	=	ratio of members who are 60 and above over hhsize
sexratio	=	no of male/hhsize
incpc1	=	household monthly per capita income at 2012 price
H_j	=	a vector of household characteristic including household head's gender, age and highest education
ownership	=	0/1 if the household rents/owns the dwelling;
pcsqm	=	per capita square meter in the dwelling

A_j	=	a vector of dummy variables for owning appliances including AC, fan, fridge, washing machine, oven, computer, TV, waterheater and generator. = 1 if yes, =0 if no
lpg	=	expenditure on LPGs
cdd25	=	CDD of the month before survey month
y2014, y2016	=	dummy variables for the years of 2014, 2016
$tinh_j$	=	vector of dummy variables for each city with Ha Noi is the base

The coefficient β_1 shows whether a household has economies of scale in electricity expenditure. The econometric is designed so that the variable *hhsize* represents the concept of doubling the household while keeping the same household composition which is control by *childrenratio* and *elderratio* variables. If $\beta_1 > 1$ then electricity bill increases more than double when household size doubles. Intuitively, this means diseconomies of scale. Vice versa, $\beta_1 < 1$ provides evidence for the existence of economies of scale.

The coefficients of *childrenratio* and *elderratio* represents for the differences in electricity consumption between a child or an elder and an adult. Children are members who are 15 and younger. Elders are members who are 60 and above. Adults are members from 16 to 59.

❖ *Estimation techniques*

Estimating the econometric function faces a technical problem. There is a weak but significant correlation between household size (*hhsize*) and income per capita (*incpc1*) at the level of 0.05 (Table 7-1). Thus besides OLS, we let income per capita enter the model non-parametrically. This approach allows us to estimate the linear coefficients of household size (*hhsize*) without the presence of income per capita (*incpc1*) in the model. The non-parametric technic is the first differencing which is discussed in Estes and Honor (1995), Yatchew (1997). Deaton and Paxson (1998) also applies this technic to solve the same problem with economies of scale of food expenditure.

	lnhhsize	lnchildrenratio	lnelderratio
lnincpc1	-0.1142*	-0.1662*	-0.0778*

Note. *: significant at the level of 0.05

Table 7-1. Correlation between income and demographic variables.

Source. Author estimated

The first differencing method contains three steps. First, it sorts the data on income per capita. Second, it takes the first differences of all variables in the original model. Third, it applies the OLS estimator to the new model of first differences of electricity bills on the first differences of all other variables except for income per capita. Let's consider the following simple function as an example.

$$\ln eleclm1_i = f(incpc1_i) + \beta\, hhsize_i + \varepsilon_i \qquad (7\text{-}8)$$

First, we sort $incpc1_i$ in increasing order $incpc1_1 < incpc1_2 < \cdots < incpc1_N$.
Second, we take the first differencing

$$(\ln eleclm1_i - \ln eleclm1_{i-1}) = \left(f(incpc1_i) - f(incpc1_{i-1})\right) + \beta(hhsize_i - hhsize_{i-1}) + \varepsilon_i - \varepsilon_{i-1} \qquad (7\text{-}9)$$

Yatchew (1977) asserts that when N increase, $\left(f(incpc1_i) - f(incpc1_{i-1})\right) \rightarrow 0$ because the derivative of f is bounded. Thus (7-10) becomes (7-11) which can be estimated by OLS.

$$(\ln eleclm1_i - \ln eleclm1_{i-1}) = \beta(hhsize_i - hhsize_{i-1}) + \varepsilon_i - \varepsilon_{i-1} \qquad (7\text{-}10)$$

Yatchew (2003) asserts that this method can lead to the less efficient of estimator in comparison to other non-parametric methods such as the "double residual method" of Robinson (1988) or the higher-order differencing of Yatchew (1997). However, this method is appropriate because it is simple to implement in popular statistic software with more options than other methods. First, since it uses OLS, it allows the incorporation of sample weight in Stata command which is not available in commands for other methods. Second, it allows the conduct of standard diagnostic tests after estimation.

7.3.3 Data

The data for non-parametric analysis is the pooled data set. In this data, households are classified to household types and household groups. Household types are classified by the number of children, adults and elders. The analysis is applied to household types having at least 100 observations. Household groups are the groups of household types having the same ratio of children:adults:elders. A common range of *z* is defined for all household types. The lower bound is the maximum of all minimum *z* of all household types. The upper bound is the minimum of all maximum of *z* of all household types. F(z) is estimated at household group level.

Group	*Type*	*N*	*min(z)*	*Max(z)*
Adults only	2. under15=0 & adult=1 & over60=0	473	5.717	9.489
	5. under15=0 & adult=4 & over60=0	1041	5.717	9.489
	3. under15=0 & adult=2 & over60=0	1,243	5.717	9.489
	6. under15=0 & adult=5 & over60=0	231	5.717	9.489
	4. under15=0 & adult=3 & over60=0	1,393	5.717	9.489
Elder only	10. under15=0 & adult=0 & over60=1	679	5.253	9.665
	19. under15=0 & adult=0 & over60=2	884	5.253	9.665
Adult:Elder = 1:1	21. under15=0 & adult=2 & over60=2	134	5.995	9.002
	11. under15=0 & adult=1 & over60=1	568	5.995	9.002
Child:Adult = 1:1	38. under15=1 & adult=1 & over60=0	186	5.164	9.751
	75. under15=2 & adult=2 & over60=0	2404	5.164	9.751
Child:Adult = 1:2	77. under15=2 & adult=4 & over60=0	303	5.577	9.449
	39. under15=1 & adult=2 & over60=0	1367	5.577	9.449
Child:Adult:Elder = 1:1:1	93. under15=2 & adult=2 & over60=2	260	5.590	8.978
	47. under15=1 & adult=1 & over60=1	116	5.590	8.978
Child:Adult = 2:1	74. under15=2 & adult=1 & over60=0	121	5.248	8.965
	147. under15=4 & adult=2 & over60=0	102	5.248	8.965

Table 7-2. Household types and household groups for non-parametric analysis
Source. Author compiled.

The data for the parametric approach is the pooled data set of three years of 2012, 2014 and 2016. We also perform the sensitivity analysis with data for the single year of 2016 which contains single price schedule. The second sensitivity analysis is with panel data in which all monetary terms are adjusted to 2014 price since we use lag of electricity expense as an independent variable.

7.4 Results and discussion

7.4.1 Non-parametric analysis

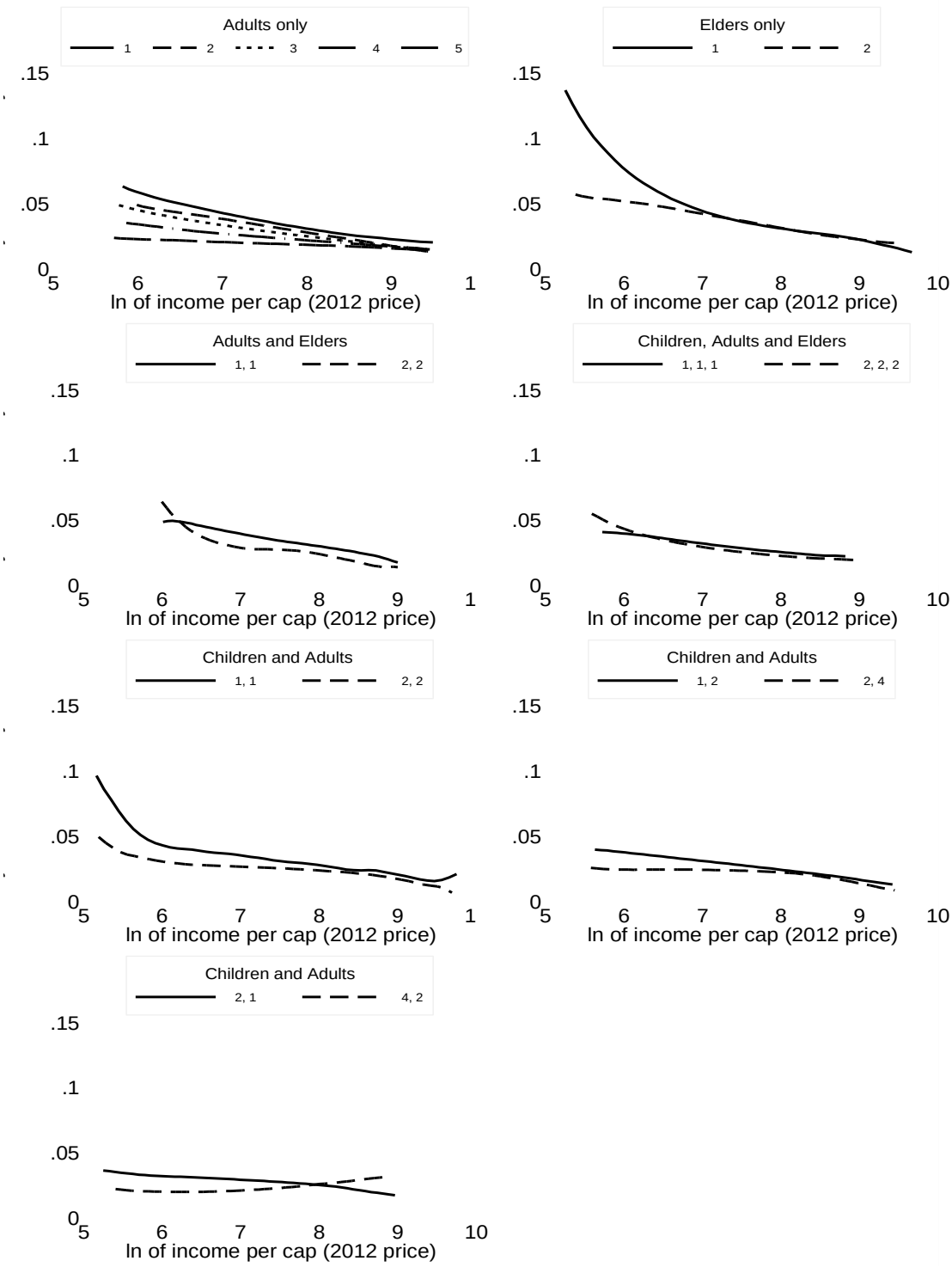


Figure 7-3. Non-parametric analysis for economies of scale
Source. Author compiled.

Figure 7-3 shows the estimates of non-parametric Engel curves for different types of households. Each graph is for a specific group of households. For example, the upper left graph is adult-only households. The upper right is elders-only households.

In each graph, straight lines and dash lines are smaller and larger households of a same household type. The horizontal axis is income per capita and the vertical axis is the ratio of electricity bill over household total income. At a given level of income per capita, if a dash line is below straight line, the household type has economies of scale in electricity expense, and vice versa the household type has diseconomies of scale.

First, except for the households of four children and two adults (in the bottom graph of Figure 7-3), all graphs have downward slopes. This means that if income per capita increases, the ratio of electricity over total household income decreases. In terms of per capita, the increase in electricity expense is slower than the increase in income. In other words, electricity is a necessary good which is in line with theory prediction.

Second, most graphs show that the straight lines are higher than the dash lines for a majority of the range of income per capita. This means that at a given income per capita, the ratio of electricity bills over total income of smaller households is higher than that of larger households with the same composition. This gives initial support for the existence of economies of scales for household electricity expenditure in Vietnam.

Third, in each graph the straight lines and the dash lines converge. There are cross points between straight lines and dash lines at high level of income per capita. This suggests that the economies of scale may not exist for rich households. It is worth noting that the cross points are at the income per capita level of about 8,103 thousand VND per month which is higher than the income level of 95 per cent of the sample. It suggests that the economies of scale may not exist for rich households.

Group	*Type*	$E(w\|z)$	*Bootstrap Std. Err.*
Adult only	2	0.0338	0.0017
	3	0.0299	0.0021
	4	0.0271	0.0024
	5	0.0225	0.0014
	6	0.0181	0.0025
Elder only	10	0.0431	0.0030
	19	0.0376	0.0053
Adult:Elder = 1:1	11	0.0337	0.0021
	21	0.0272	0.0027
Child:Adult = 1:1	38	0.0329	0.0059
	75	0.0250	0.0060
Child:Adult = 1:2	39	0.0267	0.0010
	77	0.0219	0.0033
Child:Adult = 2:1	74	0.0289	0.0372
	147	0.0214	0.0039
Child:Adult:Elder = 1:1:1	47	0.0293	0.0148
	93	0.0277	0.0015

Note. The type is detailed in Table 7-2

Table 7-3. The average electricity share weighted by the density of income per capita

Source. Author estimated.

Since the existence of cross-points in some graph cause trouble for interpreting results, we present the estimates of the Engel curves in the form of statistic for comparison. We calculate the $E(w|z)$ across household type as in Table 7-3.

Table 7-3 shows the weighted average of income fraction spent on electricity ($E(w|z)$) for different household types as the formula (7-6) in section 7.3.1. The weight is kernel density of income per capita. In Table 7-3, households in a same group are households having same composition but different size. Each size is corresponding to a household type. Household types in a same group is arranged in the pattern of ascending order of household size. In general, there is a clear pattern that within in a same group, the higher the household size, the lower the $E(w|z)$, supporting the hypothesis that economies of scale do exist in electricity expenditure at household level.

7.4.2 Parametric analysis

	OLS (lneleclm1)		*First difference (D.lneleclm1)*	
Income				
lnincpc1	0.2294***	[0.0084]		
Head characteristics				
male	-0.0117	[0.0101]	-0.0011	[0.0103]
age	0.0013***	[0.0004]	0.0016***	[0.0004]
Demographics				
lnhhsize	0.4955***	[0.0150]	0.4798***	[0.0148]
lnsexratio	0.0051***	[0.0011]	0.0055***	[0.0011]
lnchildrenratio	0.0009*	[0.0005]	0.0013**	[0.0005]
lnelderratio	0.0000	[0.0005]	-0.0003	[0.0005]
Household highest degree (base is no education)				
Primary	0.0750***	[0.0209]	0.0740***	[0.0201]
Secondary	0.1295***	[0.0204]	0.1187***	[0.0201]
High school	0.1538***	[0.0213]	0.1231***	[0.0209]
A degree	0.1264***	[0.0231]	0.1300***	[0.0228]
Housing condition				
lnpcsqm	0.1232***	[0.0085]	0.1290***	[0.0087]
ownership	-0.0709***	[0.0239]	-0.0690***	[0.0224]
Appliances				
aircon	0.2527***	[0.0140]	0.2692***	[0.0148]
fan	0.1155***	[0.0160]	0.0948***	[0.0156]
fridge	0.5222***	[0.0109]	0.5128***	[0.0108]
washing	0.1191***	[0.0110]	0.1162***	[0.0116]
oven	0.0624***	[0.0167]	0.0802***	[0.0176]
computer	0.0856***	[0.0116]	0.0901***	[0.0119]
tv	0.2761***	[0.0202]	0.2501***	[0.0189]
waterheater	0.0931***	[0.0118]	0.1117***	[0.0122]
Related goods				
generator	0.0507*	[0.0291]	0.0469*	[0.0281]
lpg1	0.0006***	[0.0001]	0.0006***	[0.0001]
Temperature and Admin				
cdd25	0.0010***	[0.0001]	0.0010***	[0.0001]

	OLS (lneleclm1)		*First difference (D.lneleclm1)*	
urban	0.1414***	[0.0099]	0.1284***	[0.0104]
year_2014	0.0653***	[0.0107]	0.0553***	[0.0121]
year_2016	0.1864***	[0.0101]	0.1924***	[0.0115]
_cons	1.3065***	[0.0750]	-0.0048	[0.0052]
N	18971		18670	
R-sq	0.727		0.568	
adj. R-sq	0.726		0.566	
F	428.9441		208.6091	

Note. Standard errors in brackets; * p<0.10, ** p<0.05, *** p<0.01

Table 7-4. Estimation of parametric approach for economies of scale in electricity expense

Source. Author compiled.

lneleclm	*Panel model*		*2016 model*	
L.lneleclm2	0.1685***	[0.0608]		
Income				
lnincpc2	0.3329***	[0.0268]		
Demographics				
lnhhsize	0.5729***	[0.0458]	0.4879***	[0.0223]
lnsexratio	0.0023	[0.0033]	0.0053***	[0.0015]
lnchildrenratio	0.0018	[0.0015]	0.0004	[0.0008]
lnelderratio	0.0003	[0.0014]	-0.0005	[0.0008]
N	3646		8974	
adj. R-sq			0.555	

Note. Standard errors in brackets; * p<0.10, ** p<0.05, *** p<0.01
The panel model is estimated with the Blundell and Bond (1998) estimators.

Table 7-5. Sensitivity analysis for economies of scale in electricity expense

Source. Author compiled.

❖ *Control variables*

Overall, the estimated result for control variables is as expected. Since price is IBTs, the electricity consumption and electricity expense move in the same direction. Thus estimated coefficients of all control variables are expected to be similar in term of sign and significance with the estimated coefficients of electricity demand function provided in Chapter 5. Asset owning, temperature (cdd25), living in urban areas, dwelling size (square meter per capita) have a positive impact on household electricity expense. Households who rent their dwelling have higher electricity expenses. Household head's gender does not impact on electricity expense.

❖ *Household composition*

From Table 7-4, there is no significant difference in electricity expense between an adult and an elder at the level of 0.1. However, electricity expense of a child is slightly higher than that of an adult at 0.05 level. The former result is as expected, but the latter is not. As discussed in the impacts of household composition on electricity consumption (Chapter 5), there is empirical evidence supporting the idea that there is no significant difference across a child, an adult and an elder in term of kWh use. The changes in household composition do not change the total electricity consumption, other things being equal. The prices that are applied to the household

do not change. Obviously, in this case the electricity expenditure of a child, an adult and an elder should be the same.

Since we have used the same data for both models of electricity use and electricity expense, the explanation for the difference in result is the difference in modelling and econometric technic. Indeed, with the OLS model, we still have insignificant difference between a child and an adult at the level of 0.05. Sensitivity analysis with data of 2016 and panel data (Table 7-5) also confirm that there is no significant difference between a child and an adult in electricity expense at the level of 0.05.

❖ *Household sizes – economies of scale*

With regard to the key variable of the chapter, the estimated result shows that when a household doubles while keeping the same composition and resources, household electricity expense increases 47 per cent. This implies that, in general, households do enjoy economies of scale in electricity expense regardless of increasing block tariffs. In other words, the quantity effect of an increase in household size dominates the price effect.

This result may come from the fact that a large fraction of the data are households with small and medium sizes. Households with four members and less account for 72 per cent of the sample. Households with six members and less account for 95 per cent of the sample. Under an increasing block tariff, the larger a family, the higher electricity consumption following by a higher average price the family must pay. As consequence, the larger a family is, the larger the price effect is. In other words, the price effect on small or medium size households is smaller than on large size households. In the case that a large fraction of the sample is small and medium size households, it is reasonable to have quantity effect dominating price effect.

7.5 Conclusion

Electricity is a kind of public goods that can be shared among household members thus providing the potential for economies of scale. So far, much has been done to confirm the economies of scale in electricity consumption. In fact, estimates from Chapter 5 also confirm the hypothesis. However, the economies of scale in electricity consumption can be ruled out under an IBTs schedule. An IBTs schedule pushes the average price up if household size increases. The increase in average price can cancel out all the economies of scale from sharing quantity. So far, there is no research on economies of scale of electricity expense under an IBTs schedule. This chapter investigate that issue in Vietnam context during 2012, 2014 and 2016.

The estimates from two methods of non-parametric and econometric as well as sensitivity analysis with different data provide evidence supporting the idea that from 2012–16 Vietnamese households did enjoy economies of scale in electricity expenditure regardless of IBTs in electricity price. When a household doubles, keeping the same composition and resources, household electricity expense increases 47 per cent. The reason may come from the fact that the household sizes in data is quite small with the mean about four members a household. The small size implies that the increase in household size only increases electricity use a small amount. Average price would increase fractionally too which is not enough to rule out the saving quantity from sharing characteristics.

However, the results from non-parametric analysis also reveal an interesting trend that is worth testing and explanation in the future. That is (i) whether there is a convergence in electricity expense per capita of all households having the same composition when income per capita increases and (ii) whether the economies of scale in electricity expenditure disappear after a certain level of high income per capita for certain household types. In other words, whether the

existence of economies of scale in household electricity expenditure depends not only on the nature of price policies but also on the wealth condition of households.

Chapter 8. Heatwaves and residential electricity demand

8.1 Introduction

Heatwaves, a climate change-related phenomenon, cause a severe health risk for people. In 2003, a heatwave killed 11,000 people in France (theguardian.com, 2003). In 2010, a heatwave in Russia killed 10,000 people (Russo, Sillmann and Fischer, 2015). Mitigation measures such as using fans or air conditioners can lessen the health risk. Applying those measures, however, could increase the socio-economic vulnerability of households due to higher electricity expenditure and less disposable income for other goods and services.

This effect is magnified in two channels. The first channel is the urban heat island (UHI) phenomenon. UHI is a phenomenon that urban areas have warmer temperature than their surrounding rural areas (Roth, 2012). Under the UHI effect, households in urban area face higher temperature and thus need higher cooling effort to mitigate the impacts of heatwaves. The second channel is the increasing block tariff scheme (IBTs). Under the IBTs, the higher electricity a household uses, the higher average price the household has to pay, therefore the higher electricity costs for cooling efforts. Both channels increase the bills households pay for mitigating the physical impacts of heatwaves. The increased bills harm poor households and can push near-poor households back to towards the poverty line.

As a tropical country, Vietnam has faced many severe heatwaves in recent years. On June 3, 2017, the temperature measured at Ha Dong station reached 41.5^{o}C breaking the records of the last 40 years (Hoàng Như, 2017). A day later, the temperature in some areas was just under 50°C (soha.vn, 2017). This raised the need for research on the impact of heatwaves on welfare of low-income households to provide a basis for government in designing appropriate policies to help poor households mitigate the physical impacts of heatwaves.

To date, there has been no such research in Vietnam. Thus this chapter aims to explore the impacts of heatwaves on household socio-economic vulnerability via the cooling effort for mitigating the physical-impacts of heatwaves. First, we investigate the marginal effect of a heatwave on household electricity demand. Second, we examine whether the marginal effect differs between rural and urban areas. The results can help to determine (i) the subsidy amount needed for targeted households in mitigating the impact of heatwaves and (ii) the necessity of different subsidy policies for urban and rural areas.

8.2 Literature review

Heatwaves are extreme weather phenomenon which occur when the ambient temperature is extremely high for a continuous period. Therefore, exploring the impacts of heatwaves on household welfare via cooling efforts requires an exploration of the relationship between temperature and electricity demand.

So far, there have been many researches on this relationship. These researches are to estimate electricity demand function elec_demand = f(temperature). There is a wide consensus that the relationship is non-linear since electricity demand increases with both low and high temperature. The differences among the researches relate to the function forms and ways to represent temperature.

In terms of function forms, the two common forms are parametric and non-parametric. The parametric approach considers the relationship as V-shape and estimates two separate slopes for cold and hot temperatures (Valor, Meneu and Caselles, 2001; Pardo, Meneu and Valor, 2002). The non-parametric does not impose any specific form for f(.). The f(.) is determined by data through a series of specification tests (Engle *et al.*, 1986; Henley and Peirson, 1997).

In terms of temperature representation, two possibilities are to use temperature directly and to use derived functions of temperature. The derived functions are cooling degree days (CDD)/ heating degree days (HDD), or transform temperature to many intervals. CDD is the amount of temperature that needs to be cooled down to reach a certain base temperature. The certain base θ is the level of thermal comfort. CDD for a day *i* is $\max(T_i - \theta, 0)$. CDD for a month is the sum of CDD for all days of a month.

Temperature representation is needed to get the direct observation for f(.). In other words, the time duration of temperature needs to be matched with the time duration of electricity load. Normally, temperature is measured daily or hourly. If electricity load is also measured daily or hourly, temperature can be used directly as an independent variable in f(.) as Henley and Peirson (1997) did. In a special case, temperature and electricity load are both measured daily. However, electricity load is not regionally disaggregatable. The temperature will be transformed to population-weighted temperature index since the higher the population, the higher the influence of temperature on electricity demand (Valor, Meneu and Caselles, 2001).

Derived function of temperature is used when electricity data is not daily but monthly data. If billing periods are fixed for all observation, temperature will be converted to CDD/HDD of a month. If billing periods differ across customers, as is the case with billing cycles, temperature can be transformed to many intervals (Engle *et al.*, 1986). According to Engle *et al.* (1986), each billing cycle is 30 days and different customers pay bills at different days in a month. In this case, electricity bills and temperature in the same month are unmatched. For example, a household pays its electricity bill on March 23 for the billing cycle from February 23 to March 23. The bill is in March but does not relate to the temperature on March 25. In that case, temperature of a month is discretized to many intervals. For each interval, electricity load is defined as the product of number of customers faced the temperature in a month and average kWh each customer consumed in a month. The former term is the total of customers of all billing cycles covering those days of a month having temperature falling in the temperature intervals. The latter term is the total kWh consumption of all billing cycles in a month divided by the total number of customers related to the billing cycles.

So far, the use of the derived function of temperature is more popular than the direct temperature due to data availability since electricity data is normally measured monthly. However, the derived function is not as good as direct temperature in reflecting the relationship between electricity load and temperature. First, the choice of the base for CDD/HDD is quite arbitrary. The level of thermal comfort is subjective. For example, Sailor and Munoz (1997) chose the level of 21˚C; Valor *et al.* (2001) chose 18˚C, Zachariadis and Pashourtidou (2007) chose 18˚C and 22˚C for HDD and CDD respectively.

Second, while CDD/HDD are a good proxy for temperature, they do not capture the concept of heatwaves since the summing up of CDD/HDD for a month loses the distribution of daily temperature. Suppose we have two months. In month 1, temperature every day is 26 ˚C. In month 2, the first three days have temperature of 35˚C and the temperature for the rest of the month are stable at 25˚C. By definition, CDD25 at the two months are the same, at 30 degree-days. However, people may have to turn on air conditioning for the first three days of the second month while they don't need air conditioning for the first month.

In this study, we propose to a way to redress the second drawback. Beside degree-day as a function representing temperature as usual, I propose to include a dummy variable for the existence of heatwaves in a month. The definition and calculation of heatwaves is adopted from Russo *et al.* (2015). The CDD/HDD capture the impact of temperature as a whole while the

heatwave dummy captures the impact of extreme value of temperature. It would provide a better estimation for impacts of temperature on electricity demand function.

8.3 Methods and data

8.3.1 Methods

In this chapter, I apply the short-run demand function as in Chapter 5. I also use price IVs and marginal prices as discussed in Chapter 5. The econometric function is as follows.

$$lnE = \alpha_0 + \alpha_1 lny + \alpha_2 lnP + \alpha_3 Z + \alpha_{4j} A_j D_j + \alpha_5 H + \varepsilon \qquad (8\text{-}1)$$

where

E = household electricity consumption
y = household (per cap) income
P = price
Z = related variables such as cdd/hdd, price of gas
A_j = vector of appliances
D_j = dummy vars = 1 if the household owns the asset j, = 0 otherwise
H = heatwave dummy = 1 if the month has heatwaves, = 0 otherwise

We introduce an additional dummy variable to capture the concept of heatwaves (H) which takes the value of 1 if there is a heatwave in the month before the survey month and takes the value of 0 otherwise.

The existence of a heatwave in a month is adopted from Russo *et al.* (2015). Russo *et al.* (2015) define a heatwave as at least three continuous days having maximum temperatures greater than daily thresholds. The daily threshold of a day is defined as the 90th percentiles of the maximum temperature of the 15 days before and 15 days after the day for all available years.

It should be noted that the Russo *et al.* (2015) definition is different from the official definition of heatwaves in Vietnam which is as follows.

> *There are three levels of disaster risk due to heatwaves*
> *1. Level 1 includes the following:*
> *a) a heatwave with the highest daily temperature up to 39°C or 40°C, lasting from 5 to 10 days;*
> *b) a heatwave with the highest daily temperature above 40°C, lasting from 3 to 5 days.*
> *2. Level 2 includes the following:*
> *a) a heatwave with the highest daily temperature up to 39°C or 40°C, lasting more than 10 days;*
> *b) a heatwave with the highest daily temperature above 40°C, lasting from 5 to 10 days.*
> *3. Level 3 includes a heatwave with the highest daily temperature above 40°C, lasting more than 10 days.* (Nguyễn Tấn Dũng, 2014a)

The official definition of heatwaves in Vietnam uses absolute daily thresholds of 39˚C or 40˚C. With this definition, heatwaves only exist in summer. In contrast, the Russo definition uses comparative daily thresholds. From Russo's point of view, each period has its own comfortable range of temperature. If temperature of a day surpasses the range, it can be considered as heatwaves. With this definition, heatwaves can exist at any time of the year, even in winter. For example, the temperature in winter is normally from 15˚C to 21˚C. If three days in a row have the maximum temperatures of 23˚C, it is abnormal for people during that period and is considered as a heatwave though the 23˚C is less than the base line of CDD25.

The Russo definition allows a better capture of the impacts of high temperatures on household electricity demand than the official definition does. Consider the previous example, the 23˚C which is already out of the comfortable temperature range in winter likely to induce people to use fan or AC to get back to the comfortable range of the period. However, official definition does not recognize it as a heatwave and thus does not take into account any impact of the temperature on household electricity consumption. Meanwhile, the Russo definition does recognize it as a heatwave and thus can capture the impact on household electricity consumption.

8.3.2 Data

All the data except for the heatwave dummy are from the pooled data set.

The data of temperature we use to define heatwave dummy variable are the daily temperature of the 14 stations from January 1, 2010 to July 16, 2018. The data is extracted from GHCN (NOAA).

The main obstacle when deriving the heatwave dummy variable is the missing data for maximum temperature. Of all 25,154 records of daily temperature for all 14 stations, 11,124 records have missing data for maximum temperature.

The first attempt is to interpolate the maximum temperature from minimum and average temperature. The original data have three variables of temperature including daily min, max and average temperature. If the relationship among the three temperatures are known precisely, Tmax can be calculated from the data of Tmin and Tavg. I check the following simplest relationship.

$$T_{avg} = (T_{max} + T_{min})/2 \qquad (8\text{-}2)$$

A derived Tavg (tavg1) is calculated from the actual Tmax and Tmin as (8-2). A paired t-test was conducted to compare the tavg1 and the actual Tavg. There was a significant difference in score for tavg1 (M=25.22, sd=5.06) and for tavg (M=24.85, sd=4.96) at the level of 0.05 (t=71.13, p-value=0). These results suggest that the desired Tavg is higher than the actual Tavg.

However, the value of the difference is quite small (the mean different is 0.37˚C). The histogram of the difference also shows that the majority range of the difference is less than 1.0˚C. Thus it is reasonable to accept the equation (8-2) to interpolate the maximum temperature. The interpolation process replaces the desired Tmax for 5,060 records. After the process, the ratio of missing data reduces to 6,064 over 25,154 records.

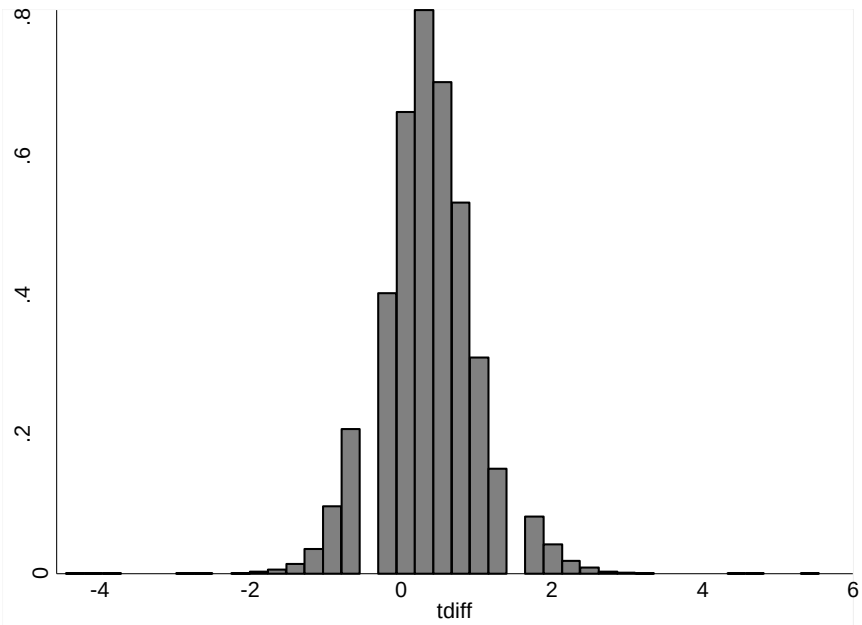


Figure 8-1. Histogram of the difference between desired Tavg and actual Tavg
Source. Author complied.

The missing maximum temperature lead to the missing data for the heatwave dummy variable. Let's denote a day having maximum temperature surpassing daily threshold as 1. Heatwave variable can be missing in three patterns. First, heatwaves are missing due to the missing max temperature of a single day such as "11.", "1.1" or .11". Second, heatwaves are missing due to the missing max temperature of two days such as "1..", ".1." or "..1". Third, heatwaves are missing due to the missing max temperature of three days in a row.

For each month, the heatwave dummy is defined by the following rule.

- If the month has at least one heatwave then H = 1
- If the month does not have any heatwave as well as does not have any missing heatwave then H = 0
- If the month does not have any heatwave but has at least one of the three patterns of missing heatwaves described above, then heatwave dummy is missing (H=.)

After merging with the pooled data set, missing data for heatwaves account for 52 per cent of the whole sample.

Heatwave	*Freq.*	*Percent*	*Cum.*
0	5,797	30.04	30.04
1	3,443	17.84	47.88
.	10,056	52.12	100
Total	19,297	100	

Table 8-1. Descriptive of heatwave dummy variable
Source. Author compiled.

8.4 Results and discussion

All three models are consistent in that cdd25 has a positive impact on household electricity demand. In all three models, at the level of 0.05, if cdd25 increase 30 degree-days (1 degree each day), the electricity consumption of a household increases 4.5 per cent at national level, 4.8 per cent in urban area and 3.3 per cent in rural area.

lnkwh1	*National*	*Urban*	*Rural*
Income			
lnincpc1	0.2638***	0.1904***	0.2486***
	[0.0167]	[0.0213]	[0.0192]
Prices			
lnmp1	-1.4327***	-0.0048	-1.8237***
	[0.2919]	[0.2306]	[0.4025]
lnrsp1	0.0005	0.0011	0.0027
	[0.0015]	[0.0026]	[0.0020]
Temperature			
cdd25	0.0015***	0.0016***	0.0011***
	[0.0001]	[0.0002]	[0.0002]
heatwave	0.0162	-0.0362*	0.0431**
	[0.0143]	[0.0209]	[0.0185]
N	9535	2712	6823
R-sq	0.592	0.711	0.536
adj. R-sq	0.588	0.701	0.53

Note. The estimation is without weight due to the significant amount of missing data on heatwaves. The table presents estimates of selected variables, the full result is in Appendix F.
* p<0.10, ** p<0.05, *** p<0.01

Table 8-2. Estimates of impacts of heatwaves on electricity demand

Source. Author estimated.

The beta coefficients of the heatwave variable are a bit strange with three different patterns in the three models. In the national model, the beta coefficient of heatwave is positive but insignificant even at the level of 0.1. In the model for urban, the beta coefficient of heatwave is negative, insignificant at the level of 0.05, but significant at the level of 0.1. Meanwhile, the beta coefficient of heatwave in rural area is positive and significant at the level of 0.05 level.

The positive and significance of the beta coefficient in rural area means that, other things being equal, a household increases its electricity use 4.3 per cent if a heatwave arises. It should be noted that the variable heatwave is designed independently from cdd25, thus the condition, other things being equal, includes the cdd25. This means that the 4.3 per cent still exists even if two months have the same level of cdd25. It provides evidence for the claim that cdd25 by itself may not be enough to capture the impact of temperature.

In the case of the urban model, the estimate results contradict our hypothesis. We predicted that due to the urban heat island, households in urban area would increase their electricity consumption exponentially in a heatwave. However, the estimates show that households do not change their electricity consumption if the significance level is 0.05. At the significance level of 0.1, households even decrease their electricity consumption. The only explanation we can think of is that under a heatwave, people tend to gather together in one place to use AC. In day time, elders and children go to commercial centers to enjoy AC. Adults tend to stay late in their offices which are equipped with AC. In night time, people in a family gather together in a room to turn on AC rather than turning on lights and fans in three other rooms as they do without heatwaves. Nevertheless, it would be worth conducting a qualitative survey to explain this strange pattern.

Since we accept the results of urban and rural models, the estimated result of national model is reasonable. The estimated coefficient of urban model is in opposite sign of the rural model.

Thus it is likely that the estimated coefficient of heatwave in national model is insignificant at the level of 0.05.

8.5 Conclusion

This chapter has investigated the impacts of heatwaves on residential electricity demand. Heatwaves have been getting more and more serious recently. However, researches on electricity demand so far has not captured fully the impact of heatwaves. Traditional approach represents the impact high temperature have through the concept of CDD. The approach does not cover the concept of heatwaves which are the extreme value of temperature over a few days in a row since it neglects the distribution of high temperature over a month. Therefore, in this chapter I propose to use an additional dummy variable to capture that extreme distribution. The way to define a heatwave is based on Russo *et al.* (2015)

Econometric analysis is performed on Vietnamese data of 2012, 2014 and 2016 for three samples of national, urban and rural areas. Though we have differences in the response patterns of electricity consumption to heatwaves in three samples, we have evidence in rural model that cooling degree day by itself is insufficient to capture the impacts of temperature on electricity demand. A variable that captures the extreme distribution of temperature is needed to fill in the gap. Besides, the results also suggest a need for further qualitative investigation with the behavior of households in urban area during heatwaves.

Chapter 9. Conclusion

❖ *Major findings*

The thesis aims to explore the factors impacting on Vietnamese residential electricity demand in the period 2012–16. The exploration focuses on four main factors which are prices, income, demographics (including household size and composition), and heatwaves. The data is the three rounds of the micro survey VHLSS in 2012, 2014 and 2016. By estimating a common form of demand function with special zoom in each factor, we get the following results.

For the price factor, the thesis examines the impact of the increasing block tariff schedule on residential demand. The two aims are: (i) to estimate the price elasticity of demand in short-run and long-run and (ii) to identify whether households respond to marginal prices or average prices. The estimated results show that (i) households electricity demand is elastic in the short-run. This contradicts studies made in other countries, but agrees with Phu (2017)'s results in Vietnam which uses different data and method. (ii) The long-run elasticity appears positive, a surprise which may be caused by the lack of price variation in the long-run data set. (iii) Households in Vietnam respond fully to marginal price rather than to average price which is in line with Nieswiadomy and Molina (1991) which finds that households tend to respond to marginal price under increasing block tariffs and respond to average price under decreasing block tariffs.

For the income factor, the thesis investigates the non-linear relationship between income and electricity demand and its implication for identifying an electricity poverty threshold. The hypothesis is the existence of an income threshold from which electricity consumption starts to increase with an increase in income. The consumed kWh per capita of households at that income threshold is the electricity poverty threshold. Estimated results are similar to other researches (Douglas F Barnes, Khandker and Samad, 2011; He and Reiner, 2016). Estimated results provide evidence that an electricity poverty threshold does exist. For the whole national sample, the electricity poverty threshold for a household of four people is 51 kWh per month. The threshold is approximately the threshold for the first block in the IBTs schedule in Vietnam. However, the empirical threshold is higher than the threshold for the cash transfer subsidy program. In addition, estimated results also suggest that there is a significant difference in the electricity poverty thresholds between urban and rural areas. The thresholds are about 118 kWh for urban households and 54 kWh for rural households.

For the role of demographic factors, the thesis answers two questions: (i) whether the IBTs schedule cancels out economies of scale in electricity expenditure in Vietnam; and (ii) whether there is a difference in electricity demand across a child, an adult and an elder in Vietnam. The estimates from two methods of non-parametric and econometric, as well as sensitivity analysis, checked with different data, provide evidences supporting the idea that from 2012–16, Vietnamese households did enjoy economies of scale in electricity expenditure regardless of IBTs in electricity price. When a household doubles, keeping the same composition and resources, household electricity expense increases 47 per cent. The results from non-parametric analysis also reveal that: (i) there is convergence in electricity expense per capita of all households having the same composition when income per capita increases; and (ii) the economies of scale in electricity expenditure disappear after a certain level of high income per capita for certain household types.

For the heatwave factor, the thesis aims to demonstrate that cooling degree days (CDD) – a popular way to represent temperature in electricity demand function – is insufficient to capture the full impacts of temperature since it neglects the extreme distribution of temperature which are heatwaves in tropical countries. An additional dummy variable to define a heatwave, based

on Russo *et al.* (2015), is employed to capture that extreme distribution. Although we have differences in the response patterns of electricity consumption to heatwaves in the three samples of national, urban and rural, we have evidence in the rural model supporting the argument that a cooling degree day by itself is insufficient to capture the impacts of temperature on electricity demand. A variable that captures the extreme distribution of temperature is needed to fill in the gap.

❖ *Policy implication*

These results suggest some implications for policy makers in Vietnam in designing an appropriate electricity price system, as well as an appropriate subsidy policy for poor households. First, EVN (2015) proposed three options for the changes in electricity tariff schedule including a single price schedule, a three-block schedule and a six-block schedule. The estimate results show that households respond fully to marginal price and electricity demand responds strongly to price change. This implies that the option of single price schedule is not appropriate if government follows the objective of stimulating electricity savings. It is obvious that increasing block tariff is more appropriate to meet the objective. In addition, since households do respond fully to marginal price, a six-block schedule will not cause any trouble for households in perceiving the price schedule.

Second, the estimation of electricity poverty thresholds provides a reference line for social welfare policy. In 2011, the Vietnamese government implemented an electricity subsidy program in cash transfers for poor households. The program started with 30,000 VND (about USD 1.5) per month in cash transfer for every quarter. Since 2014, the subsidy amount is set to 30 kWh at the prevalent prices. In July 2018, the Ministry of Finance submitted a proposal cancelling the cash transfer program due to budget shortage. The proposal faced a backlash. Thus in October 2018, the Ministry of Finance submitted a draft circular that maintained the cash transfer subsidy program at the threshold of 30 kWh as before. However, the estimated results of this research show that the minimum level of electricity consumption that a household needs to meet its basic needs is 50kWh which is far more than the current subsidy level. To pursue the objective of supporting modern energy access for poor households, the government could consider raising the level of subsidy. Additionally, the government could consider setting the threshold for the first block of the increasing block tariffs to 50kWh and set a low price for the block. In that instance, the IBTs could meet the objective of inducing electricity saving while maintain support for poor households in accessing modern energy.

Third, the estimation provides evidence for EVN that the current IBTs does not cancel out the economies of scale in electricity expenditure of households. The IBTs form of electricity price pushes up the average price for large households which can eliminate the saving from electricity sharing. This can be a penalty for large households which normally are poor households in Vietnam. To redress this issue, government has allowed to separate large households to smaller households (MOIT, 2014, 2018). With evidence from the estimation, it seems that the redress policy meets its objective. There is room for increasing price without penalizing large households which are normally poor households in Vietnam.

❖ *Limitations and future research*

Though this thesis provides meaningful results, it still faces some limitations in data and methodology that must be addressed in future research. For example, while there are 173 weather stations across Vietnam, temperature data are only available at 14, with a significant amount of missing data in maximum temperature causing trouble in identifying heatwaves. Access to data from all 173 weather stations, with full information of daily average and maximum temperature, would help to improve the accuracy of estimation.

In addition, we do not know the accurate prices of a portion of households in data which can cause a certain level of measurement errors. In Vietnam, we know the price schedules applied to 85 per cent of households who buy electricity directly from EVN. The price schedule they pay is the retail price set by government. In contrast, we do not know the price schedule applied to the remaining 15 per cent which are households in rural or remote areas who buy electricity from wholesale organizations. These wholesale organizations buy electricity from EVN at the wholesale tariff schedule. The wholesale tariff schedule is in IBTs form and set by government. We have no information on how those wholesale organizations set their own price schedules. In the data, we do not know which households belong to that portion, as well as the accurate price that the wholesale organizations sell to the households. We propose to apply the retail price schedule to the whole sample to redress the issue. First, the portion is small. Second, since the wholesale price is in IBTs form, wholesale organizations tends to set their retail price in IBTs form too. There is a positive correlation between the retail price set by government and the retail price set by the wholesale organizations. This approach can cause a certain degree of measurement errors. However, the portion of households who can buy electricity directly has been getting larger year by year. In the next few rounds of VHLSS, the phenomenon may no longer exist. We can apply the same methods in this thesis to the updated data to get better estimation.

In terms of methodology, current methods cannot estimate the long-run demand function. The long-run demand function is characterized by the lag of dependent variable in the right-hand side which can cause the dynamic panel bias. In the dynamic panel, there may be a correlation between the lag and unobserved demand shocks in error terms that results in endogeneity. The popular methods to redress the issue, such as Anderson and Hsiao (1982), Arellano and Bond (1991) or Blundell and Bond (1998), all use data at difference form. The use of data in difference form causes the time dimension to become shorter. In our context, VHLSS is rotated surveys which only allows the construction of a panel of a maximum three years of 2012, 2014 and 2016. The three years incidentally correspond to three distinctive price schedules in each year. The redress of the dynamic panel bias collapses data to the single year of 2016 which has just a single price schedule. When there is only a single price schedule (in IBTs form), the relationship between price and kWh consumption is purely the shift of the demand due to the changes in other factors, such as the number of electric appliances. This raises the need for other methods to estimate the long-run demand function.

In addition to this limitation, the results of thesis also suggest some ideas to explore household behaviors in using electricity. A close look at household behaviors should be conducted to explain three points: (i) why there is no difference in electricity demand of an adult with an elder or a child in Vietnam; (ii) why economies of scale in electricity expenditure fades off as income per capita increases; and (iii) why heatwaves seem not to have an impact on electricity demand in urban areas.

References

Agthe, D. E. and Billings, R. B. (1987) 'Equity, price elasticity, and household income under increasing block rates for water', *American Journal of Economics and Sociology*, 46(3), pp. 273–286.

Alberini, A. and Filippini, M. (2011) 'Response of residential electricity demand to price: The effect of measurement error', *Energy economics*, 33(5), pp. 889–895.

Alberini, A., Gans, W. and Velez-Lopez, D. (2011) 'Residential consumption of gas and electricity in the US: The role of prices and income', *Energy Economics*, 33(5), pp. 870–881.

Anderson, T. W. and Hsiao, C. (1982) 'Formulation and estimation of dynamic models using panel data', *Journal of econometrics*, 18(1), pp. 47–82.

Arbués, F., Villanúa, I. and Barberán, R. (2010) 'Household size and residential water demand: an empirical approach', *Australian Journal of Agricultural and Resource Economics*, 54(1), pp. 61–80.

Arellano, M. and Bond, S. (1991) 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *The review of economic studies*, 58(2), pp. 277–297.

Arellano, M. and Bover, O. (1995) 'Another look at the instrumental variable estimation of error-components models', *Journal of econometrics*, 68(1), pp. 29–51.

Barnes, Douglas F., Khandker, S. R. and Samad, H. A. (2011) 'Energy poverty in rural Bangladesh', *Energy Policy*. (Special Section on Offshore wind power planning, economics and environment), 39(2), pp. 894–904. doi: 10.1016/j.enpol.2010.11.014.

Barnes, Douglas F, Khandker, S. R. and Samad, H. A. (2011) 'Energy poverty in rural Bangladesh', *Energy Policy*, 39(2), pp. 894–904.

Barnes, R., Gillingham, R. and Hagemann, R. (1981) 'The short-run residential demand for electricity', *The review of economics and statistics*, pp. 541–552.

BCĐTW (2010) *Báo cáo tổng kết Tổng điều tra dân số và nhà ở 1/4/2009*. Ban Chỉ đạo Tổng điều tra dân số và nhà ở trung ương. Available at: https://www.gso.gov.vn/Modules/Doc_Download.aspx?DocID=11476.

Bernard, J.-T., Bolduc, D. and Belanger, D. (1996) 'Quebec residential electricity demand: a microeconometric approach', *Canadian Journal of Economics*, pp. 92–113.

Bernstein, M. A. and Griffin, J. M. (2006) *Regional differences in the price-elasticity of demand for energy*. Citeseer.

Billings, R. B. and Agthe, D. E. (1980) 'Price elasticities for water: a case of increasing block rates', *Land Economics*, 56(1), pp. 73–84.

Blundell, R. and Bond, S. (1998) 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of econometrics*, 87(1), pp. 115–143.

Bộ Tài Chính (2018) *Dự thảo Thông tư Quy định thực hiện chính sách hỗ trợ tiền điện cho hộ nghèo và hộ chính sách xã hội*. Available at:

http://www.mof.gov.vn/webcenter/portal/mbtc/r/m/fd_btccd/pduthvaba/pduthdalayka5_chitiet?id=14051&_afrLoop=26611242713907277#!%40%40%3F_afrLoop%3D26611242713907277%26id%3D14051%26rightWidth%3D0%26showFooter%3Dfalse%26showHeader%3Dfalse%26_adf.ctrl-state%3D18dro1arha_4 (Accessed: 18 February 2019).

Bond, S. R. (2002) 'Dynamic panel data models: a guide to micro data methods and practice', *Portuguese economic journal*, 1(2), pp. 141–162.

Branch, E. R. (1993) 'Short run income elasticity of demand for residential electricity using consumer expenditure survey data', *The Energy Journal*, pp. 111–121.

Bravo, V. *et al.* (1983) 'A first approach to defining basic energy needs', *United Nations University, Japan, UNUP-454*.

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

Browning, M., Chiappori, P.-A. and Lewbel, A. (2013) 'Estimating consumption economies of scale, adult equivalence scales, and household bargaining power', *Review of Economic Studies*, 80(4), pp. 1267–1303.

Cameron, A. C. and Trivedi, P. K. (2010) *Microeconometrics using stata*. Stata press College Station, TX.

Châu Anh (2015) 'Chuyên gia chê biểu giá điện của EVN', 19 September. Available at: http://www.tienphong.vn/kinh-te/chuyen-gia-che-bieu-gia-dien-cua-evn-911177.tpo (Accessed: 11 September 2017).

Chicoine, D. L. and Ramamurthy, G. (1986) 'Evidence on the specification of price in the study of domestic water demand', *Land Economics*, 62(1), pp. 26–32.

Cragg, J. G. and Donald, S. G. (1993) 'Testing identifiability and specification in instrumental variable models', *Econometric Theory*, 9(2), pp. 222–240.

Deaton, A. and Paxson, C. (1998) 'Economies of scale, household size, and the demand for food', *Journal of political economy*, 106(5), pp. 897–930.

Dharmaratna, D. and Harris, E. (2012) 'Estimating residential water demand using the Stone-Geary functional form: the case of Sri Lanka', *Water resources management*, 26(8), pp. 2283–2299.

Đình Dũng (2015) 'Hội thảo cơ cấu biểu giá điện: Nhiều quan điểm đồng nhất', 22 September. Available at: http://baocongthuong.com.vn/hoi-thao-co-cau-bieu-gia-dien-nhieu-quan-diem-dong-nhat.html (Accessed: 11 September 2017).

Đỗ Hoàng Anh Tuấn (2015) *Thông tư 190/2014/TT-BTC Quy định thực hiện chính sách hỗ trợ tiền điện cho hộ nghèo và hộ chính sách xã hội*. Available at: http://vbpl.vn/tw/Pages/vbpq-toanvan.aspx?ItemID=46166&Keyword=190/2014/TT-BTC (Accessed: 15 September 2018).

DTI, D. of T. and I. (2007) *Fuel Poverty, The National Achives*. Available at: http://webarchive.nationalarchives.gov.uk/20070603164901/http://www.dti.gov.uk/energy/fuel-poverty/index.html (Accessed: 17 September 2018).

Dubin, J. A. and McFadden, D. L. (1984) 'An econometric analysis of residential electric appliance holdings and consumption', *Econometrica: Journal of the Econometric Society*, pp. 345–362.

Engle, R. F. *et al.* (1986) 'Semiparametric estimates of the relation between weather and electricity sales', *Journal of the American statistical Association*, 81(394), pp. 310–320.

Espey, J. A. and Espey, M. (2004) 'Turning on the lights: A meta-analysis of residential electricity demand elasticities', *Journal of Agricultural and Applied Economics*, 36(1), pp. 65–81.

Estes, E. and Honore, B. (1995) 'Partial regression using one nearest neighbor', *V Un $ published Manuscript, Princeton University*.

EVN (2015) *Đề án cải tiến cơ cấu biểu giá*. Available at: http://minhbach.moit.gov.vn/App_File/userfiles/files/De%20an%20cai%20tien%20co%20cau%20bieu%20gia-full.pdf (Accessed: 15 September 2017).

Fan, J. and Gijbels, I. (1991) *Local linear smoothers in regression function estimation*. Department of Statistics [University of North Carolina at Chapel Hill].

Fell, H., Li, S. and Paul, A. (2014) 'A new look at residential electricity demand using household expenditure data', *International Journal of Industrial Organization*, 33, pp. 37–47.

Filippini, M. and Pachauri, S. (2004) 'Elasticities of electricity demand in urban Indian households', *Energy policy*, 32(3), pp. 429–436.

Fisher, F. M. and Kaysen, C. (1962) *A study in econometrics: The demand for electricity in the United States*. North-Holland Pub. Co.

Foster, H. S. and Beattie, B. R. (1981) 'Urban residential demand for water in the United States: reply', *Land Economics*, 57(2), pp. 257–265.

Foster, V., Tre, J.-P. and Wodon, Q. (2000) 'Energy prices, energy efficiency, and fuel poverty', *Unpublished paper. Latin America and Caribbean Regional Studies Program, Washington, DC: The World Bank*. Available at: http://www.mediaterre.org/docactu,bWF4aW0vZG9jcy9wZTE=,1.pdf (Accessed: 30 September 2016).

Frondel, M. and Kussel, G. (2018) 'Switching on electricity demand response: Evidence for German households'.

Frondel, M., Kussel, G. and Sommer, S. (2019) 'Heterogeneity in the price response of residential electricity demand: A dynamic approach for Germany', *Resource and Energy Economics*, 57, pp. 119–134.

Furuta, H. (2014) *Users' Manual for Handling Resampled Micro Data of Vietnam Household Living Standart Survey*. Available at: http://www.sinfonica.or.jp/information/research/vietnam/Manual_VHLSS_overall_v1.1.pdf (Accessed: 22 February 2018).

Garcia-Cerrutti, L. M. (2000) 'Estimating elasticities of residential energy demand from panel county data using dynamic random variables models with heteroskedastic and correlated error terms', *Resource and Energy economics*, 22(4), pp. 355–366.

Gaudin, S., Griffin, R. C. and Sickles, R. C. (2001) 'Demand specification for municipal water management: evaluation of the Stone-Geary form', *Land Economics*, 77(3), pp. 399–422.

GSOVN (2008) *Phương án Khảo sát mức sống hộ gia đình và Điều tra quyền số chỉ số giá tiêu dùng năm 2008*.

GSOVN (2010) *Phương án khảo sát mức sống dân cư năm 2010*.

GSOVN (2011a) 'Hội nghị huấn nghiệp vụ Khảo sát mức sống hộ gia đình (VHLSS) 2012, Điều tra Giá sinh hoạt theo không gian (SCOLI) 2012, Bộ chỉ tiêu thống kê phát triển giới của quốc gia, chỉ tiêu trẻ em và phụ nữ'. Available at: http://www.gso.gov.vn/default.aspx?tabid=382&ItemID=12061 (Accessed: 16 January 2018).

GSOVN (2011b) *Phương án khảo sát mức sống dân cư năm 2012*.

GSOVN (2013) *Phương án Khảo sát mức sống dân cư và Điều tra quyền số chỉ số giá tiêu dùng năm 2014*.

GSOVN (2015) *Phương án Khảo sát mức sống dân cư năm 2016*. Available at: https://thuvienphapluat.vn/van-ban/Van-hoa-Xa-hoi/Quyet-dinh-1095-QD-TCTK-khao-sat-muc-song-dan-cu-2016-2015-304223.aspx.

GSOVN (2018) 'Annual average consumer price index, gold, USD price index and core inflation (Previous year = 100)'. Available at: https://www.gso.gov.vn/default_en.aspx?tabid=780 (Accessed: 22 July 2018).

Ha-Duong, M. (2016) 'Affordable and reliable power for all in Vietnam'. *Renewable Energy Week in Vietnam 2016*, Can Tho and Ha Noi. Available at: https://minh.haduong.com/files/HaDuong-2016-Presentation-AffordableReliablePower-EN.pdf (Accessed: 26 February 2019).

Halvorsen, R. (1973) *Demand for electric power in the United States*. Institute for Economic Research, University of Washington.

H.Anh (2018) 'Bộ Tài chính muốn bỏ chính sách cho không tiền điện với người nghèo', *dantri.com.vn*, 24 July. Available at: read:https://dantri.com.vn/kinh-doanh/bo-tai-chinh-muon-bo-chinh-sach-cho-khong-tien-dien-voi-nguoi-ngheo-2018072400144186.htm (Accessed: 18 February 2019).

Hausmann, J., Kinnucan, M. and McFaddden, D. (1979) 'A two-level electricity demand model: Evaluation of the Connecticut time-of-day pricing test', *Journal of Econometrics*, 10(3), pp. 263–289.

He, X. and Reiner, D. (2016) 'Electricity demand and basic needs: Empirical evidence from China's households', *Energy Policy*, 90, pp. 212–221. doi: 10.1016/j.enpol.2015.12.033.

Henley, A. and Peirson, J. (1997) 'Non-linearities in electricity demand and temperature: parametric versus non-parametric methods', *Oxford Bulletin of Economics and Statistics*, 59(1), pp. 149–162.

Henson, S. E. (1984) 'Electricity demand estimates under increasing-block rates', *Southern Economic Journal*, pp. 147–156.

Hoàng Như (2017) *Nắng nóng Hà Nội vượt kỷ lục 40 năm, zing.vn*. Available at: https://news.zing.vn/nang-nong-ha-noi-vuot-ky-luc-40-nam-post751893.html (Accessed: 21 October 2018).

Holtedahl, P. and Joutz, F. L. (2004) 'Residential electricity demand in Taiwan', *Energy economics*, 26(2), pp. 201–224.

Houthakker, H. S. (1951) 'Some calculations on electricity consumption in Great Britain', *Journal of the Royal Statistical Society. Series A (General)*, 114(3), pp. 359–371.

Houthakker, H. S. and Taylor, L. D. (1970) 'Consumer demand in the United States'.

Houthakker, H. S., Verleger, P. K. and Sheehan, D. P. (1974) 'Dynamic demand analyses for gasoline and residential electricity', *American Journal of Agricultural Economics*, 56(2), pp. 412–418.

Hsiao, C. and Mountain, D. (1985) 'Estimating the short-run income elasticity of demand for electricity by using cross-sectional categorized data', *Journal of the American Statistical Association*, 80(390), pp. 259–265.

Hsing, Y. (1994) 'Estimation of residential demand for electricity with the cross-sectionally correlated and time-wise autoregressive model', *Resource and Energy Economics*, 16(3), pp. 255–263.

Hung, M.-F. and Chie, B.-T. (2013) 'Residential water use: Efficiency, affordability, and price elasticity', *Water resources management*, 27(1), pp. 275–291.

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

Ito, K. (2014) 'Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing', *The American Economic Review*, 104(2), pp. 537–563.

Kavousian, A., Rajagopal, R. and 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, pp. 184–194.

Khandker, S. R., Barnes, D. F. and Samad, H. A. (2012) 'Are the energy poor also income poor? Evidence from India', *Energy Policy*, 47, pp. 1–12. doi: 10.1016/j.enpol.2012.02.028.

Kiviet, J. F. (1995) 'On bias, inconsistency, and efficiency of various estimators in dynamic panel data models', *Journal of econometrics*, 68(1), pp. 53–78.

Kmenta, J. (1997) *Elements of Econometrics: Second Edition*. University of Michigan Press; Subsequent edition.

Le Viet Phu (2017) *Electricity Price, Residential Electricity Demand, and Renewable Energy Development Policies in Vietnam*. Working Paper. Fulbright University Vietnam. Available at: https://fsppm.fuv.edu.vn/attachment.aspx?id=65139 (Accessed: 6 December 2018).

Liebman, J. B. and Zeckhauser, R. J. (2004) *Schmeduling*. Working Paper.

Mai Phương (2015) *Những vấn đề đặt ra trong quản lý điện nông thôn (Issues on rural electricity management), icon - Trang tin điện tử ngành điện*. Available at:

http://icon.com.vn/vn-s83-124254-637/Nhung-van-de-dat-ra-trong-quan-ly-dien-nong-thon-.aspx (Accessed: 15 September 2018).

Martínez-Espiñeira*, R. and Nauges, C. (2004) 'Is all domestic water consumption sensitive to price control?', *Applied economics*, 36(15), pp. 1697–1703.

McFadden, D., Puig, C. and Kirschner, D. (1977) 'Determinants of the long-run demand for electricity', in. *Proceedings of the American Statistical Association*, Business and Economics Section, pp. 109–19.

MOIT, M. of I. and T. (2014) *Circular 16/2014/TT-BCT - Regulation on electricity price*. Available at: https://thuvienphapluat.vn/van-ban/Thuong-mai/Thong-tu-16-2014-TT-BCT-thuc-hien-gia-ban-dien-Bo-Cong-Thuong-232110.aspx.

MOIT, M. of I. and T. (2018) *Circular 25/2018/TT-BCT AMENDING AND SUPPLEMENTING THE CIRCULAR NO. 16/2014/TT-BCT DATED MAY 29, 2014 OF THE MINISTER OF INDUSTRY AND TRADE ON ELECTRICITY PRICES*. Available at: https://thuvienphapluat.vn/van-ban/Tai-chinh-nha-nuoc/Thong-tu-25-2018-TT-BCT-sua-doi-Thong-tu-16-2014-TT-BCT-ve-thuc-hien-gia-ban-dien-394753.aspx (Accessed: 1 April 2019).

MOIT, M. of I. and T. and DEA, D. E. A. (2017) *Vietnam Energy Outlook Report 2017*.

National Assembly (2004) *Electricity law*. Available at: https://thuvienphapluat.vn/van-ban/Doanh-nghiep/Luat-Dien-Luc-2004-28-2004-QH11-18056.aspx.

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

Nguyễn Công Nghiệp (2012) *Công văn Số: 2067/BTC-NSNN Hướng dẫn chi trả kinh phí hỗ trợ hộ nghèo tiền điện*. Available at: https://vanban.luatminhkhue.vn/xem-vb/467/cong-van-2067-btc-nsnn-cua-bo-tai-chinh-ve-viec-huong-dan-chi-tra-kinh-phi-ho-tro-ho-ngheo-tien-dien.aspx (Accessed: 19 September 2018).

Nguyen Tan Dung (2007) *QĐ 16/2007/QĐ-TTg về việc Duyệt quy hoạch tổng thể mạng lưới quan trắc tài nguyên và môi trươngnf quốc giá đến năm 202, 16/2007/QĐ-TTg*. Available at: https://thuvienphapluat.vn/van-ban/Tai-nguyen-Moi-truong/Quyet-dinh-16-2007-QD-TTg-quy-hoach-tong-the-mang-luoi-quan-trac-tai-nguyen-moi-truong-quoc-gia-den-nam-2020-16752.aspx.

Nguyễn Tấn Dũng (2011) *Quyết định 2409/QĐ-TTg về việc kinh phí hỗ trợ hộ nghèo tiền điện*. Available at: https://thuvienphapluat.vn/van-ban/Tai-chinh-nha-nuoc/Quyet-dinh-2409-QD-TTg-kinh-phi-ho-tro-ho-ngheo-tien-dien-133171.aspx (Accessed: 5 June 2017).

Nguyễn Tấn Dũng (2013) *69/2013/QĐ-TTg Quy định về cơ chế điều chỉnh mức giá bán lẻ điện bình quân*. Available at: https://thuvienphapluat.vn/van-ban/Thuong-mai/Quyet-dinh-69-2013-QD-TTg-co-che-dieu-chinh-muc-gia-ban-le-dien-binh-quan-213916.aspx (Accessed: 5 June 2017).

Nguyễn Tấn Dũng (2014a) *Quy định chi tiết, hướng dẫn thi hành một số điều của luật phòng, chống thiên tai, 66/2014/NĐ-CP*. Available at: https://thuvienphapluat.vn/van-ban/Tai-nguyen-Moi-truong/Nghi-dinh-66-2014-ND-CP-huong-dan-Luat-Phong-chong-thien-tai-238656.aspx.

Nguyễn Tấn Dũng (2014b) *Quyết định 60/2014/QĐ-TTg Quy định tiêu chí hộ chính sách xã hội được hỗ trợ tiền điện*. Available at: https://thuvienphapluat.vn/van-ban/Tai-chinh-nha-nuoc/Quyet-dinh-60-2014-QD-TTg-tieu-chi-ho-chinh-sach-xa-hoi-duoc-ho-tro-tien-dien-255839.aspx (Accessed: 18 February 2019).

Nguyễn Tấn Dũng (2015) *Quyết định về việc bổ sung kinh phí 2015 cho các địa phương để hỗ trợ tiền điện hộ nghèo, hô chính sách xã hội*. Decision 936/QĐ-TTg. Prime Minister. Available at: https://thuvienphapluat.vn/van-ban/Tai-chinh-nha-nuoc/Quyet-dinh-936-QD-TTg-2015-kinh-phi-ho-tro-tien-dien-ho-ngheo-ho-chinh-sach-xa-hoi-279571.aspx (Accessed: 19 September 2018).

Nieswiadomy, M. L. and Molina, D. J. (1989) 'Comparing residential water demand estimates under decreasing and increasing block rates using household data', *Land Economics*, 65(3), pp. 280–289.

Nieswiadomy, M. L. and Molina, D. J. (1991) 'A note on price perception in water demand models', *Land economics*, 67(3), pp. 352–359.

Nordin, J. A. (1976) 'A proposed modification of Taylor's demand analysis: comment', *The Bell Journal of Economics*, pp. 719–721.

OECD (no date) *What are equivalent scales?* Available at: http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf.

Opaluch, J. J. (1982) 'Urban residential demand for water in the United States: further discussion', *Land Economics*, 58(2), pp. 225–227.

Opaluch, J. J. (1984) 'A test of consumer demand response to water prices: reply', *Land Economics*, 60(4), pp. 417–421.

Pachauri, S. *et al.* (2004) 'On Measuring Energy Poverty in Indian Households', *World Development*, 32(12), pp. 2083–2104. doi: 10.1016/j.worlddev.2004.08.005.

Pardo, A., Meneu, V. and Valor, E. (2002) 'Temperature and seasonality influences on Spanish electricity load', *Energy Economics*, 24(1), pp. 55–70.

Parti, M. and Parti, C. (1980) 'The total and appliance-specific conditional demand for electricity in the household sector', *The Bell Journal of Economics*, pp. 309–321.

Phung, D. T. and Nguyen, P. (no date) *Vietnam Household Living Standard Survey (VHLSS), 2002 and 2004 - Basic information*. GSOVN. Available at: https://www.google.com.vn/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwi-hMaBg9zYAhXCTrwKHSnCApsQFgguMAE&url=http%3A%2F%2Fsiteresources.worldbank.org%2FINTLSMS%2FResources%2F3358986-1181743055198%2F3877319-1207149468624%2FBINFO_VHLSS_02_04.pdf&usg=AOvVaw00TqOUnjuJRLpC98MF15cJ.

Poulsen, M. and Forrest, J. (1988) 'Correlates of energy use: domestic electricity consumption in Sydney', *Environment and Planning A*, 20(3), pp. 327–338.

Prime Minister (2006) *Decision of Approving the Roadmap and Conditions for Formation and Development of Different Levels of the Electricity Market in Vietnam*. Available at:

https://thuvienphapluat.vn/van-ban/Linh-vuc-khac/Quyet-dinh-26-2006-QD-TTg-lo-trinh-dieu-kien-hinh-thanh-phat-trien-cac-cap-do-thi-truong-dien-luc-Viet-Nam-10848.aspx.

Prime Minister (2013) *Decision of Regulation on Schedule, Conditions and Structure of Electricity Sector for Formulation and Development of Electricity Market Levels in Vietnam*. Available at: https://thuvienphapluat.vn/van-ban/thuong-mai/quyet-dinh-63-2013-qd-ttg-lo-trinh-dieu-kien-co-cau-nganh-dien-phat-trien-cap-do-thi-truong-dien-luc-212741.aspx.

Reiss, P. C. and White, M. W. (2005) 'Household electricity demand, revisited', *The Review of Economic Studies*, 72(3), pp. 853–883.

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

Roodman, D. (2009) 'How to do xtabond2: An introduction to difference and system GMM in Stata', *The stata journal*, 9(1), pp. 86–136.

Roth, M. (2012) 'Urban heat islands', in *Handbook of Environmental Fluid Dynamics, Volume Two*. CRC Press, pp. 160–177.

Russo, S., Sillmann, J. and Fischer, E. M. (2015) 'Top ten European heatwaves since 1950 and their occurrence in the coming decades', *Environmental Research Letters*, 10(12), p. 124003. doi: 10.1088/1748-9326/10/12/124003.

Sailor, D. J. and Muñoz, J. R. (1997) 'Sensitivity of electricity and natural gas consumption to climate in the USA—methodology and results for eight states', *Energy*, 22(10), pp. 987–998.

Salari, M. and Javid, R. J. (2017) 'Modeling household energy expenditure in the United States', *Renewable and Sustainable Energy Reviews*, 69, pp. 822–832.

Sanquist, T. F. *et al.* (2012) 'Lifestyle factors in US residential electricity consumption', *Energy Policy*, 42, pp. 354–364.

Shin, J.-S. (1985) 'Perception of price when price information is costly: evidence from residential electricity demand', *The review of economics and statistics*, pp. 591–598.

Sinnott, R. W. (1984) 'Virtues of the Haversine', *Sky Telesc.*, 68, p. 159.

soha.vn (2017) *Nắng nóng khủng khiếp ở Hà Nội, mẹ mang giường cho con nằm dưới đường trên cao*. Available at: http://soha.vn/toan-canh-nang-nong-khung-khiep-2017-e71681.htm (Accessed: 21 October 2018).

STAIGER, D. and STOCK, J. H. (1997) 'INSTRUMENTAL VARIABLES REGRESSION WITH WEAK INSTRUMENTS', *Econometrica*, 65(3), pp. 557–586.

Stock, J. H. and Yogo, M. (2005) 'Testing for Weak Instruments in Linear IV Regression', *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, p. 80.

Taylor, L. D. (1975) 'The demand for electricity: a survey', *The Bell Journal of Economics*, pp. 74–110.

Taylor, L. D., Gail, B. and Philip, V. (1977) *The residential demand for energy*. Electric Power Research Institute (EPRI, EA-235).

Taylor, L. D. and Houthakker, H. S. (2009) *Consumer demand in the United States: Prices, income, and consumption behavior*. Springer Science & Business Media.

Terza, J. V. (1986) 'Determinants of household electricity demand: A two-stage probit approach', *Southern Economic Journal*, pp. 1131–1139.

theguardian.com (2003) *Heat wave killed 11,000 in France, The Guardian*. Available at: https://www.theguardian.com/world/2003/aug/29/france (Accessed: 21 October 2018).

Thục Quyên (2014) '15 năm dấu ấn điện khi hóa nông thôn', *Báo Tiền Phong*, 7 May. Available at: https://www.tienphong.vn/kinh-te/15-nam-dau-an-dien-khi-hoa-nong-thon-702970.tpo (Accessed: 15 September 2018).

Tiến Hiệp (2018) *Duy trì và đẩy mạnh phong trào "Hộ gia đình tiết kiệm điện" trên địa bàn Thủ đô, EVNHANOI*. Available at: http://evnhanoi.vn/tin-tuc-evnhanoi/tiet-kiem-dien/4682-duy-tri-va-day-manh-phong-trao-ho-gia-dinh-tiet-kiem-dien-tren-dia-ban-thu-do (Accessed: 15 February 2019).

Tonn, B. and Eisenberg, J. (2007) 'The aging US population and residential energy demand', *Energy policy*, 35(1), pp. 743–745.

Valor, E., Meneu, V. and Caselles, V. (2001) 'Daily Air Temperature and Electricity Load in Spain', *Journal of Applied Meteorology*, 40(8), pp. 1413–1421. doi: 10.1175/1520-0450(2001)040<1413:DATAEL>2.0.CO;2.

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

Yatchew, A. (2003) *Semiparametric regression for the applied econometrician*. Cambridge University Press.

Zachariadis, T. and Pashourtidou, N. (2007) 'An empirical analysis of electricity consumption in Cyprus', *Energy Economics*, 29(2), pp. 183–198.

Zhou, S. and Teng, F. (2013) 'Estimation of urban residential electricity demand in China using household survey data', *Energy Policy*, 61, pp. 394–402.

Appendix A. Policy context

From	*To*	*No of block*	*Average price*	*Legal documents*	*Poverty Status*	
					kWh	*Prices*
5-1997	9-1999	5		QD 31/1997/VGCP-TLSX		
10-1999	9-2002	5		QD 193/1999/QĐ-TTg		
10-2002	12-2006	5		QD 124/2002/QĐ-TTg		
1-2007	2-2009	6	842.00	QD 276/2006/QĐ-TTg		
3-2009	2-2010	7	948.50	TT 05/2009/TT-BCT		
3-2010	2-2011	7	1,058.00	TT 08/2010/TT-BCT		
3-2011	12-2011	7	1,242.00	TT 05/2011/TT-BCT	50	993.00
1-2012	6-2012	7	1,304.00	TT 42/2011/TT-BCT	50	993.00
7-2012	12-2012	7	1,369.00	TT 17/2012/TT-BCT	50	993.00
1-2013	7-2013	7	1,437.00	TT 38/2012/TT-BCT	50	993.00
8-2013	5-2014	7	1,508.85	TT 19/2013/TT-BCT	50	993.00
6-2014	3-2015	6	1,508.85	QD 4887/QĐ-BCT		
4-2015	11-2017	6	1,622.01	QD 2256/QĐ-BCT		
12-2017	2-2019	6	1,720.65	QD 4495/QĐ-BTC		
3-2019	present	6	1,864.44	QD 648/QD-BCT		

Note. Unit '000 Vietnamdong; AP: average selling price; VAT excluded.
Poverty status – Special first block and price for poor households.

Table A-1. Evolution of electricity price in Vietnam since 1997

Source. Author compiled.

Appendix B. Data descriptive

Variable	*Description*
year	Survey year
tinh	City code
huyen	District code
xa	Ward code
diaban	Village/ Neighborhood code
hoso	House hold ID in each Neighborhood
urban	1-urban; 0-rural
svmonth	The month conducting the survey with the household
month	Survey month - 1 = elec consumption period
reg8	8-region coding
reg6	6-region coding
hhwt	Household weight
popwt	Population weight
male	The gender of household head; 1 - male 0 - female
age	Age of household head
edu	Education level of household head
hhsize	[derived] Number of people in the household
highestedu	[derived] The maximum education level of people in the household
hhmale	[derived] Number of male members in households
hhover60	[derived] number of members >= 60
hhunder15	[derived] number of members <= 15
hhadult	[derived] number of members from 16-59
sexratio	[derived] Male/hhsize
elderratio	[derived] Ratio of people over 60 in a household
childrenratio	[derived] Ratio of children under 15 in a household
lnhhsize	[derived] ln of household size
lnelderratio	[derived] ln of elder ratio
lnchildrenratio	[derived] ln of children ratio
lnsexratio	[derived] ln of sex ratio
inchhy	Annual household income. Unit. '000 VND
lninchhy	[derived] Ln of Annual household income. Unit. '000 VND
inchhy1	[derived] Annual household income. Unit. '000 VND at 2012 price
lninchhy1	[derived] Ln of Annual household income. Unit. '000 VND at 2012 price
inchhy2	[derived] Annual household income. Unit. '000 VND at 2014 price
lninchhy2	[derived] Ln of Annual household income. Unit. '000 VND at 2014 price
inchhy3	[derived] Annual household income. Unit. '000 VND at 2016 price
lninchhy3	[derived] Ln of Annual household income. Unit. '000 VND at 2016 price
inchh	[derived] Monthly household income. Unit. '000
lninchh	[derived] Ln of Monthly household income. Unit. '000
inchh1	[derived] Monthly household income. Unit. '000 at 2012 price
lninchh1	[derived] Ln of Monthly household income. Unit. '000 at 2012 price
inchh2	[derived] Monthly household income. Unit. '000 at 2014 price
lninchh2	[derived] Ln of Monthly household income. Unit. '000 at 2014 price
inchh3	[derived] Monthly household income. Unit. '000 at 2016 price
lninchh3	[derived] Ln of Monthly household income. Unit. '000 at 2016 price
incpc	Monthly Income per Capita. Unit. '000 VND
lnincpc	[derived] Ln of Monthly Income per Capita. Unit. '000 VND
incpc1	[derived] Monthly Income per Capita. Unit. '000 VND at 2012 price
lnincpc1	[derived] Ln of Monthly Income per Capita. Unit. '000 VND at 2012 price
incpc2	[derived] Monthly Income per Capita. Unit. '000 VND at 2014 price
lnincpc2	[derived] Ln of Monthly Income per Capita. Unit. '000 VND at 2014 price
incpc3	[derived] Monthly Income per Capita. Unit. '000 VND at 2016 price
lnincpc3	[derived] Ln of Monthly Income per Capita. Unit. '000 VND at 2016 price

Variable	*Description*
poor	Government Certified Poverty Status. 1-yes; 0-no
exphh	[derived] Monthly household exp. Unit. '000 VND
lnexphh	[derived] Ln of Monthly household exp. Unit. '000 VND
exphh1	[derived] Monthly household exp. Unit. '000 VND at 2012 price
lnexphh1	[derived] Ln of Monthly household exp. Unit. '000 VND at 20
exphh2	[derived] Monthly household exp. Unit. '000 VND at 2014 price
lnexphh2	[derived] Ln of Monthly household exp. Unit. '000 VND at 20
exphh3	[derived] Monthly household exp. Unit. '000 VND at 2016 price
lnexphh3	[derived] Ln of Monthly household exp. Unit. '000 VND at 20
exppc	[derived] Monthly per cap exp. '000 VND (best)
lnexppc	[derived] Ln of Monthly per cap exp. '000 VND (best)
exppc1	[derived] Monthly per cap exp. '000 VND (best) at 2012 price
lnexppc1	[derived] Ln of Monthly per cap exp. '000 VND (best) at 2012
exppc2	[derived] Monthly per cap exp. '000 VND (best) at 2014 price
lnexppc2	[derived] Ln of Monthly per cap exp. '000 VND (best) at 2014
exppc3	[derived] Monthly per cap exp. '000 VND (best) at 2016 price
lnexppc3	[derived] Ln of Monthly per cap exp. '000 VND (best) at 2016
sqm	Total residential area? m2
flat	Main accommodation? 1-Appartment; 2-Flat
pillar	the main material as poles (or pillars, or carrying walls)
roof	the main material of roof
wall	the material of external wall
dwelling	Type of dwelling
ownership	Who owns the the main housing accommodation
dwprice	Price of the dwelling justified by owners
dwprice1	[derived] Price of the dwelling justified by owners at 2012 price
dwprice2	[derived] Price of the dwelling justified by owners at 2014 price
dwprice3	[derived] Price of the dwelling justified by owners at 2016 price
aircon	Air conditioner(s)
fan	Electric fan(s)
fridge	Refrigerator(s)
washing	Washing machine(s), (clothes-) drying machine(s)
oven	Microwave oven(s), baking oven(s)
computer	Computer(s)
tv	Telivisions
waterheater	(Bath) water heater(s)
generator	Electricity generator(s)
pump	Pumping machine(s)
printer	Printer(s)
fax	Fax machine(s)
mobile	Mobile telephone(s)
video	Video player(s), DVD player(s), digital player(s), satellite antenna
hifi	Hi-fi system of various kinds
radio	Radio/radio-cassette player(s)
diskplayer	Disk player(s)
camera	Camera(s), video recorder(s)
cooker1	Gas cooker(s), magnetic cooker(s)
cooker2	Electric cooker(s), electric rice cooker(s), pressure cooker(s)
vacuum	Vacuum cleaner(s), dehumidifier(s), water filter(s)
juicer	Juice extractor(s), citrus juicer(s)
coal	Coal in last month. Unit. '000 VND
coal1	[derived] Coal in last month. Unit. '000 VND at 2012 price
coal2	[derived] Coal in last month. Unit. '000 VND at 2014 price
coal3	[derived] Coal in last month. Unit. '000 VND at 2016 price
briquette	Coal briquette in last month. Unit. '000 VND

Variable	*Description*
briquette1	[derived] Coal briquette in last month. Unit. '000 VND at 2012 price
briquette2	[derived] Coal briquette in last month. Unit. '000 VND at 2014 price
briquette3	[derived] Coal briquette in last month. Unit. '000 VND at 2016 price
petro	Petroleum in last month. Unit. '000 VND
petro1	[derived] Petroleum in last month. Unit. '000 VND at 2012 price
petro2	[derived] Petroleum in last month. Unit. '000 VND at 2014 price
petro3	[derived] Petroleum in last month. Unit. '000 VND at 2016 price
kerosene	Kerosene in last month. Unit. '000 VND
kerosene1	[derived] Kerosene in last month. Unit. '000 VND at 2012 price
kerosene2	[derived] Kerosene in last month. Unit. '000 VND at 2014 price
kerosene3	[derived] Kerosene in last month. Unit. '000 VND at 2016 price
mazut	Mazut oil in last month. Unit. '000 VND
mazut1	[derived] Mazut oil in last month. Unit. '000 VND at 2012 price
mazut2	[derived] Mazut oil in last month. Unit. '000 VND at 2014 price
mazut3	[derived] Mazut oil in last month. Unit. '000 VND at 2016 price
diesel	Diesel oil in last month. Unit. '000 VND
diesel1	[derived] Diesel oil in last month. Unit. '000 VND at 2012 price
diesel2	[derived] Diesel oil in last month. Unit. '000 VND at 2014 price
diesel3	[derived] Diesel oil in last month. Unit. '000 VND at 2016 price
lpg	LPG in last month. Unit. '000 VND
lpg1	[derived] LPG in last month. Unit. '000 VND at 2012 price
lpg2	[derived] LPG in last month. Unit. '000 VND at 2014 price
lpg3	[derived] LPG in last month. Unit. '000 VND at 2016 price
gas	Natural gas in last month. Unit. '000 VND
gas1	[derived] Natural gas in last month. Unit. '000 VND at 2012 price
gas2	[derived] Natural gas in last month. Unit. '000 VND at 2014 price
gas3	[derived] Natural gas in last month. Unit. '000 VND at 2016 price
firewood	Firewood, husk, sawdust?
firewood1	[derived] Firewood, husk, sawdust? at 2012 price
firewood2	[derived] Firewood, husk, sawdust? at 2014 price
firewood3	[derived] Firewood, husk, sawdust? at 2016 price
agrwaste	Farm by-products? (straw, sugar cane leaves, maize/jute/hemp/seagrass stems,…)
agrwaste1	[derived] Farm by-products? (straw, sugar cane leaves, maize/jute/hemp/seagrass
agrwaste2	[derived] Farm by-products? (straw, sugar cane leaves, maize/jute/hemp/seagrass
agrwaste3	[derived] Farm by-products? (straw, sugar cane leaves, maize/jute/hemp/seagrass
enoth	Other types of fuel?
enoth1	[derived] Other types of fuel? at 2012 price
enoth2	[derived] Other types of fuel? at 2014 price
enoth3	[derived] Other types of fuel? at 2016 price
mainlight	Which is the main lighting in your household?
eleclm	Elec. bill for the last month. Unit: '000 VND
eleclm1	[derived] Elec. bill for the last month. Unit: '000 VND at 2012 price
eleclm2	[derived] Elec. bill for the last month. Unit: '000 VND at 2014 price
eleclm3	[derived] Elec. bill for the last month. Unit: '000 VND at 2016 price
kwh	Elec. in kWh for the last month
kwh1	[derived] kWh for last month from elec bill
lnkwh1	[derived] Ln of derived kWh
elecly	Elec. bills on electricity over the past 12 months. Unit: '000 VND
elecly1	[derived] Elec. bills on electricity over the past 12 months. Unit: '000 VND at
elecly2	[derived] Elec. bills on electricity over the past 12 months. Unit: '000 VND at
elecly3	[derived] Elec. bills on electricity over the past 12 months. Unit: '000 VND at
ensubsidy	Amount of subsidy for energy in the past 12 months. Unit: '000 VND
ensubsidy1	[derived] Amount of subsidy for energy in the past 12 months. Unit: '000 VND at
ensubsidy2	[derived] Amount of subsidy for energy in the past 12 months. Unit: '000 VND at
ensubsidy3	[derived] Amount of subsidy for energy in the past 12 months. Unit: '000 VND at

Variable	*Description*
elecpoor	Elec. has been sufficient to meet needs over the last 30 days?
mp	[derived] Marginal prices - including VAT. Unit. 000 VND
mp1	[derived] [derived] Marginal prices - including VAT. Unit. 000 VND at 2012 price
mp2	[derived] [derived] Marginal prices - including VAT. Unit. 000 VND at 2014 price
mp3	[derived] [derived] Marginal prices - including VAT. Unit. 000 VND at 2016 price
rsp	[derived] Rate Structure Premium = kwhcum*mp-billcum
rsp1	[derived] [derived] Rate Structure Premium = kwhcum*mp-billcum at 2012 price
rsp2	[derived] [derived] Rate Structure Premium = kwhcum*mp-billcum at 2014 price
rsp3	[derived] [derived] Rate Structure Premium = kwhcum*mp-billcum at 2016 price
ap	[derived] Average prices - including VAT. Unit. 000 VND
ap1	[derived] [derived] Average prices - including VAT. Unit. 000 VND at 2012 price
ap2	[derived] [derived] Average prices - including VAT. Unit. 000 VND at 2014 price
ap3	[derived] [derived] Average prices - including VAT. Unit. 000 VND at 2016 price
billcum	[derived] Accumulated bills of all previous block
billcum1	[derived] [derived] Accumulated bills of all previous block at 2012 price
billcum2	[derived] [derived] Accumulated bills of all previous block at 2014 price
billcum3	[derived] [derived] Accumulated bills of all previous block at 2016 price
kwhcum	[derived] Accumulated kWh of all previous block
w	[derived] Elec share of household monthly income in $ term
wkwh	[derived] kWh/household monthly expenditure in 2012 price
pgroup	[derived] Price group id
pcsqm	[derived] Squared mettre per capita
lnpcsqm	[derived] Ln of squared metre per capita
longcen	Center longitude of the ward
latcen	Center latitude of the ward
stationid	The nearest weather station id
station	The nearest weather station name
cdd20	[derived] Cooling Degree Day at the base of 20 oC
cdd21	[derived] Cooling Degree Day at the base of 21 oC
cdd22	[derived] Cooling Degree Day at the base of 22 oC
cdd23	[derived] Cooling Degree Day at the base of 23 oC
cdd24	[derived] Cooling Degree Day at the base of 24 oC
cdd25	[derived] Cooling Degree Day at the base of 25 oC
cdd26	[derived] Cooling Degree Day at the base of 26 oC
cdd27	[derived] Cooling Degree Day at the base of 27 oC
cdd28	[derived] Cooling Degree Day at the base of 28 oC
hdd10	[derived] Heating Degree Day at the base of 10 oC
hdd11	[derived] Heating Degree Day at the base of 11 oC
hdd12	[derived] Heating Degree Day at the base of 12 oC
hdd13	[derived] Heating Degree Day at the base of 13 oC
hdd14	[derived] Heating Degree Day at the base of 14 oC
hdd15	[derived] Heating Degree Day at the base of 15 oC
hdd16	[derived] Heating Degree Day at the base of 16 oC
hdd17	[derived] Heating Degree Day at the base of 17 oC
hdd18	[derived] Heating Degree Day at the base of 18 oC
p12	[derived] coefficient to adjust monetary terms of corresponding year to 2012 prices
p14	[derived] coefficient to adjust monetary terms of corresponding year to 2014 prices
p16	[derived] coefficient to adjust monetary terms of corresponding year to 2016 prices

Table B-1. List of variables
Source. Author compiled.

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Admin					
year		Percent	Cum.		
2012	4,914	25.46	25.46		
2014	5,047	26.15	51.62		
2016	9,336	48.38	100.00		
Month (the month before survey month)		Percent	Cum.		
2011m12	6	0.03	0.03		
2012m2	385	2.00	2.03		
2012m3	833	4.31	6.34		
2012m4	26	0.13	6.47		
2012m5	1,186	6.15	12.62		
2012m6	10	0.05	12.67		
2012m7	2	0.01	12.68		
2012m8	1,139	5.90	18.58		
2012m9	12	0.06	18.64		
2012m10	13	0.07	18.71		
2012m11	1,303	6.75	25.46		
2014m2	2,214	11.47	36.94		
2014m3	43	0.22	37.16		
2014m5	2,293	11.88	49.04		
2014m6	17	0.09	49.13		
2014m7	2	0.01	49.14		
2014m8	238	1.23	50.37		
2014m9	4	0.02	50.39		
2014m11	237	1.23	51.62		
2016m2	2,220	11.50	63.12		
2016m5	2,288	11.85	74.98		
2016m8	2,498	12.95	87.92		
2016m11	2,331	12.08	100.00		
urban		Percent	Cum.		
Rural	12,937	67.04	67.04		
Urban	6,360	32.96	100.00		
reg6		Percent	Cum.		
Red River Delta	4,744	24.58	24.58		
Midlands and Northern Mountains	2,275	11.79	36.37		
Northern and Coastal Central	4,171	21.61	57.98		
Central Highlands	1,097	5.69	63.67		
Southeast	3,437	17.81	81.48		
Mekong Delta	3,573	18.52	100.00		
reg8		Percent	Cum.		
Red River Delta	4,490	23.27	23.27		
East Northern Mtns	2,033	10.54	33.80		
West Northern Mtns	496	2.57	36.37		
North Central Coast	2,294	11.89	48.26		
South Central Coast	1,526	7.91	56.17		
Central Highlands	1,097	5.69	61.86		
Southeast	3,787	19.63	81.48		
Mekong Delta	3,573	18.52	100.00		

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Head characteristics					
Male		Percent	Cum.		
Female	5,195	26.92	26.92		
Male	14,102	73.08	100.00		
Education		Percent	Cum.		
None	3,274	16.97	16.97		
Primary School	4,753	24.63	41.60		
Secondary School	5,744	29.77	71.36		
High School	3,070	15.91	87.27		
A degree	1,535	7.95	95.23		
.	921	4.77	100.00		
age	19,297	51.31	14.20	14.00	105.00
Demographic					
hhsize	19,297	3.77	1.57	1.00	15.00
sexratio	19,296	0.48	0.21	0.00	1.00
elderratio	19,296	0.17	0.30	0.00	1.00
childrenratio	19,296	0.21	0.21	0.00	1.00
highestedu		Percent	Cum.		
None	1,130	5.86	5.86		
Primary School	3,165	16.40	22.26		
Secondary School	5,495	28.48	50.74		
High School	5,868	30.41	81.14		
A degree	3,423	17.74	98.88		
.	216	1.12	100.00		
Economics					
inchhy	19,291	122,270.80	131,988.60	1,830.00	5,286,860.00
inchh	19,291	10,189.23	10,999.05	152.50	440,571.70
inchh1	19,291	9,233.03	10,046.38	152.50	440,571.70
incpc	19,291	2,838.26	2,853.47	102.00	110,143.00
incpc1	19,291	2,571.80	2,602.97	91.93	110,143.00
exphh	19,297	9,829.20	9,366.13	298.62	191,673.50
exphh1	19,297	8,905.32	8,391.30	269.12	167,211.80
exppc	19,297	2,818.19	2,734.67	152.00	62,647.86
exppc1	19,297	2,552.27	2,444.89	132.60	54,652.66
Housing condition					
sqm	19,278	80.97	53.70	5.00	750.00
pcsqm	19,278	24.91	20.30	2.00	360.00
flat		Percent	Cum.		
Apartment	698	3.61	3.61		
Flat	18,581	96.29	99.91		
.	18	0.09	100.00		
pillar		Percent	Cum.		
Reinforcement concrete	7,158	37.09	37.09		

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Bricks/stones	8,920	46.22	83.31		
Iron/steel/good wood	1,792	9.29	92.60		
Poor-quality wood/bamboo	1,380	7.15	99.75		
Others	30	0.16	99.91		
.	18	0.09	100.00		
roof		Percent	Cum.		
Reinforcement concrete	4,392	22.76	22.76		
Tiles (cement, terracotta)	5,507	28.54	51.30		
Roof slabs (cement, metal)	8,982	46.55	97.85		
Leave/straw/rolled roofing/ Others	397	2.06	99.91		
.	18	0.09	100.00		
wall		Percent	Cum.		
Concrete/ Bricks/stones	16,179	83.84	83.84		
Wood/Metal	1,855	9.62	93.46		
Calcareous soil/straw/ Bamboo	1,245	6.45	99.91		
.	18	0.09	100.00		
dwelling		Percent	Cum.		
Villa/ Houses with a private kitchen	4,341	22.50	22.50		
Houses – shared kitchen or bathroom	2,299	11.91	34.41		
Semi-permanent houses	11,285	58.48	92.89		
Temporary and other types of house	1,355	7.02	99.91		
.	18	0.09	100.00		
ownership		Percent	Cum.		
No	1,031	5.34	5.34		
Yes	18,248	94.56	99.91		
.	18	0.09	100.00		
Appliances					
aircon	19,229	0.15	0.36	0.00	1.00
fan	19,229	0.91	0.29	0.00	1.00
fridge	19,229	0.63	0.48	0.00	1.00
washing	19,229	0.31	0.46	0.00	1.00
oven	19,229	0.08	0.27	0.00	1.00
computer	19,229	0.22	0.41	0.00	1.00
tv	19,229	0.94	0.24	0.00	1.00
waterheater	19,229	0.24	0.43	0.00	1.00
generator	19,229	0.02	0.13	0.00	1.00
pump	19,229	0.49	0.50	0.00	1.00
printer	19,229	0.02	0.13	0.00	1.00
fax	19,229	0.00	0.04	0.00	1.00
mobile	19,229	0.88	0.32	0.00	1.00
video	19,229	0.48	0.50	0.00	1.00
hifi	19,229	0.14	0.35	0.00	1.00
radio	19,229	0.02	0.16	0.00	1.00
diskplayer	19,229	0.01	0.11	0.00	1.00
camera	19,229	0.04	0.20	0.00	1.00
cooker1	19,229	0.78	0.41	0.00	1.00
cooker2	19,229	0.85	0.35	0.00	1.00

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
vacuum	19,229	0.05	0.22	0.00	1.00
juicer	19,229	0.21	0.40	0.00	1.00
Energy					
coal	19,294	0.38	7.02	0.00	350.00
coal1	19,294	0.35	6.48	0.00	305.33
briquette	19,294	3.81	31.14	0.00	2,400.00
briquette1	19,294	3.58	30.08	0.00	2,400.00
petro	19,294	324.99	527.70	0.00	18,720.00
petro1	19,294	295.76	476.46	0.00	16,330.93
kerosene	19,294	1.16	13.43	0.00	1,200.00
kerosene1	19,294	1.08	12.70	0.00	1,200.00
mazut	19,294	0.44	7.30	0.00	300.00
mazut1	19,294	0.40	6.66	0.00	270.37
diesel	19,294	2.37	58.79	0.00	4,000.00
diesel1	19,294	2.20	54.92	0.00	3,800.00
lpg	19,294	107.63	114.26	0.00	4,500.00
lpg1	19,294	98.35	105.18	0.00	4,055.52
gas	19,294	0.22	7.40	0.00	780.00
gas1	19,294	0.20	6.59	0.00	680.46
firewood	19,294	41.45	110.82	0.00	4,320.00
firewood1	19,294	38.00	102.03	0.00	4,320.00
agrwaste	19,294	5.64	29.12	0.00	800.00
agrwaste1	19,294	5.19	26.80	0.00	720.98
enoth	19,294	1.88	20.56	0.00	800.00
enoth1	19,294	1.70	18.47	0.00	697.90
Electricity					
eleclm	19,297	241.61	254.67	2.00	7,320.00
eleclm1	19,297	218.02	226.41	1.74	6,385.81
kwh	19,297	146.73	145.01	2.00	5,862.00
kwh1	19,297	134.20	111.71	1.23	2,661.54
lnkwh1	19,297	4.59	0.85	0.20	7.89
elecly	19,297	2,809.84	2,798.05	16.00	42,000.00
elecly1	19,297	2,536.11	2,495.09	16.00	36,639.91
mp	19,297	1.86	0.38	1.37	2.85
mp1	19,297	1.68	0.32	1.37	2.48
rsp	19,297	43.17	61.70	0.00	253.93
rsp1	19,297	38.89	55.18	0.00	221.53
ap	19,297	1.66	0.21	1.37	2.75
ap1	19,297	1.51	0.15	1.37	2.40
w	19,291	0.03	0.02	0.00	0.60
Temperature					
cdd20	19,280	197.41	98.30	0.56	331.67
cdd21	19,280	171.75	90.49	0.00	300.67
cdd22	19,280	146.67	81.99	0.00	269.67
cdd23	19,280	122.27	72.86	0.00	238.67
cdd24	19,280	98.54	63.33	0.00	207.67
cdd25	19,280	75.67	53.45	0.00	176.67
cdd26	19,280	54.23	43.34	0.00	145.67
cdd27	19,280	35.26	33.04	0.00	115.00

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
cdd28	19,280	20.02	23.43	0.00	85.78
hdd10	19,280	0.06	0.72	0.00	15.56
hdd11	19,280	0.18	1.35	0.00	23.56
hdd12	19,280	0.44	2.42	0.00	35.00
hdd13	19,280	0.89	4.03	0.00	50.00
hdd14	19,280	1.64	6.28	0.00	66.44
hdd15	19,280	2.86	9.30	0.00	84.44
hdd16	19,280	4.55	13.27	0.00	103.89
hdd17	19,280	6.68	18.19	0.00	125.78
hdd18	19,280	9.20	23.88	0.00	151.33
Price schedule		Percent	Cum.		
2011m3 - 2011m12	6	0.03	0.03		
2012m1 - 2012m6	2,440	12.64	12.67		
2012m7 - 2012m12	2,468	12.79	25.46		
2013m8 - 2014m5	4,550	23.58	49.04		
2014m6 - 2015m3	497	2.57	51.62		
2015m4 - 2017m11	9,336	48.38	100		

Note. The summary statistics are calculated with household weight.
Table B-2. Summary statistics of the pooled data set
Source. Author compiled.

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Admin					
year		Percent	Cum.		
2012	1,746	31.59	31.59		
2014	1,851	33.50	65.09		
2016	1,929	34.91	100.00		
Month (the month before survey month)		Percent	Cum.		
2012m2	2	0.03	0.03		
2012m3	6	0.10	0.13		
2012m5	9	0.16	0.29		
2012m7	2	0.04	0.33		
2012m8	904	16.35	16.68		
2012m9	18	0.33	17.01		
2012m10	16	0.29	17.30		
2012m11	789	14.29	31.59		
2014m11	1,851	33.50	65.09		
2016m8	1,929	34.91	100.00		
urban		Percent	Cum.		
Rural	3,733	67.56	67.56		
Urban	1,793	32.44	100.00		
reg6		Percent	Cum.		
Red River Delta	1,441	26.08	26.08		
Midlands and Northern Mountains	678	12.27	38.35		
Northern and Coastal Central	1,239	22.42	60.77		
Central Highlands	321	5.81	66.58		
Southeast	802	14.52	81.10		
Mekong Delta	1,044	18.90	100.00		
reg8		Percent	Cum.		
Red River Delta	1,365	24.70	24.70		
East Northern Mtns	641	11.60	36.30		
West Northern Mtns	113	2.05	38.35		
North Central Coast	679	12.28	50.63		
South Central Coast	468	8.47	59.10		
Central Highlands	321	5.81	64.91		
Southeast	895	16.19	81.10		
Mekong Delta	1,044	18.90	100.00		
Head characteristics					
Male		Percent	Cum.		
Female	1,433	25.94	25.94		
Male	4,093	74.06	100.00		
Edu		Percent	Cum.		
None	1,047	18.94	18.94		
Primary School	1,407	25.47	44.40		
Secondary School	1,662	30.08	74.49		
High School	800	14.48	88.97		
A degree	377	6.83	95.80		
.	232	4.2	100		
age	5,526	52.71	13.85	16.00	99.00

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Demographics					
hhsize	5,526	3.77	1.55	1.00	13.00
sexratio	5,526	0.47	0.21	0.00	1.00
elderratio	5,526	0.19	0.31	0.00	1.00
childrenratio	5,526	0.20	0.21	0.00	0.75
highestedu		Percent	Cum.		
None	315	5.70	5.70		
Primary School	1,000	18.09	23.79		
Secondary School	1,613	29.19	52.98		
High School	1,602	28.99	81.97		
A degree	939	16.99	98.96		
.	57	1.04	100.00		
Economics					
inchhy	5,525	119,837.10	120,294.50	2,765.00	2,646,980.00
inchh	5,525	9,986.42	10,024.54	230.42	220,581.70
inchh1	5,525	9,143.08	9,109.58	201.01	198,794.00
incpc	5,525	2,739.24	2,723.58	115.00	68,058.00
incpc1	5,525	2,506.84	2,492.19	100.32	64,972.00
exphh	5,526	9,618.07	10,323.82	320.34	191,673.50
exphh1	5,526	8,805.44	9,304.97	320.34	167,211.80
exppc	5,526	2,745.99	3,159.33	232.53	62,647.86
exppc1	5,526	2,510.81	2,843.06	232.53	54,652.66
Housing condition					
sqm	5,525	82.13	52.47	9.00	608.00
pcsqm	5,525	25.47	20.63	2.00	450.00
flat		Percent	Cum.		
Apartment	168	3.04	3.04		
Flat	5,357	96.94	99.99		
.	1	0.01	100.00		
pillar		Percent	Cum.		
Reinforcement concrete	2,057	37.23	37.23		
Bricks/stones	2,614	47.31	84.53		
Iron/steel/good wood	465	8.41	92.94		
Poor-quality wood/bamboo	375	6.79	99.73		
Others	14	0.25	99.99		
.	1	0.01	100.00		
roof		Percent	Cum.		
Reinforcement concrete	1,253	22.67	22.67		
Tiles (cement, terracotta)	1,767	31.98	54.65		
Roof slabs (cement, metal)	2,394	43.32	97.97		
Leave/straw/rolled roofing/ Others	112	2.02	99.99		
.	1	0.01	100.00		
wall		Percent	Cum.		
Concrete/ Bricks/stones	4,697	84.99	84.99		
Wood/Metal	483	8.74	93.73		
Calcareous soil/straw/ Bamboo	346	6.26	99.99		
.	1	0.01	100.00		

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
dwelling		Percent	Cum.		
Villa/ Houses with a private kitchen	1,248	22.58	22.58		
Houses – shared kitchen or bathroom	707	12.79	35.37		
Semi-permanent houses	3,228	58.42	93.79		
Temporary and other types of house	343	6.20	99.99		
.	1	0.01	100.00		
ownership		Percent	Cum.		
No	168	3.04	3.04		
Yes	5,358	96.95	99.99		
.	1	0.01	100.00		
Appliances					
aircon	5,513	0.13	0.34	0.00	1.00
fan	5,513	0.91	0.28	0.00	1.00
fridge	5,513	0.63	0.48	0.00	1.00
washing	5,513	0.31	0.46	0.00	1.00
oven	5,513	0.07	0.25	0.00	1.00
computer	5,513	0.20	0.40	0.00	1.00
tv	5,513	0.94	0.23	0.00	1.00
waterheater	5,513	0.25	0.43	0.00	1.00
generator	5,513	0.02	0.13	0.00	1.00
pump	5,513	0.50	0.50	0.00	1.00
printer	5,513	0.02	0.12	0.00	1.00
fax	5,513	0.00	0.03	0.00	1.00
mobile	5,513	0.87	0.33	0.00	1.00
video	5,513	0.49	0.50	0.00	1.00
hifi	5,513	0.15	0.36	0.00	1.00
radio	5,513	0.03	0.16	0.00	1.00
diskplayer	5,513	0.01	0.09	0.00	1.00
camera	5,513	0.04	0.19	0.00	1.00
cooker1	5,513	0.78	0.41	0.00	1.00
cooker2	5,513	0.86	0.35	0.00	1.00
vacuum	5,513	0.06	0.23	0.00	1.00
juicer	5,513	0.20	0.40	0.00	1.00
Energy					
coal	5,526	0.75	14.86	0.00	900.00
coal1	5,526	0.70	13.59	0.00	811.10
briquette	5,526	3.69	32.28	0.00	1,000.00
briquette1	5,526	3.46	30.53	0.00	1,000.00
petro	5,526	300.82	423.34	0.00	9,400.00
petro1	5,526	276.40	385.88	0.00	8,200.36
kerosene	5,526	1.45	18.02	0.00	600.00
kerosene1	5,526	1.35	16.49	0.00	580.00
mazut	5,526	0.74	10.46	0.00	250.00
mazut1	5,526	0.71	10.00	0.00	250.00
diesel	5,526	2.43	61.55	0.00	3,000.00
diesel1	5,526	2.27	57.94	0.00	3,000.00
lpg	5,526	109.45	113.90	0.00	3,500.00
lpg1	5,526	101.18	107.25	0.00	3,500.00
gas	5,526	0.47	21.17	0.00	1,560.00
gas1	5,526	0.43	19.21	0.00	1,405.91
firewood	5,526	38.67	103.27	0.00	4,500.00
firewood1	5,526	35.94	96.64	0.00	4,055.52

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
agrwaste	5,526	5.13	31.43	0.00	1,000.00
agrwaste1	5,526	4.83	30.44	0.00	1,000.00
enoth	5,526	1.73	15.68	0.00	400.00
enoth1	5,526	1.61	14.70	0.00	400.00
Electricity					
eleclm	5,526	240.76	250.53	5.00	3,500.00
eleclm1	5,526	219.69	224.92	4.36	3,053.33
kwh	5,526	148.70	132.01	2.00	1,666.00
kwh1	5,526	135.72	111.74	3.06	1,319.16
lnkwh1	5,526	4.60	0.84	1.12	7.18
elecly	5,526	2,840.83	2,877.98	50.00	42,000.00
elecly1	5,526	2,592.42	2,589.83	50.00	36,639.91
mp	5,526	1.82	0.36	1.37	2.85
mp1	5,526	1.68	0.31	1.37	2.48
rsp	5,526	41.77	61.22	0.00	253.93
rsp1	5,526	38.15	55.48	0.00	221.53
ap	5,526	1.64	0.20	1.37	2.65
ap1	5,526	1.50	0.15	1.37	2.31
w	5,525	0.03	0.02	0.00	0.60
Temperature					
cdd20	5,526	214.29	77.05	16.67	324.44
cdd21	5,526	185.33	73.92	9.67	293.44
cdd22	5,526	157.42	69.06	5.11	262.44
cdd23	5,526	130.38	62.87	1.67	231.44
cdd24	5,526	104.07	55.72	0.00	200.44
cdd25	5,526	78.83	47.42	0.00	169.44
cdd26	5,526	55.18	37.92	0.00	138.44
cdd27	5,526	34.01	28.18	0.00	107.44
cdd28	5,526	17.32	19.62	0.00	77.67
hdd10	5,526	0.00	0.00	0.00	0.00
hdd11	5,526	0.00	0.00	0.00	0.00
hdd12	5,526	0.00	0.00	0.00	0.00
hdd13	5,526	0.00	0.00	0.00	0.00
hdd14	5,526	0.00	0.01	0.00	0.11
hdd15	5,526	0.01	0.10	0.00	1.67
hdd16	5,526	0.03	0.33	0.00	5.56
hdd17	5,526	0.11	0.74	0.00	10.56
hdd18	5,526	0.29	1.61	0.00	16.78
Price schedule		Percent	Cum.		
2012m1 - 2012m6	16	0.29	0.29		
2012m7 - 2012m12	1,729	31.30	31.59		
2014m6 - 2015m3	1,851	33.50	65.09		
2015m4 - 2017m11	1,929	34.91	100.00		

Note. The summary statistics are calculated with household weight.
Table B-3. Summary statistics of the panel data set
Source. Author compiled.

Appendix C. Price elasticity

lnkwh1	*OLS*	*IV*
k		0.153
		[0.3219]
Income		
lnincpc1	0.1249***	0.2417***
	[0.0092]	[0.0144]
Prices		
lnrsp1	0.0129***	0.0037***
	[0.0003]	[0.0013]
lnmp1	3.8332***	-1.6497***
	[0.0842]	[0.2277]
lnapmp1	3.5368***	-0.2525
	[0.1043]	[0.5527]
Head characteristics		
male	0.0005	-0.013
	[0.0062]	[0.0120]
age	0.0001	0.0016***
	[0.0003]	[0.0004]
Demographics		
lnhhsize	0.1714***	0.5525***
	[0.0126]	[0.0335]
lnsexratio	0.0044***	0.0044***
	[0.0009]	[0.0012]
lnchildrenratio	0.0003	0.001
	[0.0003]	[0.0007]
lnelderratio	-0.0002	0
	[0.0003]	[0.0006]
Household highest degree (base is no education)		
Primary	0.0846***	0.0649***
	[0.0180]	[0.0172]
Secondary	0.1291***	0.1205***
	[0.0198]	[0.0210]
High school	0.1484***	0.1433***
	[0.0205]	[0.0213]
A degree	0.1215***	0.1214***
	[0.0190]	[0.0255]
Housing condition		
Lnpcsqm	0.0369***	0.1369***
	[0.0053]	[0.0094]
ownership	-0.0372***	-0.0733**
	[0.0114]	[0.0313]
Appliances		
aircon	-0.0383***	0.3215***
	[0.0075]	[0.0223]
fan	0.1143***	0.1018***
	[0.0182]	[0.0273]
fridge	0.2207***	0.5483***

	[0.0118]	[0.0269]
washing	-0.0082	0.1590***
	[0.0058]	[0.0168]
oven	-0.0555***	0.0840***
	[0.0076]	[0.0126]
computer	-0.0259***	0.1158***
	[0.0066]	[0.0139]
tv	0.2440***	0.2636***
	[0.0199]	[0.0237]
waterheater	0.0160*	0.1084***
	[0.0086]	[0.0231]
Related goods		
generator	-0.0398**	0.0802***
	[0.0157]	[0.0308]
lpg1	0.0001***	0.0007***
	[0.0000]	[0.0002]
Admin and Temperature		
cdd25	0.0003***	0.0012***
	[0.0001]	[0.0001]
urban	0.0493***	0.1593***
	[0.0072]	[0.0153]
year_2014	-0.0578***	0.0892***
	[0.0081]	[0.0146]
year_2016	-0.2114***	0.1788***
	[0.0135]	[0.0246]
_cons	1.4821***	1.5337***
	[0.1100]	[0.1012]
N	18971	18971
R-sq	0.873	0.602
adj. R-sq	0.872	0.6

Note. Model is estimated with weight and province dummy variables.
Table C-1. Full results of the perceived price model
Source. Author estimated.

Appendix D. Electricity poverty

lnkwh1pc	*National*	*Rural*	*Urban*
Income quantiles (selected first 20 quantiles out of 100 quantiles)			
_Iqincpc1_2	-0.0199	-0.1177	-0.0164
	[0.0886]	[0.1034]	[0.0923]
_Iqincpc1_3	0.1047	0.0218	0.0878
	[0.0748]	[0.0904]	[0.0784]
_Iqincpc1_4	0.1631**	0.0077	0.1600**
	[0.0780]	[0.0733]	[0.0798]
_Iqincpc1_5	0.1681**	0.1457*	0.1125
	[0.0702]	[0.0811]	[0.0879]
_Iqincpc1_6	0.1315*	0.108	0.2153**
	[0.0731]	[0.0820]	[0.0975]
_Iqincpc1_7	0.2007***	0.1121	0.1206
	[0.0731]	[0.0816]	[0.0846]
_Iqincpc1_8	0.1615**	0.0746	0.1605*
	[0.0758]	[0.0709]	[0.0853]
_Iqincpc1_9	0.1903**	0.1317*	0.2841***
	[0.0774]	[0.0776]	[0.0867]
_Iqincpc1_10	0.2030**	0.1214	0.3180***
	[0.0804]	[0.0822]	[0.0770]
_Iqincpc1_11	0.3290***	0.0738	0.3394***
	[0.0722]	[0.0673]	[0.1019]
_Iqincpc1_12	0.3006***	0.1914**	0.3306***
	[0.0894]	[0.0846]	[0.0721]
_Iqincpc1_13	0.3769***	0.1294	0.2555***
	[0.0851]	[0.0881]	[0.0788]
_Iqincpc1_14	0.2756***	0.1233	0.3516***
	[0.0734]	[0.0774]	[0.0901]
_Iqincpc1_15	0.3608***	0.2610***	0.4142***
	[0.0678]	[0.0757]	[0.0840]
_Iqincpc1_16	0.3619***	0.2787***	0.2799***
	[0.0645]	[0.0976]	[0.0918]
_Iqincpc1_17	0.3257***	0.3129***	0.3825***
	[0.0677]	[0.0845]	[0.0810]
_Iqincpc1_18	0.4407***	0.2406***	0.3046***
	[0.0896]	[0.0793]	[0.0938]
_Iqincpc1_19	0.3827***	0.2585***	0.2848***
	[0.0725]	[0.0700]	[0.0768]
_Iqincpc1_20	0.4844***	0.2924***	0.3307***
	[0.0759]	[0.0838]	[0.0796]
Prices			
lnmp1	-1.1447***	-1.5356***	-0.1579
	[0.1966]	[0.2485]	[0.2854]
lnrsp1	0.0028***	0.0026**	0.0034**
	[0.0009]	[0.0013]	[0.0016]
Head characteristics			
male	-0.01	0.0019	-0.0158
	[0.0114]	[0.0153]	[0.0133]
age	0.0013***	0.0012**	0.0007
	[0.0004]	[0.0005]	[0.0005]
Demographics			

lnkwh1pc	*National*	*Rural*	*Urban*
lnhhsize	-0.4733***	-0.5135***	-0.4752***
	[0.0297]	[0.0341]	[0.0593]
lnsexratio	0.0046***	0.0050***	0.0035***
	[0.0011]	[0.0016]	[0.0012]
lnchildrenratio	0.0006	0.0015**	-0.0002
	[0.0006]	[0.0007]	[0.0009]
lnelderratio	-0.0001	-0.0002	0.0003
	[0.0006]	[0.0007]	[0.0008]
Household highest degree (base is no education)			
Primary	0.0692***	0.0917***	0.0195
	[0.0166]	[0.0202]	[0.0310]
Secondary	0.1213***	0.1540***	0.0584
	[0.0210]	[0.0213]	[0.0429]
High school	0.1401***	0.1699***	0.0779*
	[0.0211]	[0.0238]	[0.0422]
A degree	0.1262***	0.1547***	0.0607
	[0.0246]	[0.0262]	[0.0446]
Housing condition			
lnpcsqm	0.1312***	0.1412***	0.1066***
	[0.0092]	[0.0135]	[0.0134]
ownership	-0.0773***	-0.0382	-0.0646**
	[0.0280]	[0.0297]	[0.0261]
Appliances			
aircon	0.3051***	0.3624***	0.2267***
	[0.0231]	[0.0384]	[0.0289]
fan	0.0941***	0.1242***	-0.0384
	[0.0261]	[0.0229]	[0.0561]
fridge	0.5308***	0.5755***	0.3703***
	[0.0280]	[0.0353]	[0.0395]
washing	0.1507***	0.1605***	0.1166***
	[0.0149]	[0.0207]	[0.0210]
oven	0.0878***	0.0236	0.0875***
	[0.0123]	[0.0289]	[0.0187]
computer	0.1081***	0.1076***	0.0835***
	[0.0129]	[0.0200]	[0.0124]
tv	0.2528***	0.2479***	0.2412***
	[0.0226]	[0.0219]	[0.0566]
waterheater	0.1057***	0.1506***	0.0463***
	[0.0213]	[0.0240]	[0.0157]
Related goods			
generator	0.0771**	0.1393***	-0.0029
	[0.0307]	[0.0478]	[0.0329]
lpg1	0.0006***	0.0009***	0.0003*
	[0.0001]	[0.0001]	[0.0001]
Admin and Temperature			
cdd25	0.0011***	0.0011***	0.0011***
	[0.0001]	[0.0001]	[0.0002]
year_2014	0.0782***	0.0772***	0.0602***
	[0.0137]	[0.0180]	[0.0188]
year_2016	0.1661***	0.2088***	0.0828***

lnkwh1pc	*National*	*Rural*	*Urban*
	[0.0179]	[0.0280]	[0.0166]
urban		0.1494***	
		[0.0143]	
_cons	2.6511***	2.7660***	3.0113***
	[0.1119]	[0.1511]	[0.1765]
N	18971	13171	5800
R-sq	0.633	0.557	0.665
adj. R-sq	0.629	0.55	0.654

Note. All income quantiles above the 20th quantiles are positive and significant at the level of 0.05

Table D-1. Full results of electricity poverty models with full samples

Source. Author estimated.

lnkwh1pc	*National*	*Rural*	*Urban*
Income quantiles (selected first 10 quantiles out of 100 quantiles)			
_Iqincpc1_2	0.0737	0.033	-0.0503
	[0.0835]	[0.0938]	[0.0986]
_Iqincpc1_3	0.1589**	0.1259	0.0664
	[0.0679]	[0.1039]	[0.0840]
_Iqincpc1_4	0.2087***	0.0941	0.1479*
	[0.0806]	[0.0846]	[0.0758]
_Iqincpc1_5	0.2195***	0.2552***	0.128
	[0.0638]	[0.0861]	[0.0853]
_Iqincpc1_6	0.1639**	0.1927**	0.1978*
	[0.0726]	[0.0883]	[0.1042]
_Iqincpc1_7	0.2517***	0.1931**	0.0987
	[0.0739]	[0.0781]	[0.0895]
_Iqincpc1_8	0.2188***	0.1498**	0.1679*
	[0.0701]	[0.0757]	[0.0902]
_Iqincpc1_9	0.2478***	0.2244**	0.2724***
	[0.0768]	[0.0881]	[0.0945]
_Iqincpc1_10	0.2562***	0.2271***	0.3517***
	[0.0787]	[0.0872]	[0.0863]
Prices			
lnmp1	-1.1723***	-1.5815***	-0.2229
	[0.2039]	[0.2638]	[0.3170]
lnrsp1	0.0025***	0.0023*	0.0025
	[0.0009]	[0.0012]	[0.0018]
Head characteristics			
male	-0.0098	0.0028	-0.0153
	[0.0115]	[0.0153]	[0.0139]
age	0.0013***	0.0013**	0.0007
	[0.0004]	[0.0005]	[0.0005]
Demographics			
lnhhsize	-0.4689***	-0.5076***	-0.4597***
	[0.0302]	[0.0329]	[0.0706]
lnsexratio	0.0046***	0.0050***	0.0035***
	[0.0011]	[0.0015]	[0.0011]

lnkwh1pc	*National*	*Rural*	*Urban*
lnchildrenratio	0.0006	0.0015**	-0.0003
	[0.0006]	[0.0007]	[0.0009]
lnelderratio	-0.0001	-0.0002	0.0002
	[0.0006]	[0.0008]	[0.0008]
Household highest degree (base is no education)			
Primary	0.0684***	0.0932***	0.0208
	[0.0168]	[0.0207]	[0.0312]
Secondary	0.1211***	0.1568***	0.0604
	[0.0214]	[0.0217]	[0.0434]
High school	0.1400***	0.1727***	0.0812*
	[0.0215]	[0.0244]	[0.0424]
A degree	0.1257***	0.1581***	0.0642
	[0.0251]	[0.0268]	[0.0448]
Housing condition			
lnpcsqm	0.1323***	0.1428***	0.1094***
	[0.0094]	[0.0132]	[0.0145]
ownership	-0.0777***	-0.038	-0.0689**
	[0.0283]	[0.0302]	[0.0295]
Appliances			
aircon	0.3070***	0.3660***	0.2316***
	[0.0239]	[0.0401]	[0.0320]
fan	0.0943***	0.1240***	-0.0389
	[0.0262]	[0.0230]	[0.0582]
fridge	0.5364***	0.5829***	0.3871***
	[0.0302]	[0.0358]	[0.0394]
washing	0.1526***	0.1627***	0.1223***
	[0.0153]	[0.0211]	[0.0251]
oven	0.0881***	0.0236	0.0883***
	[0.0124]	[0.0292]	[0.0190]
computer	0.1097***	0.1093***	0.0865***
	[0.0132]	[0.0202]	[0.0134]
tv	0.2526***	0.2472***	0.2460***
	[0.0230]	[0.0222]	[0.0575]
waterheater	0.1065***	0.1531***	0.0479***
	[0.0214]	[0.0243]	[0.0152]
Related goods			
generator	0.0784**	0.1421***	-0.0027
	[0.0312]	[0.0490]	[0.0336]
lpg1	0.0006***	0.0009***	0.0003*
	[0.0001]	[0.0001]	[0.0002]
Admin and Temperature			
cdd25	0.0011***	0.0011***	0.0011***
	[0.0001]	[0.0001]	[0.0002]
urban	0.1509***		
	[0.0143]		
year_2014	0.0804***	0.0790***	0.0648***
	[0.0139]	[0.0177]	[0.0211]
year_2016	0.1711***	0.2153***	0.0940***
	[0.0182]	[0.0257]	[0.0170]

lnkwh1pc	*National*	*Rural*	*Urban*
_cons	2.5941***	2.6602***	2.9843***
	[0.1096]	[0.1525]	[0.1692]
N	18965	13165	5798
R-sq	0.627	0.547	0.645
adj. R-sq	0.623	0.541	0.634

Note. All income quantiles above the 10th quantiles are positive and significant at the level of 0.05
Table D-2. Full results of electricity poverty models after removing outliers
Source. Author estimated.

	National model		**Urban model**		**Rural model**	
Model without weight						
Endogeneity test	F stats	p-value	F stats	p-value	F stats	p-value
	593.563	0.0000	146.373	0.0000	364.783	0.0000
Weak instrument test *		Cri. val		Cri. val		Cri. val
Cragg & Donal min eigenvalue	344.885	7.03	124.642	7.03	225.441	7.03
Model with weight						
Endogeneity test	F stats	p-value	F stats	p-value	F stats	p-value
	557.491	0.0000	112.63	0.0000	425.948	0.0000
Weak instrument test **	F-stat	Cri. val	F-stat	Cri. val	F-stat	Cri. val
lnmp1	254.873	10	39.7951	10	152.241	10
lnrsp1	304.939	10	107.392	10	347.16	10

Note. Endogeneity test is Durbin-Wu-Hausman test; * Stock & Yogo test,
** F test for join significance of instruments in first stage models
Table D-3. Endogeneity tests and weak instrument tets
Source. Author estimated.

Appendix E. Economies of scale

Test of the functional form of the conditional mean for the First difference model

Ramsey RESET test using powers of the fitted values of D.lneleclm1

Ho: model has no omitted variables

F(3, 18580) = 0.82

Prob > F = 0.4816

lneleclm2	*Fixed effects*		*xtabond2*	
Lag of dependent variables				
L.lneleclm2	-0.2629***	[0.0229]	0.1685***	[0.0608]
Income				
lnincpc2	0.0978***	[0.0281]	0.3329***	[0.0268]
Head characteristics				
male	-0.0642	[0.0816]	-0.0629**	[0.0291]
age	-0.0048	[0.0030]	-0.0006	[0.0012]
Demographics				
lnhhsize	0.5169***	[0.0655]	0.5729***	[0.0458]
lnsexratio	0.0066	[0.0043]	0.0023	[0.0033]
lnchildrenratio	-0.0007	[0.0021]	0.0018	[0.0015]
lnelderratio	-0.0027	[0.0022]	0.0003	[0.0014]
Household highest degree (base is no education)				
Primary	0.0697	[0.0674]	0.1734***	[0.0567]
Secondary	0.0462	[0.0745]	0.2651***	[0.0563]
High school	0.0153	[0.0799]	0.3465***	[0.0596]
A degree	0.0881	[0.0897]	0.4332***	[0.0670]
Housing condition				
lnpcsqm	0.1906***	[0.0370]	0.1933***	[0.0244]
ownership	0.1136	[0.0836]	0.0347	[0.0614]
Related goods				
lpg2	0.0004***	[0.0001]	0.0010***	[0.0002]
Admin and Temperature				
cdd25	0.0009***	[0.0002]	0.0010***	[0.0002]
urban	-0.0496	[0.2721]	0.2368***	[0.0393]
svmonth_4	0	[.]	0	[.]
svmonth_6	0	[.]	0	[.]
svmonth_8	0	[.]	0	[.]
svmonth_9	0.2262***	[0.0214]	0.0823	[0.1898]
svmonth_10	0	[.]	0	[.]
svmonth_11	0	[.]	0	[.]
svmonth_12	0	[.]	-0.0296	[0.1829]
_cons	4.3778***	[0.3697]	0	[.]

lneleclm2	*Fixed effects*	*xtabond2*
N	3646	3646
df_m	17	24

Note. xtabond2 is Blundell and Bond (1998) estimators.

Table E-1. Full result of sensitivity analysis for economies of scale with panel data

Source. Author estimated.

lneleclm1	*OLS*		*First difference*	
Income				
lnincpc	0.2273***	[0.0117]		
Head characteristics				
male	-0.021	[0.0146]	0.0002	[0.0140]
age	0.0008	[0.0006]	0.0013**	[0.0006]
Demographics				
lnhhsize	0.4866***	[0.0221]	0.4879***	[0.0223]
lnsexratio	0.0060***	[0.0016]	0.0053***	[0.0015]
lnchildrenratio	0.0011	[0.0008]	0.0004	[0.0008]
lnelderratio	-0.0003	[0.0008]	-0.0005	[0.0008]
Household highest degree (base is no education)				
Primary	0.0807***	[0.0302]	0.1158***	[0.0288]
Secondary	0.1426***	[0.0291]	0.1575***	[0.0284]
High school	0.1463***	[0.0304]	0.1588***	[0.0295]
A degree	0.1513***	[0.0325]	0.1695***	[0.0321]
Housing condition				
lnpcsqm	0.1212***	[0.0122]	0.1319***	[0.0123]
ownership	-0.0115	[0.0406]	0.0357	[0.0359]
Appliances				
aircon	0.2393***	[0.0185]	0.2445***	[0.0182]
fan	0.1256***	[0.0242]	0.1063***	[0.0238]
fridge	0.5395***	[0.0167]	0.5293***	[0.0163]
washing	0.1177***	[0.0150]	0.1207***	[0.0154]
oven	0.0363	[0.0224]	0.0636***	[0.0237]
computer	0.0836***	[0.0163]	0.1006***	[0.0162]
tv	0.2822***	[0.0293]	0.2845***	[0.0275]
waterheater	0.0778***	[0.0164]	0.0747***	[0.0164]
Related goods				
generator	0.1117***	[0.0417]	0.1286***	[0.0395]
lpg	0.0006***	[0.0001]	0.0005***	[0.0001]
Admin and Temperature				
cdd25	0.0008***	[0.0001]	0.0008***	[0.0001]
urban	0.1189***	[0.0141]	0.1283***	[0.0140]
_cons	1.6193***	[0.1113]	-0.0035	[0.0074]

lneleclm1	*OLS*	*First difference*
N	9105	8974
R-sq	0.716	0.559
adj. R-sq	0.713	0.555
df_m	87	86
F	195.6865	107.5484

Table E-2. Full results of sensitivity analysis for economies of scale with 2016 data only
Source. Author estimated.

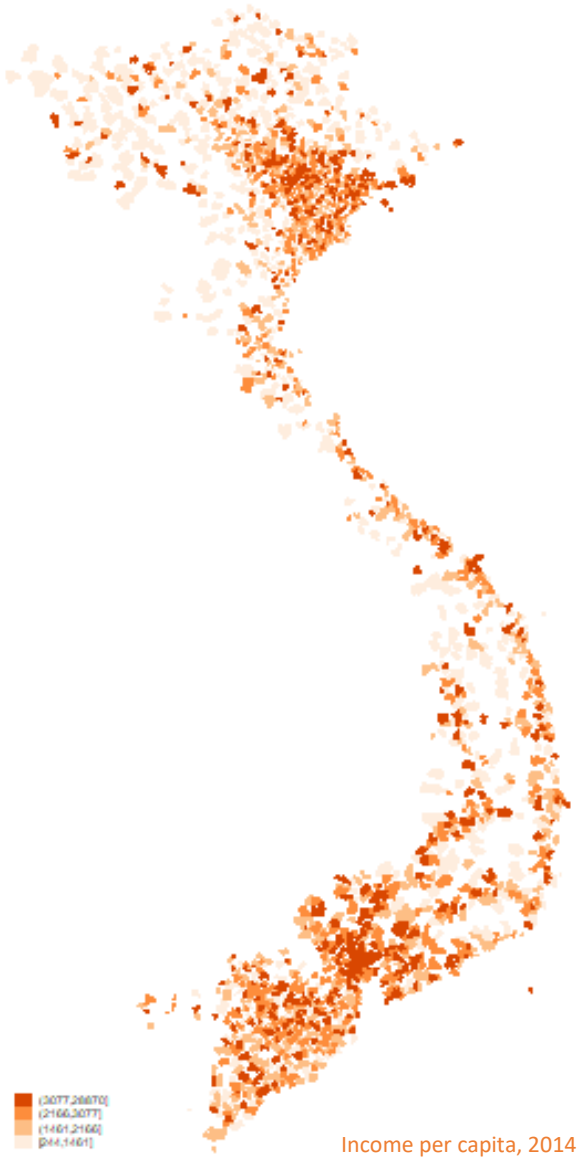
Appendix F. Heatwaves

lnkwh1	*National*	*Urban*	*Rural*
Income			
lnincpc1	0.2638***	0.1904***	0.2486***
	[0.0167]	[0.0213]	[0.0192]
Prices			
lnmp1	-1.4327***	-0.0048	-1.8237***
	[0.2919]	[0.2306]	[0.4025]
lnrsp1	0.0005	0.0011	0.0027
	[0.0015]	[0.0026]	[0.0020]
Head characteristics			
male	-0.0241	-0.0403**	-0.002
	[0.0183]	[0.0168]	[0.0211]
age	0.0012**	0.0001	0.0011
	[0.0006]	[0.0010]	[0.0009]
Demographics			
lnhhsize	0.5500***	0.4897***	0.4992***
	[0.0353]	[0.0502]	[0.0299]
lnsexratio	0.0060***	0.0052**	0.0055**
	[0.0016]	[0.0024]	[0.0024]
lnchildrenratio	0.0011**	0.0001	0.0015*
	[0.0005]	[0.0010]	[0.0008]
lnelderratio	0.0005	0.0012	0.0003
	[0.0008]	[0.0017]	[0.0010]
Household highest degree (base is no education)			
Primary	0.0588*	-0.0141	0.0878***
	[0.0310]	[0.0928]	[0.0248]
Secondary	0.1380***	0.0233	0.1821***
	[0.0301]	[0.0849]	[0.0246]
High school	0.1707***	0.0666	0.2056***
	[0.0292]	[0.0877]	[0.0272]
A degree	0.1612***	0.0621	0.1884***
	[0.0341]	[0.0887]	[0.0357]
Housing condition			
lnpcsqm	0.1470***	0.1147***	0.1411***
	[0.0182]	[0.0205]	[0.0204]
ownership	-0.1574***	-0.1584***	-0.0693
	[0.0507]	[0.0469]	[0.0646]
Appliances			
aircon	0.2992***	0.1860***	0.3750***
	[0.0382]	[0.0345]	[0.0486]
fan	0.0992***	-0.0562	0.1277***
	[0.0284]	[0.0630]	[0.0219]
fridge	0.5943***	0.4194***	0.5884***
	[0.0452]	[0.0724]	[0.0422]
washing	0.1872***	0.1321***	0.1929***
	[0.0271]	[0.0344]	[0.0363]
oven	0.0505**	0.0752***	-0.0165
	[0.0220]	[0.0255]	[0.0474]

lnkwh1	*National*	*Urban*	*Rural*
computer	0.0914***	0.0645***	0.0959***
	[0.0260]	[0.0212]	[0.0371]
tv	0.2663***	0.2277***	0.2542***
	[0.0247]	[0.0443]	[0.0300]
waterheater	0.1281***	0.0563*	0.1591***
	[0.0266]	[0.0288]	[0.0297]
Related goods			
generator	0.1028*	0.0501	0.1157
	[0.0564]	[0.0558]	[0.0840]
lpg1	0.0007***	0.0003	0.0011***
	[0.0002]	[0.0002]	[0.0002]
Admin and Temperature			
cdd25	0.0015***	0.0016***	0.0011***
	[0.0001]	[0.0002]	[0.0002]
heatwave	0.0162	-0.0362*	0.0431**
	[0.0143]	[0.0209]	[0.0185]
urban	0.1816***		
	[0.0212]		
year_2014	0.1081***	0.0962***	0.0929***
	[0.0170]	[0.0229]	[0.0189]
year_2016	0.2662***	0.1433***	0.2730***
	[0.0325]	[0.0360]	[0.0382]
_cons	1.1830***	1.9592***	1.4114***
	[0.1459]	[0.2082]	[0.1989]
N	9535	2712	6823
R-sq	0.592	0.711	0.536
adj. R-sq	0.588	0.701	0.53

Table F-1. Full results of the heatwave model
Source. Author estimates.

Annex 1. Assigning Official National Administration Unit Code to Vietnam GADM Shapefile 2018 at Ward Level



Assigning Official National Administration Unit Code to Vietnam GADM Shapefile 2018 at Ward Level

Technical report

July - 2018

ABSTRACT

This report assigns official national administration unit codes to Vietnam GADM Shapefile 2018 at ward level. The output is a new shapefile with official administration unit codes. These codes allow to join geographical data in shapefile with social-economic data to perform spatial econometric analysis or graph the map of social economic data at ward level. The assigning process finishes with 11,154 out of 11,163 wards (99.91%) assigned official admin codes.

Hoai-Son NGUYEN, Minh HA-DUONG

TECHNICAL REPORT

Assigning Official National Administration Unit Code to Vietnam GADM Shapefile 2018 at Ward Level

This report was prepared by:

Hoai-Son NGUYEN [1, 2]
Minh HA-DUONG [1, 3]

[1] Clean Energy and Sustainable Development Lab (CleanED), 18 Hoang Quoc Viet, Cau Giay, Ha Noi, Vietnam
[2] National Economics University (NEU), Vietnam
[3] International Research Center on Environment and Development (CIRED), National Center for Scientific Research (CNRS), France
Email: hoaisonkt@gmail.com; minh.haduong@gmail.com

Cover Illustration. Average income per capita at ward level of Vietnam 2014
Source: Author compiled from GADM shapefile and VHLSS 2014

With financial support from **Wellcome Trust Seed Awards**
Grant number 205764/Z/16/Z
"Assessing energy precarity and heat related health risks
from climate change in subtropical Asian cities"
coordinated by Dr. Leslie Mabon, Robert Gordon University

Cite as: .....

July - 2018

Contents

List of figures 4
List of tables 5
Summary 6
Objective 6
Results 6
Output format 6
1 Introduction 7
2 Administration hierarchy, shape files and official nation administration list 8
2.1 Administration hierarchy in Vietnam 8
2.2 GADM shapefile 8
2.3 GSO list 9
3 Methods 11
3.1 Objectives 11
3.2 Mismatches classification and resolution 11
4 Results 12
4.1 Matching process 12
4.1.1 Normalize and check for candidate key 12
4.1.2 Deal with differences in writing style convention 12
4.1.3 Deal with administration changes 13
4.2 Matching results 14
5 Example of using shapefile with social economic data 15
6 Conclusion 17

List of figures

Figure 1. Report objectives 7
Figure 2. Administration hierarchy in Vietnam 8
Figure 3. Objectives of the report in detail 11
Figure 4. Method to assign cdd25 to each household 16

List of tables

Table 1. Shapefile and the corresponding Stata datasets 8
Table 2. The structure of the shapefile 9
Table 3. Example of records in shapefile 9
Table 4. Example of records in unnormalized form in shapefile – admin unit type in prarentheses 9
Table 5. Example of records in unnormalized form in shapefile – district type in prefix 9
Table 6. Structure of GSO list 2014 10
Table 7. Example of GSO list 2014 data 10
Table 8. Comparison of admin unit number between shapefile and GSO list 10
Table 9. Example of differences in writing style – Leading zeros in name 12
Table 10. Example of differences in writing style – Capitalization 12
Table 11. Example of differences in writing style – Tone marks position 13
Table 12. Example of differences in writing style – Others 13
Table 13. Example of administration changes – District changes 13
Table 14. Example of administration changes – Ward changes 14
Table 15. Matching results 14
Table 16. List of unmatched cases 15
Table 17. Example of final shapefile 15
Table 18. Data description of VHLSS and GHCN 16
Table 19. Structure of the final example data 16
Table 20. Pearson'R correlation between temperature and average monthly income in Vietnam, Jun 2014 17

Summary

Objective

Shapefile is an important source for spatial econometrics and visualization in analysis. Spatial econometrics requires to connect social-economic data with the geographical information of analysis units such as central longitudes, latitudes or polygon borders. Shapefile is the popular form to store that geographical data. The shapefile format stores the data on geometric shapes like points, lines, and polygons. These shapes, together with the social-economic data linked to each shape, generate a dataset for spatial econometrics and visualization.

As we wrote this, Global Administrative Areas (GADM) shapefile of Vietnam version 3.4 (GADM 2018) seems to be the best choice for researchers. First, the shapefile is free. Second, the version 3.4 is updated to April 2018. It is the most updated shapefile to this moment and is suitable for using with recent updated data such as Vietnam Living Standard Survey (VHLSS) 2016 or Enterprise Survey 2016. Under the condition that the official shapefile provided by Vietnamese government is not updated regularly and hard to access, the GADM shapefile turns to be the better choice.

However, the GADM shapefile has only the administration unit names but does not have administration unit codes. It poses a challenge in joining with social-economic data since these data in Vietnam use administration unit codes instead of unit names. Thus, this report aims to assign national administration unit codes to the GADM shapefile. The new shapefile is then available to plug in any social-economic data in Vietnam to perform spatial economic analysis.

We employ the official administration list from General Statistics Offices (GSO) 2014 (GSO 2015) as a medium to assign the admin code to shapefile. Shapefile has the name of each admin unit and geographical data of that unit. Social-economic data has data on each admin unit and the national code of that admin unit. GSO list has both national code and name of each admin unit. By joining the GSO list with the shapefile via admin unit name, we have new shapefile including not only admin unit names and geographical data but also admin unit codes which can serve as a key to join with social economic data later.

Results

The assigning process is done by constructing a map table. The map table includes matched cases in admin unit names in shapefile and GSO list. The map table is filled in in three phases. The first is after the normalization of both files to ensure that each field contains only one information. The second is after dealing with differences in writing styles. The last one is after adjusting for administration changes from 2014 – the year of GSO list and 2018 – the year of shapefile.

In the map table, 11,154 out of 11,170 wards (99.86%) are matched between GADM shapefile and GSO list. Only 16 cases are not matched in which 9 cases are from shapefile and 7 cases are from GSO list. Thus, in the new shapefile, we assigned national code to 11,154 wards out of total 11,163 wards (99.91%) in the original GADM shapefile.

Output format

All the outputs are stored in a zip file named “vnshp.zip”. The zip file has three folders and a license file.

- The “VN shapefile” folder stores new shapefile in shapefile format including vnshp.dbf, vnshp.shp and vnshp.shx.
- The “Supplementary” folder stores (i) the map table in csv format, (ii) Stata do file (script) for matching process, (iii) Example folder storing data and script of example of using new shapefile.

The shapefile is constructed under financial support from Wellcome Trust Seed Awards. The license of the shapefile is according to the Attribution – NonCommercial 4.0 Generic (CC BY – NC 4.0) of Creative Commons (Creative Commons Accessed 2018-07-18). You are free to copy and redistribute the material in any medium or format as well as remix, transform, and build upon the material. The license is under the term that (i) You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. And (ii) You may not use the material for commercial purposes.

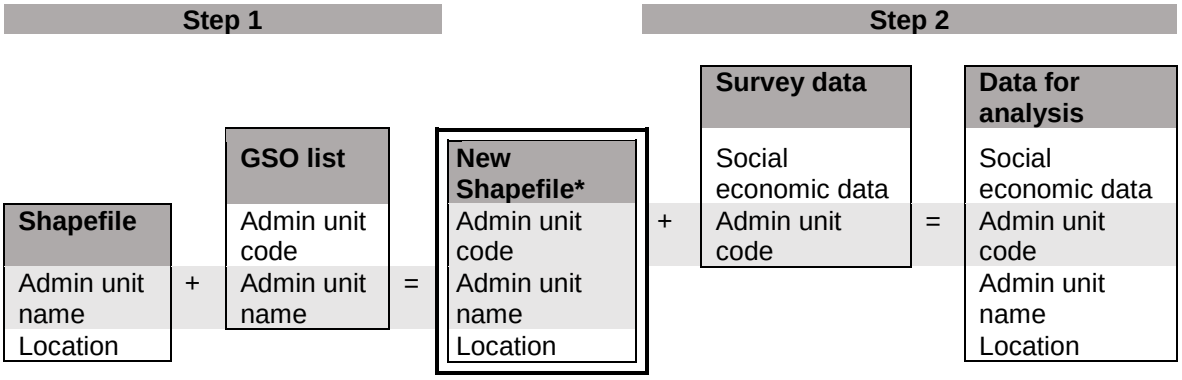
1 Introduction

Shapefile is a useful source for spatial econometrics and visualization in analysis. Spatial econometrics require to connect social-economic data with the location of analysis units such as central longitudes, latitudes or polygon borders. In a social economic dataset, each analysis unit has data on the name or official code of its related administration unit. Meanwhile, shapefiles consist of a list of administration units with their location. Merging social economic data with shapefile is done by matching administration unit name or code.

Currently, Vietnam has two sources of shapefile including official and free sources. The official shape file has full information of administration units including location, names and official national codes which match with the admin unit code in national surveys. However, this source is hard to access. In addition, the shapefile is not updated regularly. The most updated official shapefile we can access is from 2008. The shapefile did not catch up with the changes in administration units from 2008 to present. In addition, the national admin code in that shapefile followed older system which already change after 2008. Thus, the shape file is no longer suitable for analysis with social economic data after 2008.

By contrast, the free sources have more updated shapefiles. To our best knowledge, the best free source of shapefile so far is Global Administrative Areas (GADM). The most updated shapefile of Vietnam in GADM is April 2018 which is suitable for using with recent updated data such as Vietnam Living Standard Survey (VHLSS) 2016 or Enterprise Survey 2016. However, the free shapefile has the list of administration unit names but official codes. The lack of official admin code imposes a challenge in merging with social economic data. In this case, the shapefile has admin units' name and location data. Meanwhile, survey data normally has the admin units' official codes and social-economic data. Merging the two datasets needs to match admin unit names in shapefile with admin unit codes in survey data.

This report aims to improve the free GADM shapefile by assigning administration official code to the file. We perform the task by utilizing the official administration list issued by General Statistics Office (GSO). The official list contains both administration unit names and codes. Merging the list with shapefile by admin unit name results in a new shapefile having admin unit codes. The new shapefile is then able to merge with any survey data which has admin unit code only.



Note. * New shapefile is the output of the report
Figure 1. Report objectives
Source: Authors compiled.

This report focuses on shapefile at ward level. The GSO list at Dec 31, 2014 is employed for the report. All data processing is done with Stata 14. The report contains six parts. The first is the introduction. The second is a brief review on Vietnam administration hierarchy as well as GSO list and GADM shapefile. The third part describes the methods assigning admin unit code from GSO list to shapefile. The forth is the result following by a small example of how to use the new shapefile with survey data. The last part is a conclusion.

2 Administration hierarchy, shape files and official nation administration list

2.1 Administration hierarchy in Vietnam

The administration hierarchy in Vietnam includes 3 tiers (Vietnamese National Assembly 2013, 2015)

- 1st tier is city level including Province/Municipality
- 2nd tier is district level including urban/rural district, town, Provincial city, Municipality city
- 3rd tier is ward level including ward, commune, township

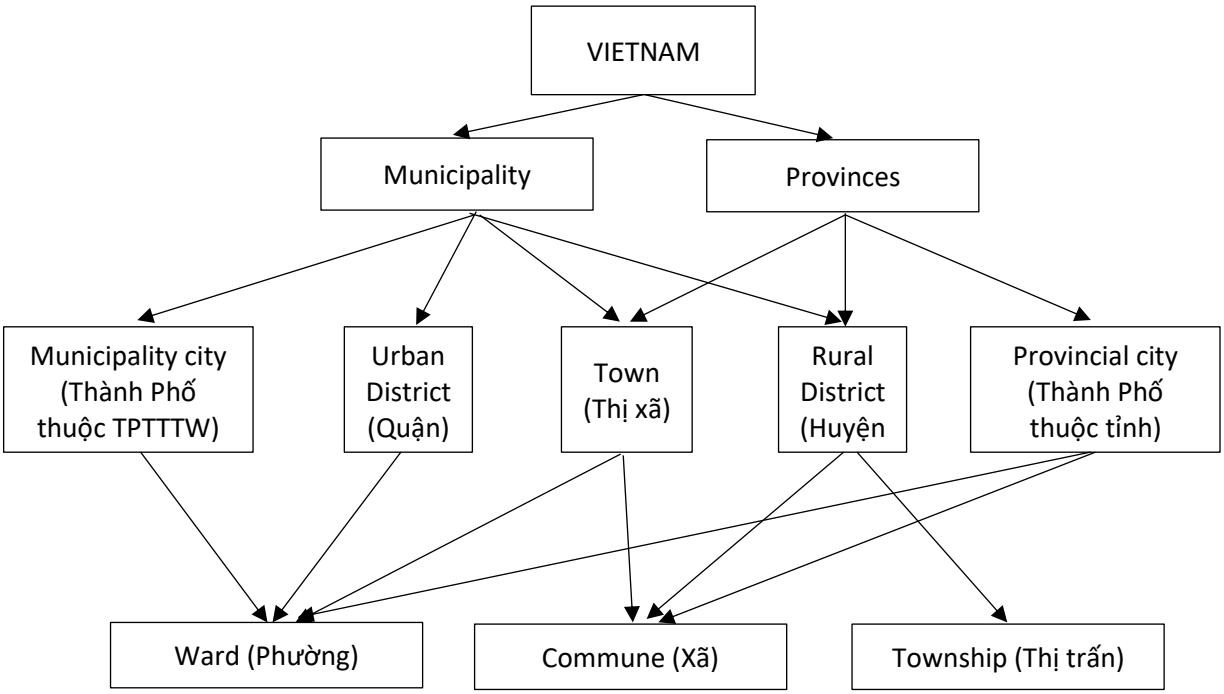


Figure 2. Administration hierarchy in Vietnam
Source: Author compiled

2.2 GADM shapefile

The most updated GADM shapefile is version 3.4, April 2018. The original GADM shape at ward level include five files with the same name “gadm36_VNM_3” with different suffixes including .shp, .dbf, .shx, .prj and .cpg (GADM 2018). We only need two files with the suffixes of .dbf an .shp. The .shp file contains the geometry data of each wards with a list of its vertices. The .dbf file contains wards’ attributes with one record per ward. The relationship between the twofile is one-to-one based on record number. Attribute records in the dbf file must be in the same order as records in the shp file (Environmental Systems Research Institute 1998).

For the convenience in processing data with Stata, the original shapefile is converted to Stata datasets by a user-written command shp2dta (Crow 2015). The converted process is detail in following Table 1.

	Shapefile	Stata dataset converted	Description
1	Gadm36_VNM_3.shp	vncoord_centroids.dta	Wards’ polygon data
2	Gadm36_VNM_3.dbf	vndb_centroids.dta	Wards’ attribute including central longitude and latitude

Table 1. Shapefile and the corresponding Stata datasets
Source: Authors compiled.

The vndb_centroids.dta file covers general attributes of each ward. The vncoord_centroids.dta has information of the polygon information of the ward in term of the ward vertices longitude and latitude. The two files are connected by a field named “id” in former which is correspond to the values taken on by variable _ID in the latter. Since we focus on assigning national code to each ward, from hereafter we operate everything in the vndb_centroids file. From hereafter, the shapefile means the vndb_centroids file.

Fields	Description
id	Area ID to connect with vncoord_centroids – the polygon file
x_center	x-coordinate of area centroid (central longitude)
y_center	y-coordinate of area centroid (central latitude)
country	Country name
city	City name
district	District name
ward	Ward name
wardtype	Ward type

Note. Fields in bold are primary key
Table 2. The structure of the shapefile
Source: Authors compiled

In the shapefile, each record describes attributes of a single ward. There are two ways to characterize the identity of each ward (each row). By definition, “id” field is the first one. The “id” is unique for each row. By nature, the combination of “city, district, ward and ward type” is the second one. The file does not have two distinct rows having the same values for these four attributes. There is also no proper subset of these four attributes for which the above condition holds.

We have two candidate keys for the file. However, the “id” field is for polygon file connection purpose only. It is not the ward national official code nor presented in GSO list. Thus, the combination of “city, district, ward, wardtype” is selected as the primary composite key for the file.

id	city	district	ward	wardtype	x_center	y_center
1	An Giang	An Phú	An Phú	Thị trấn	105.0868	10.79434
2	An Giang	An Phú	Đa Phước	Xã	105.1162	10.74601
3	An Giang	An Phú	Khánh An	Xã	105.108	10.94508

Table 3. Example of records in shapefile
Source: Authors compiled

It should be noted that the shapefile is not normalized in term that each field contain only one information. Some ward or district name includes the ward type, district type in parentheses. These cases are for two wards/district having same name in a same district/city but differing ward type/district type.

Id	City	District	Ward	Wardtype
208	Bạc Liêu	Phước Long	Phước Long (Thị trấn)	Thị trấn
209	Bạc Liêu	Phước Long	Phước Long (Xã)	Xã
2301	Đồng Tháp	Hồng Ngự	Thường Thới Tiền	Xã
2302	Đồng Tháp	Hồng Ngự (Thị xã)	An Bình A	Xã

Table 4. Example of records in unnormalized form in shapefile – admin unit type in prarentheses
Source: Author compiled.

Besides, in shapefile all the district names do not include district type in prefix but Bắc Kạn district name.

Id	City	district	ward
567	Bắc Kạn	Thành Phố Bắc Kạn	Đức Xuân
568	Bắc Kạn	Thành Phố Bắc Kạn	Dương Quang

Table 5. Example of records in unnormalized form in shapefile – district type in prefix
Source: Author compiled.

The national training center is a special admin unit. It is a military area and does not belong to any ward.

2.3 GSO list

In this report, we assign admin code of 2014 from GSO (2015) to the shapefile.

Fields	Description
ward_code	**Wards' National code**
district_code	Districts' National code
city_code	City national code
ward	Ward name
district	District name
city	City name
wardtype	Ward type

Table 6. Structure of GSO list 2014
Source: Authors compiled

In the GSO list, each record describes a ward including ward name, ward type, the district and the city where the ward locates. Each ward, district and city name have a corresponding national admin code. The field ward_code is the primary key for the file. There are no two distinct rows having the same values of ward_code.

ward _code	district _code	city _code	Ward	district	city	wardtype
1	1	1	Phường Phúc Xá	Quận Ba Đình	Thành phố Hà Nội	Phường
4	1	1	Phường Trúc Bạch	Quận Ba Đình	Thành phố Hà Nội	Phường
6	1	1	Phường Vĩnh Phúc	Quận Ba Đình	Thành phố Hà Nội	Phường
7	1	1	Phường Cống Vị	Quận Ba Đình	Thành phố Hà Nội	Phường

Table 7. Example of GSO list 2014 data
Source: Authors compiled.

Table 7 shows that the data in GSO list is not normalized. All the fields of ward, district and city are not atomic. The fields include both admin unit names and admin unit types. For example, the ward in in the first row is "Phường Phúc Xá". The first part "Phường" is ward type which mean "a ward" as in Figure 2. The second part "Phúc Xá" is the ward name which is like ward name in shapefile. Similarly, the district in the first row is "Quận Ba Đình". The part "Quận" means "Urban district" and "Ba Đình" is the name of the district.

If the fields are normalized, we have a foreign key to join with the shapefile. The normalization is to separate the admin unit type from the admin unit name in each filed. We have four new fields of city, district, ward and ward type which is the same the primary composite key in the shapefile. Thus, the four new fields can act as a foreign composite key in joining with the shape file.

The new four fields are also a candidate key. There are no two records having the same value of the four fields. In that case, the relationship between the shapefile and the GSO list is one-to-one on the four fields basis. It allows to assign each ward code in GSO list to a corresponding one and only one ward id in the shapefile.

Table 8 below compares statistic of shapefile and GSO list according to the administration hierarchy.

Admin unit	Shapefile	GSO 2014 list
City level: Provinces and Municipalities	63	63
District level: Municipality cities, Urban districts, Towns, Rural Districts and Provincial cities	710	704
Ward level: Wards, Communes, Townships and other	11,163	11,161
Ward	1,568	1,545
Commune	8,972	9,001
Town	601	615
Island	2	
National Training Center	2	

Table 8. Comparison of admin unit number between shapefile and GSO list
Source: Authors compiled

3 Methods

3.1 Objectives

The report assigns national admin codes including city code, district code, ward code to shapefile. The output is the new shapefile that have the structure as in Figure 3 below.

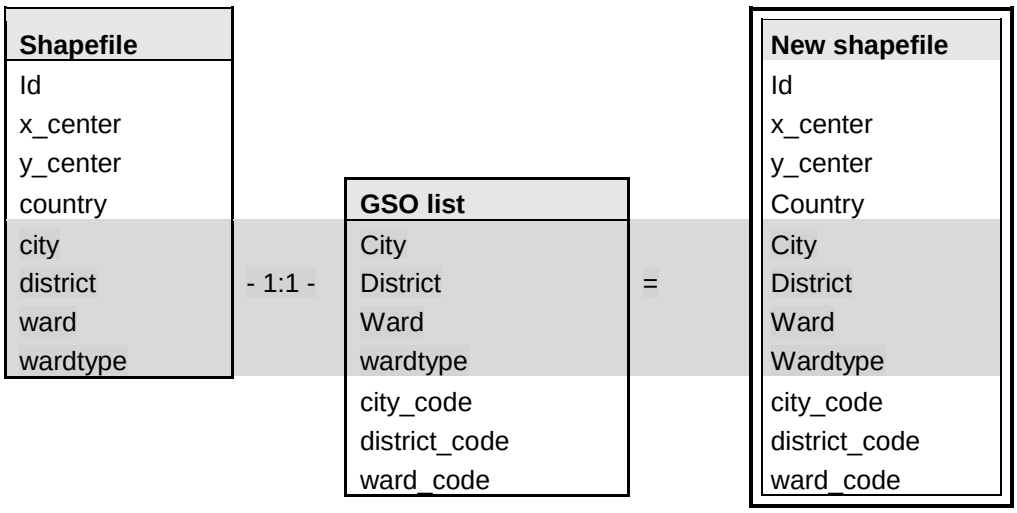


Figure 3. Objectives of the report in detail
Source: Authors compiled

3.2 Mismatches classification and resolution

We merge the shapefile with GSO list by matching the combination of city, district, ward and wardtype at each file. There are three possibilities where unmatched cases arise. We resolve them in this order.

- The first are the mismatches due to unnormalized form of the four fields in both files as mentioned in section 2.

- The second are the mismatches due to differences in writing style convention. This come from the fact that all the four fields are in string format.

- The third are the mismatches due to the administration changes. The GSO list is for 2014 while the shapefile is for 2018. From 2014 to 2018, some administration changes can arise such as changes in ward name, ward type or transfer from one district to another.

Thus, we propose overall organization of the procedure (implemented in Stata) to match shapefile and GSO list.

- Normalize both the shapefile and GSO list as indicated in section 2. The rule of thumb is that a field should contain only one attribute value, and not include the values of another field. In shape file, the ward type and district type are in parentheses removed from the ward name and district name. In GSO list, administration unit type is separated from admin unit name. In both files, the normalized “city, district, ward and wardtype” are stored in four new fields.

- Check that the normalized fields in both files are candidate key. The combination of the four fields are non-zero and unique.

- Initialize a map table with 10 columns. Four columns are for the normalized keys from GSO, four for the normalized keys from GADM, one for mismatches classification and one for comment. Determining the matching keys. Adding all the matches to the map table.

- Deal with the rows in GSO and GDAM files that are not in the map table. We match them manually case by case and add them to the map table.
  - We solve for cases of differences in writing style first.
  - Administration change cases are solved last with supported legal documents. Ideally, the changes in administration should be fill firstly in the map table since they are deterministic. However, we

can only do that if we have full information on administration changes from 2014 to 2018 at the beginning. That seems impossible to get. Therefore, we search for changes in admin unit case by case after solving all unmatched cases that we have more information.

- Use the map table to assign the ward codes, district code and city code to the shapefile

4 Results

4.1 Matching process

4.1.1 Normalize and check for candidate key

After normalization, checking procedure confirms that all new four files of city district ward wardtype of the two files still are candidate key. They are non-zero and unique in both files. After normalization, there are 10,953 cases are matched by normalized key. There are 418 unmatched cases, in which 210 cases are from shapefile and 208 cases from GSO list.

4.1.2 Deal with differences in writing style convention

In dealing with differences in writing style, we found that the differences are categorized to (i) leading zero in name; (ii) capitalization; (iii) tone marks position and (iv) others. Below are examples of those mismatch categories.

(i) Leading zeros in name

The ward/district names in GSO list are in form of 2 digits such as “01” if the names are numeric. Meanwhile, the name in shp has no leading zero such as “1”

File	Ward	Wardtype	district	city
Gso	06	Phường	Quận 4	Hồ Chí Minh
Shp	1	Phường	Quận 10	Hồ Chí Minh

Table 9. Example of differences in writing style – Leading zeros in name
Source: Author compiled.

(ii) Capitalization

Some mismatched cases come from the differences in upper and lower cases as the following example.

File	ward	Wardtype	District	city
gso	Đại Áng	Xã	Thanh Trì	Hà Nội
shp	Đại áng	Xã	Thanh Trì	Hà Nội

Table 10. Example of differences in writing style – Capitalization
Source: Author compiled.

(iii) Tone marks position

The unmatched cases come from the differences in the position of tone marks. According to Wikipedia (Accessed 2018-07-18), in Vietnamese:

> *In syllables where the vowel part consists of more than one vowel (such as diphthongs and triphthongs), the placement of the tone is still a matter of debate. Generally, there are two methodologies, an "old style" and a "new style". While the "old style" emphasizes aesthetics by placing the tone mark as close as possible to the center of the word (by placing the tone mark on the last vowel if an ending consonant part exists and on the next-to-last vowel if the ending consonant doesn't exist, as in hóa, hủy), the "new style" emphasizes linguistic principles and tries to apply the tone mark on the main vowel (as in hoá, huỷ). In both styles, when one vowel already has a quality diacritic on it, the tone mark must be applied to it as well, regardless of where it appears in the syllable (thus thuế is acceptable while thúê is not). In the case of the ươ diphthong, the mark is placed on the ơ. The u in qu is considered part of the consonant. Currently, the new style is usually used in textbooks published by Nhà Xuất bản Giáo dục, while most people still prefer the old style in casual uses.*

File	Ward	Wardtype	District	City
Shp	Hoà Long	Xã	Bà Rịa	Bà Rịa - Vũng Tàu
Gso	Hòa Long	Xã	Bà Rịa	Bà Rịa - Vũng Tàu
Shp	Phước Hoà	Xã	Tân Thành	Bà Rịa - Vũng Tàu
Gso	Phước Hòa	Xã	Tân Thành	Bà Rịa - Vũng Tàu
Shp	Tân Hoà	Xã	Tân Thành	Bà Rịa - Vũng Tàu
Gso	Tân Hòa	Xã	Tân Thành	Bà Rịa - Vũng Tàu

Table 11. Example of differences in writing style – Tone marks position
Source: Author compiled.

(iv) Others

The differences in this category mainly comes from the differences in transcription across ethnic group in Vietnamese. In these cases, the ward names have the same pronunciation but different in transcription such as “Bắc Ngà” and “Pắc Ngà”. These cases normally happen in mountain areas where there are many ethnic group live. The manual matching for the cases is not hard for Vietnamese but is difficult for foreign researchers. There are 22 out of 32 unmatched cases in category “Others” are due to the reason. The other cases are due to Roman numerals, blank space related mismatch or irregular character in ward name.

File	city_norm	district_norm	ward_norm	wardtype_norm	Note
shp	Cà Mau	Cà Mau	Tân Thành ()	Phường	Have parentheses in names
gso	Cà Mau	Cà Mau	Tân Thành	Phường	
shp	Gia Lai	Đăk Đoa	H'Neng	Xã	Blank space after apostrophe
gso	Gia Lai	Đăk Đoa	H' Neng	Xã	
shp	Đồng Tháp	Lấp Vò	Tân Khánh Trung	Xã	Redundant blank space between words
gso	Đồng Tháp	Lấp Vò	Tân Khánh Trung	Xã	
shp	Gia Lai	Mang Yang	Hà Ra	Xã	Same pronunciation but different transcription
gso	Gia Lai	Mang Yang	Hra	Xã	
shp	Sơn La	Bắc Yên	Bắc Ngà	Xã	
gso	Sơn La	Bắc Yên	Pắc Ngà	Xã	
shp	Sóc Trăng	Mỹ Xuyên	Hòa Tú 2	Xã	Roman numerals
gso	Sóc Trăng	Mỹ Xuyên	Hòa Tú II	Xã	

Table 12. Example of differences in writing style – Others
Source: Author compiled.

4.1.3 Deal with administration changes

During the period from 2014 to 2018, there are some administration changes. There is no change in city level. The changes are only at district and ward level. At district level, some districts changed their names, some were created from existing wards.

Shp				gso				
city	district	ward	wardtype	city	district	ward	wardtype	Legal documents
Bình Phước	Phú Riềng	Bình Sơn	Xã	Bình Phước	Bù Gia Mập	Bình Sơn	Xã	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Bình Tân	Xã	Bình Phước	Bù Gia Mập	Bình Tân	Xã	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015

Table 13. Example of administration changes – District changes
Source: Author compiled.

At ward level, some ward types changed such as from commune/town to wards. Some wards change their name going with changes in their type.

Shp				gso				
city	**district**	**Ward**	**wardtype**	**city**	**district**	**ward**	**wardtype**	**Legal documents**
Bạc Liêu	Giá Rai	1	Phường	Bạc Liêu	Giá Rai	Giá Rai	Thị trấn	Nghị quyết 930 ngày 15/5/2015
Bạc Liêu	Giá Rai	Hộ Phòng	Phường	Bạc Liêu	Giá Rai	Hộ Phòng	Thị trấn	Nghị quyết 930 ngày 15/5/2015
Quảng Nam	Điện Bàn	Điện Dương	Phường	Quảng Nam	Điện Bàn	Điện Dương	Xã	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015

Table 14. Example of administration changes – Ward changes
Source: Author compiled.

4.2 Matching results

With the four fileds of original files, there is no single matched case between GSO and Shapefile. After normalizing the databases format, 10,953 row pairs match. Resolving differences in writing style match another 141 row pairs. Accounting for administration changes match another 60. In total, 11,154 wards are matched. There are only 16 wards do not match, in which nine wards are from shape file and seven wards from GSO list.

Type of matched cases	Count	Percentage
Normalization	**10,953**	**98.06%**
Atomic normalization	10,953	98.06%
Differences in writing styles	**141**	**1.27%**
Leading zero in names	90	0.81%
Capitalization	8	0.07%
Tone mark position	10	0.09%
Others	33	0.30%
Administration changes	**60**	**0.53%**
Ward changes	35	0.31%
District changes	25	0.22%
Unmatched cases	**16**	**0.14%**
Unmatched	16	0.14%
Total	**11,170**	**100.00%**

Table 15. Matching results
Source: Author compiled.

Table 16 below shows the list of unmatched cases.

File	id	City	District	ward	wardtype
Shp	315	Bắc Giang	Lục Ngạn	Cấm Sơn	Trung tâm huấn luyện
Shp	350	Bắc Giang	Sơn Động	Cấm Sơn	Trung tâm huấn luyện
Shp	3096	Hải Phòng	Bạch Long Vĩ	Bạch Long Vĩ	Đảo
Shp	3314	Hậu Giang	Long Mỹ	Bình Thạnh	Phường
Shp	3320	Hậu Giang	Long Mỹ	Thuận An	Phường
Shp	3322	Hậu Giang	Long Mỹ	Vĩnh Tường	Phường
Shp	8515	Quảng Trị			
Shp	10556	Trà Vinh	Duyên Hải	1	Phường
Shp	10557	Trà Vinh	Duyên Hải	2	Phường
Gso		Hậu Giang	Long Mỹ	Long Mỹ	Thị trấn
Gso		Khánh Hòa	Trường Sa	Sinh Tồn	Xã
Gso		Khánh Hòa	Trường Sa	Song Tử Tây	Xã
Gso		Khánh Hòa	Trường Sa	Trường Sa	Thị trấn
Gso		Thanh Hóa	Nông Cống	Minh Thọ	Xã
Gso		Thanh Hóa	Đông Sơn	Đông Xuân	Xã
Gso		Trà Vinh	Duyên Hải	Duyên Hải	Thị trấn

Table 16. List of unmatched cases
Source: Author compiled.

The final updated shape file has the structure as the following.

id	x_center	y_center	ward _code	district _code	city _code	ward	wardtype	district	city
3497	106.7295	10.77221	27184	771	79	Phường 1	Phường	Quận 10	Hồ Chí Minh
3512	106.6499	10.75614	27247	772	79	Phường 1	Phường	Quận 11	Hồ Chí Minh
3550	106.5956	10.76355	27160	770	79	Phường 1	Phường	Quận 3	Hồ Chí Minh
3564	106.6545	10.7518	27298	773	79	Phường 1	Phường	Quận 4	Hồ Chí Minh
3579	106.6694	10.75343	27325	774	79	Phường 1	Phường	Quận 5	Hồ Chí Minh

Table 17. Example of final shapefile
Source: Author compiled.

5 Example of using shapefile with social economic data

The shapefile now has national administration codes. These codes are the key to plug social-economic data to shapefile. This part describes a small example of using shapefile and VHLSS with the national codes.

Suppose we doubt that in Vietnam, poor households live in hot areas in summer. If the hypothesis is true, it may suggest that the poor households are more vulnerable during summer thus, facing higher level of welfare inequality.

To check the hypothesis, we calculate the correlation between households' income per capita and the Cooling Degree Day (CDD) that the household face in June 2014 for simplicity. CDD is the amount of temperature that need to be cooled down to reach a certain base temperature for every day of a month. The higher CDD of an area is, the hotter the weather of the area is. In this example, 25^0C is chosen as the base. The formula of cdd25 is the following

Cdd25 = ∑(tavg-25) for all days of a month which have average daily temperature (tvag) higher than 25oC.

The data on household income per capita is extracted from VHLSS 2014. The data on temperature to calculate cdd25 comes from Global Historical Climatology Network (GHCN) of National Centers for Environmental Information (NOOA); GHCN provides daily temperature of 15 weather stations across Vietnam.

VHLSS data on income		GHCN data on temperature	
Variable	**Description**	**Variable**	**Description**
tinh	GSO code of province (city_code)	station	Station code
huyen	GSO code of district (district_code)	name	Station name
xa	GSO code of ward (ward_code)	latitude	Latitude of the station
diaban	Enumerator Area	longitude	Longitude of the station
hoso	Household number	cdd25*	Cooling degree day at 25^0C
inc_month	Average monthly Income per capita		

Note. * original GHCN provide data on daily temperature, cdd25 is calculated based on the temperature

Table 18. Data description of VHLSS and GHCN

Source: Author compiled.

Our mission is to assign cdd25 on GHCN data to each household in VHLSS. We employ shape file to carry out the task. The method contains two steps detailed in Figure 3. The step 1 is proximity matching. The y-center and x-center are the latitude and longitude of the central point of each ward which together determine the position of the ward. For each ward, we calculate the distance from the ward to each station and choose the cdd25 of the nearest station as the cdd25 of the ward. The distance is calculated basing on latitude and longitude. The second step is merging the ward with cdd25 to each household to get household income. It should be noted that city_code, district_code and ward_code in shapefile are tinh, huyen, xa in VHLSS corresponding.

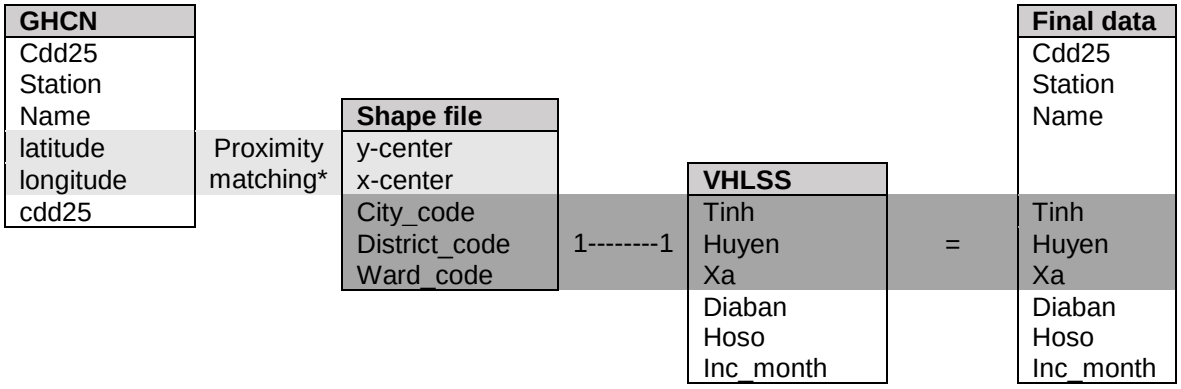


Figure 4. Method to assign cdd25 to each household

Source: Authors compiled

The Stata code for method in Figure 3 is provided in appendix B. The final data has the structure as the following.

Variable	Description
Station	Station code
Name	Station name
Cdđ25	Cooling degree day at the base of 25
Id	Id to merge with the polygon file
Tinh	GSO code of province
City	City name
Huyen	GSO code of district
District	District name
Xa	GSO code of ward
Ward	Ward name
Wardtype	Ward type
Diaban	Enumerator Area
Hoso	Household number
In_month	Average monthly Income per capita

Table 19. Structure of the final example data

Source: Authors compiled

With the above data, we can calculate the correlation between temperature and household income. Roughly speaking, we find no evidence for the hypothesis that in Vietnam, poor households live in hot areas.

	cdd25	inc_month
cdd25	1	
inc_month	0.0020	1
	(0.8489)	

Note. Number in parentheses is p-value.

Table 20. Pearson'R correlation between temperature and average monthly income in Vietnam, Jun 2014

6 Conclusion

In this report we assign official administration code to shapefile from GADM. As we wrote this, the GADM shapefile is the free and the most updated shapefile. However, it does not have official admin code for each admin unit. Without the admin code, data from shapefile cannot join with social economic data for analysis. Thus, we employ the official administration list from GSO 2014 as a medium to perform the task.

The assigning process was done by constructing a map table. The map table includes matched cases in admin unit names in shapefile and admin unit names in GSO list. The map table is filled in in three phases. The first is after the normalization of both files to ensure that each field contains only one information. The second is after dealing manually with differences in writing styles. The last one is after adjusting for administration changes from 2014 – the year of GSO list and 2018 – the year of shapefile. After the assigning process, there are 11,154 out of 11,163 wards (99.91%) assigned official admin codes. Only 16 cases are not matched between the shapefile and GSO list in which 9 cases are from shapefile.

The new shapefile with the official administration unit code are now available to plug in any social economic data in Vietnam. It can save time for researchers in doing spatial econometrics or graphing social-economic data at ward level. It is particularly useful for foreign researchers in analyzing Vietnamese data with shapefile since they do not have to match case by case manually in Vietnamese.

Though we already assigned ward codes to 99.91% of original GADM shapefile, we still have three points to improve in future. First, we only assigned GSO code at 2014 to the shape file. The reason is that GSO code 2014 is old enough to use with social-economic data at 2012 and at the same time, is updated enough to use with the lastest data at 2016. But soon, GSO codes at 2014 will be outdated when social-economic data for 2018 comes out. Thus, in next version, we will add GSO code at 2016 and 2018 to the shapefile.

Second, we matched manually some cases that is differences in writing stypes while these cases should be matched by program script. In this report, it is reasonable since the number of cases in this category is relatively small. There are 51 cases in total including 8 cases of capitalization, 10 cases of tone mark position and 33 special cases. However, in the next version, when the number of unmatched cases in this category may increase, we will develop script to handle these cases.

Finally, there are still 16 unmatched cases in which nine come from the original GADM shapefile. We highly appreciate any comment or feedback that help us to solve the unmatched cases. Please let us know if you have any idea on the issue. Thank you in advance for your support!

References

Creative Commons. n.d. “Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).” Accessed July 18, 2018. https://creativecommons.org/licenses/by-nc/4.0/.

Crow, Kevin. 2015. *SHP2DTA: Stata Module to Converts Shape Boundary Files to Stata Datasets*. https://EconPapers.repec.org/RePEc:boc:bocode:s456718.

Environmental Systems Research Institute. 1998. “ESRI Shapefile Technical Description.” ESRI White Paper. Environmental Systems Research Institute. https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf.

GADM. 2018. “GADM Shapefile of Vietnam. Version 3.4.” Global Administrative Areas (GADM). https://biogeo.ucdavis.edu/data/gadm3.6/shp/gadm36_VNM_shp.zip.

GSO. 2015. “GSO List of Administratinon Unit at Dec 31 2014.” General Statistics Office Of Vietnam (GSO). https://www.gso.gov.vn/dmhc2015/.

Vietnamese National Assembly. 2013. *Vietnam’s 2013 Constitution*. http://moj.gov.vn/vbpq/lists/vn%20bn%20php%20lut/view_detail.aspx?itemid=28814.

———. 2015. *Law on organizing the local government*. *77/2015/QH13*. https://thuvienphapluat.vn/van-ban/Bo-may-hanh-chinh/Law-No-77-2015-QH13-organizing-the-local-Government-284413.aspx.

Wikipedia. n.d. “Vietnamese Alphabet.” Wikipedia. Accessed July 18, 2018. https://en.wikipedia.org/wiki/Vietnamese_alphabet#Tone_marks.

Appendix A. **The list of manual matching cases in map table**.

shp_city_norm	shp_district_norm	shp_ward_norm	shp_wardtype_norm	gso_city_norm	gso_district_norm	gso_ward_norm	gso_wardtype_norm	type	comment
Bà Rịa - Vũng Tàu	Bà Rịa	Hòa Long	Xã	Bà Rịa - Vũng Tàu	Bà Rịa	Hoà Long	Xã	4	
Bà Rịa - Vũng Tàu	Tân Thành	Phước Hòa	Xã	Bà Rịa - Vũng Tàu	Tân Thành	Phước Hoà	Xã	4	
Bà Rịa - Vũng Tàu	Tân Thành	Tân Hòa	Xã	Bà Rịa - Vũng Tàu	Tân Thành	Tân Hoà	Xã	4	
Bình Phước	Phú Riềng	Bình Sơn	Xã	Bình Phước	Bù Gia Mập	Bình Sơn	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Bình Tân	Xã	Bình Phước	Bù Gia Mập	Bình Tân	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Bù Nho	Xã	Bình Phước	Bù Gia Mập	Bù Nho	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Long Bình	Xã	Bình Phước	Bù Gia Mập	Long Bình	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Long Hà	Xã	Bình Phước	Bù Gia Mập	Long Hà	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Long Hưng	Xã	Bình Phước	Bù Gia Mập	Long Hưng	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Long Tân	Xã	Bình Phước	Bù Gia Mập	Long Tân	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Phú Riềng	Xã	Bình Phước	Bù Gia Mập	Phú Riềng	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Phú Trung	Xã	Bình Phước	Bù Gia Mập	Phú Trung	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bình Phước	Phú Riềng	Phước Tân	Xã	Bình Phước	Bù Gia Mập	Phước Tân	Xã	6	Nghị quyết 931/NQ-UBTVQH ngày 15/5/2015
Bạc Liêu	Giá Rai	1	Phường	Bạc Liêu	Giá Rai	Giá Rai	Thị trấn	5	Nghị quyết 930 ngày 15/5/2015
Bạc Liêu	Giá Rai	Hộ Phòng	Phường	Bạc Liêu	Giá Rai	Hộ Phòng	Thị trấn	5	Nghị quyết 930 ngày 15/5/2015
Bạc Liêu	Giá Rai	Láng Tròn	Phường	Bạc Liêu	Giá Rai	Phong Thạnh Đông A	Xã	5	Nghị quyết 930 ngày 15/5/2015

Bắc Giang	Lục Ngạn	Cấm Sơn	Trung tâm huấn luyện					8	
Bắc Giang	Sơn Động	Cấm Sơn	Trung tâm huấn luyện					8	
Bắc Kạn	Bắc Kạn	Huyền Tụng	Phường	Bắc Kạn	Bắc Kạn	Huyền Tụng	Xã	5	Nghị quyết 892/NQ-UBTVQH13 ngày 11/3/2015
Bắc Kạn	Bắc Kạn	Xuất Hóa	Phường	Bắc Kạn	Bắc Kạn	Xuất Hoá	Xã	5	Nghị quyết 892/NQ-UBTVQH13 ngày 11/3/2015
Cao Bằng	Phục Hoà	Triệu Ẩu	Xã	Cao Bằng	Phục Hoà	Triệu ẩu	Xã	3	
Cà Mau	Cà Mau	Tân Thành ()	Phường	Cà Mau	Cà Mau	Tân Thành	Phường	7	
Cà Mau	Đầm Dơi	Tạ An Khương Nam	Xã	Cà Mau	Đầm Dơi	Tạ An Khương Nam	Xã	7	
Cà Mau	Đầm Dơi	Tạ An Khương Đông	Xã	Cà Mau	Đầm Dơi	Tạ An Khương Đông	Xã	7	
Gia Lai	Mang Yang	Hà Ra	Xã	Gia Lai	Mang Yang	Hra	Xã	7	
Gia Lai	Đăk Đoa	H'Neng	Xã	Gia Lai	Đăk Đoa	H' Neng	Xã	7	
Hà Nội	Hoàn Kiếm	Chương Dương Độ	Phường	Hà Nội	Hoàn Kiếm	Chương Dương	Phường	7	
Hà Nội	Thanh Trì	Đại Áng	Xã	Hà Nội	Thanh Trì	Đại áng	Xã	3	
Hà Tĩnh	Kỳ Anh	Kỳ Liên	Phường	Hà Tĩnh	Kỳ Anh	Kỳ Liên	Xã	5	Nghị quyết 903/NQ-UBTVQH13 ngày 10/4/2015
Hà Tĩnh	Kỳ Anh	Kỳ Long	Phường	Hà Tĩnh	Kỳ Anh	Kỳ Long	Xã	5	Nghị quyết 903/NQ-UBTVQH13 ngày 10/4/2015
Hà Tĩnh	Kỳ Anh	Kỳ Phương	Phường	Hà Tĩnh	Kỳ Anh	Kỳ Phương	Xã	5	Nghị quyết 903/NQ-UBTVQH13 ngày 10/4/2015
Hà Tĩnh	Kỳ Anh	Kỳ Thịnh	Phường	Hà Tĩnh	Kỳ Anh	Kỳ Thịnh	Xã	5	Nghị quyết 903/NQ-UBTVQH13 ngày 10/4/2015
Hà Tĩnh	Kỳ Anh	Kỳ Trinh	Phường	Hà Tĩnh	Kỳ Anh	Kỳ Trinh	Xã	5	Nghị quyết 903/NQ-UBTVQH13 ngày 10/4/2015

Hà Tĩnh	Kỳ Anh	Sông Trí	Phường	Hà Tĩnh	Kỳ Anh	Kỳ Anh	Thị trấn	5	Nghị quyết 903/NQ-UBTVQH13 ngày 10/4/2015
Hải Phòng	Bạch Long Vĩ	Bạch Long Vĩ	Đảo					8	
Hậu Giang	Long Mỹ	Bình Thạnh	Phường					8	
				Hậu Giang	Long Mỹ	Long Mỹ	Thị trấn	8	
Hậu Giang	Long Mỹ	Thuận An	Phường					8	
Hậu Giang	Long Mỹ	Trà Lồng	Phường	Hậu Giang	Long Mỹ	Trà Lồng	Thị trấn	5	http://baochinhphu.vn/Hoat-dong-dia-phuong/Thanh-lap-thi-xa-Long-My-tinh-Hau-Giang/234611.vgp
Hậu Giang	Long Mỹ	Vĩnh Tường	Phường					8	
Hồ Chí Minh	1	Cầu kho	Phường	Hồ Chí Minh	1	Cầu Kho	Phường	3	
Hồ Chí Minh	12	Tân Chánh Hiệp	Xã	Hồ Chí Minh	12	Tân Chánh Hiệp	Phường	5	Quyết định số 1195/QĐ-UB ngày 18/3/1997
Hồ Chí Minh	Cần Giờ	Long Hoà	Xã	Hồ Chí Minh	Cần Giờ	Long Hòa	Xã	4	
Hồ Chí Minh	Củ Chi	Phú Hoà Đông	Xã	Hồ Chí Minh	Củ Chi	Phú Hòa Đông	Xã	4	
Hồ Chí Minh	Tân Phú	Hoà Thạnh	Phường	Hồ Chí Minh	Tân Phú	Hòa Thạnh	Phường	4	
Hồ Chí Minh	Tân Phú	Tân Thới Hoà	Phường	Hồ Chí Minh	Tân Phú	Tân Thới Hòa	Phường	4	
				Khánh Hòa	Trường Sa	Sinh Tồn	Xã	8	
				Khánh Hòa	Trường Sa	Song Tử Tây	Xã	8	
				Khánh Hòa	Trường Sa	Trường Sa	Thị trấn	8	
Kon Tum	Ia H' Drai	Ia Dal	Xã	Kon Tum	Sa Thầy	Ia Dal	Xã	6	Nghị quyết 890/NQ-UBTVQH13 ngày 11/3/2015
Kon Tum	Ia H' Drai	Ia Dom	Xã	Kon Tum	Sa Thầy	Ia Dom	Xã	6	Nghị quyết 890/NQ-UBTVQH13 ngày 11/3/2015
Kon Tum	Ia H' Drai	Ia Tơi	Xã	Kon Tum	Sa Thầy	Ia Tơi	Xã	6	Nghị quyết 890/NQ-UBTVQH13 ngày 11/3/2015
Kon Tum	Ngọc Hồi	Plei Cần	Thị trấn	Kon Tum	Ngọc Hồi	Plei Kần	Thị trấn	7	
Kon Tum	Ngọc Hồi	Pờ Y	Xã	Kon Tum	Ngọc Hồi	Bờ Y	Xã	7	
Kon Tum	Tu Mơ Rông	Ngok Yêu	Xã	Kon Tum	Tu Mơ Rông	Ngọc Yêu	Xã	7	

Kon Tum	Tu Mơ Rông	Ngọk Lây	Xã	Kon Tum	Tu Mơ Rông	Ngọc Lây	Xã	7	
Kon Tum	Tu Mơ Rông	Đắk Tơ Kan	Xã	Kon Tum	Tu Mơ Rông	Đắk Tờ Kan	Xã	7	
Kon Tum	Đắk Glei	Đắk Plô	Xã	Kon Tum	Đắk Glei	Đắk Blô	Xã	7	
Lai Châu	Mường Tè	Pa Ủ	Xã	Lai Châu	Mường Tè	Pa ủ	Xã	3	
Long An	Bến Lức	Thạnh Phú	Xã	Long An	Bến Lức	Thanh Phú	Xã	7	
Long An	Tân Thạnh	Nhơn Hòa	Xã	Long An	Tân Thạnh	Nhơn Hoà	Xã	4	
Lạng Sơn	Bắc Sơn	Đồng Ý	Xã	Lạng Sơn	Bắc Sơn	Đồng ý	Xã	3	
Ninh Bình	Tam Điệp	Yên Bình	Phường	Ninh Bình	Tam Điệp	Yên Bình	Xã	5	Nghị Quyết số 904/NQ-UBNTQH13 ngày 10/4/2015
Quảng Nam	Điện Bàn	Vĩnh Điện	Phường	Quảng Nam	Điện Bàn	Vĩnh Điện	Thị trấn	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Nam	Điện Bàn	Điện An	Phường	Quảng Nam	Điện Bàn	Điện An	Xã	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Nam	Điện Bàn	Điện Dương	Phường	Quảng Nam	Điện Bàn	Điện Dương	Xã	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Nam	Điện Bàn	Điện Nam Bắc	Phường	Quảng Nam	Điện Bàn	Điện Nam Bắc	Xã	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Nam	Điện Bàn	Điện Nam Trung	Phường	Quảng Nam	Điện Bàn	Điện Nam Trung	Xã	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Nam	Điện Bàn	Điện Nam Đông	Phường	Quảng Nam	Điện Bàn	Điện Nam Đông	Xã	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Nam	Điện Bàn	Điện Ngọc	Phường	Quảng Nam	Điện Bàn	Điện Ngọc	Xã	5	Quyết định số 889/NQ-UBTVQH13 ngày 11/3/2015
Quảng Ngãi	Nghĩa Hành	Hành Tín Đông	Xã	Quảng Ngãi	Nghĩa Hành	Hành Tín Đông	Xã	7	
Quảng Ninh	Đông Triều	Hưng Đạo	Phường	Quảng Ninh	Đông Triều	Hưng Đạo	Xã	5	Nghị quyết 891/NQ-UBTVQH13 ngày 11/3/2015
Quảng Ninh	Đông Triều	Kim Sơn	Phường	Quảng Ninh	Đông Triều	Kim Sơn	Xã	5	Nghị quyết 891/NQ-UBTVQH13 ngày 11/3/2015
Quảng Ninh	Đông Triều	Mạo Khê	Phường	Quảng Ninh	Đông Triều	Mạo Khê	Thị trấn	5	Nghị quyết 891/NQ-UBTVQH13 ngày 11/3/2015
Quảng Ninh	Đông Triều	Xuân Sơn	Phường	Quảng Ninh	Đông Triều	Xuân Sơn	Xã	5	Nghị quyết 891/NQ-UBTVQH13 ngày 11/3/2015

Quảng Ninh	Đông Triều	Đông Triều	Phường	Quảng Ninh	Đông Triều	Đông Triều	Thị trấn	5	Nghị quyết 891/NQ-UBTVQH13 ngày 11/3/2015
Quảng Ninh	Đông Triều	Đức Chính	Phường	Quảng Ninh	Đông Triều	Đức Chính	Xã	5	Nghị quyết 891/NQ-UBTVQH13 ngày 11/3/2015
Quảng Trị	Cồn Cỏ	Cồn Cỏ	Đảo					8	
Sóc Trăng	Mỹ Xuyên	Hòa Tú 2	Xã	Sóc Trăng	Mỹ Xuyên	Hòa Tú II	Xã	7	
Sóc Trăng	Trần Đề	Đại Ân 2	Xã	Sóc Trăng	Trần Đề	Đại Ân 2	Xã	7	
Sơn La	Bắc Yên	Bắc Ngà	Xã	Sơn La	Bắc Yên	Pắc Ngà	Xã	7	
Sơn La	Bắc Yên	Xín Vàng	Xã	Sơn La	Bắc Yên	Xím Vàng	Xã	7	
Sơn La	Mai Sơn	Cò Nòi	Xã	Sơn La	Mai Sơn	Cò Nòi	Xã	7	
Sơn La	Mai Sơn	Mường Chanh	Xã	Sơn La	Mai Sơn	Mương Chanh	Xã	7	
Sơn La	Mai Sơn	Nà Ớt	Xã	Sơn La	Mai Sơn	Nà Ơt	Xã	7	
Sơn La	Mộc Châu	Quy Hướng	Xã	Sơn La	Mộc Châu	Qui Hướng	Xã	7	
Sơn La	Quỳnh Nhai	Nậm Ét	Xã	Sơn La	Quỳnh Nhai	Nậm ét	Xã	3	
Sơn La	Thuận Châu	Muội Nọi	Xã	Sơn La	Thuận Châu	Muổi Nọi	Xã	7	
Sơn La	Thuận Châu	Mường É	Xã	Sơn La	Thuận Châu	Mường é	Xã	3	
Sơn La	Thuận Châu	Nong Lay	Xã	Sơn La	Thuận Châu	Noong Lay	Xã	7	
Sơn La	Thuận Châu	Phỏng Lái	Xã	Sơn La	Thuận Châu	Phổng Lái	Xã	7	
Sơn La	Thuận Châu	Phỏng Lập	Xã	Sơn La	Thuận Châu	Phổng Lập	Xã	7	
Sơn La	Thuận Châu	Thôn Mòn	Xã	Sơn La	Thuận Châu	Thôm Mòn	Xã	7	
Sơn La	Thuận Châu	Tòng Cọ	Xã	Sơn La	Thuận Châu	Tông Cọ	Xã	7	
Sơn La	Thuận Châu	Tòng Lệnh	Xã	Sơn La	Thuận Châu	Tông Lạnh	Xã	7	
Sơn La	Thuận Châu	É Tòng	Xã	Sơn La	Thuận Châu	é Tòng	Xã	3	
				Thanh Hóa	Nông Cống	Minh Thọ	Xã	8	
Thanh Hóa	Sầm Sơn	Quảng Châu	Phường	Thanh Hóa	Quảng Xương	Quảng Châu	Xã	6	Nghị quyết 39/2016/NQ-HĐND ngày 8/12/2016
Thanh Hóa	Sầm Sơn	Quảng Cư	Phường	Thanh Hóa	Sầm Sơn	Quảng Cư	Xã	5	Nghị quyết 39/2016/NQ-HĐND ngày 8/12/2016
Thanh Hóa	Sầm Sơn	Quảng Hùng	Xã	Thanh Hóa	Quảng Xương	Quảng Hùng	Xã	6	Nghị quyết 91/2014/NQ-HĐND ngày 4/7/2014
Thanh Hóa	Sầm Sơn	Quảng Minh	Xã	Thanh Hóa	Quảng Xương	Quảng Minh	Xã	6	Nghị quyết 91/2014/NQ-HĐND ngày 4/7/2014

Thanh Hóa	Sầm Sơn	Quảng Thọ	Phường	Thanh Hóa	Quảng Xương	Quảng Thọ	Xã	6	Nghị quyết 39/2016/NQ-HĐND ngày 8/12/2016
Thanh Hóa	Sầm Sơn	Quảng Vinh	Phường	Thanh Hóa	Quảng Xương	Quảng Vinh	Xã	6	Nghị quyết 39/2016/NQ-HĐND ngày 8/12/2016
Thanh Hóa	Sầm Sơn	Quảng Đại	Xã	Thanh Hóa	Quảng Xương	Quảng Đại	Xã	6	Nghị quyết 91/2014/NQ-HĐND ngày 4/7/2014
				Thanh Hóa	Đông Sơn	Đông Xuân	Xã	8	
Thái Nguyên	Phổ Yên	Ba Hàng	Phường	Thái Nguyên	Phổ Yên	Ba Hàng	Thị trấn	5	Nghị quyết sô 932/NQ-UBTVQH13 ngày 15/5/2015
Thái Nguyên	Phổ Yên	Bãi Bông	Phường	Thái Nguyên	Phổ Yên	Bãi Bông	Thị trấn	5	Nghị quyết sô 932/NQ-UBTVQH13 ngày 15/5/2015
Thái Nguyên	Phổ Yên	Bắc Sơn	Phường	Thái Nguyên	Phổ Yên	Bắc Sơn	Thị trấn	5	Nghị quyết sô 932/NQ-UBTVQH13 ngày 15/5/2015
Thái Nguyên	Phổ Yên	Đồng Tiến	Phường	Thái Nguyên	Phổ Yên	Đồng Tiến	Xã	5	Nghị quyết sô 932/NQ-UBTVQH13 ngày 15/5/2015
Thái Nguyên	Sông Công	Lương Sơn	Phường	Thái Nguyên	Thái Nguyên	Lương Sơn	Xã	5	Nghị quyết sô 932/NQ-UBTVQH13 ngày 15/5/2015
Thái Nguyên	Thái Nguyên	Chùa Hang	Phường	Thái Nguyên	Đồng Hỷ	Chùa Hang	Thị trấn	5	Nghị quyết số 422/NQ-UBTVQH14 ngày 18/8/2017
Thái Nguyên	Thái Nguyên	Huống Thượng	Xã	Thái Nguyên	Đồng Hỷ	Huống Thượng	Xã	6	Nghị quyết số 422/NQ-UBTVQH14 ngày 18/8/2017
Thái Nguyên	Thái Nguyên	Linh Sơn	Xã	Thái Nguyên	Đồng Hỷ	Linh Sơn	Xã	6	Nghị quyết số 422/NQ-UBTVQH14 ngày 18/8/2017
Thái Nguyên	Thái Nguyên	Sơn Cẩm	Xã	Thái Nguyên	Phú Lương	Sơn Cẩm	Xã	6	Nghị quyết số 422/NQ-UBTVQH14 ngày 18/8/2017
Thái Nguyên	Thái Nguyên	Đồng Bẩm	Phường	Thái Nguyên	Thái Nguyên	Đồng Bẩm	Xã	5	Nghị quyết số 422/NQ-UBTVQH14 ngày 18/8/2017
Thái Nguyên	Thái Nguyên	Đồng Liên	Xã	Thái Nguyên	Phú Bình	Đồng Liên	Xã	6	Nghị quyết số 422/NQ-UBTVQH14 ngày 18/8/2017
Thừa Thiên Huế	Huế	Đúc	Phường	Thừa Thiên Huế	Huế	Phường Đúc	Phường	7	
Tiền Giang	Gò Công Đông	Tăng Hòa	Xã	Tiền Giang	Gò Công Đông	Tăng Hoà	Xã	4	
Tiền Giang	Tân Phước	Thạnh Hòa	Xã	Tiền Giang	Tân Phước	Thạnh Hoà	Xã	4	
Trà Vinh	Duyên Hải	1	Phường					8	
Trà Vinh	Duyên Hải	2	Phường					8	

				Trà Vinh	Duyên Hải	Duyên Hải	Thị trấn	8	
Trà Vinh	Duyên Hải	Đôn Châu	Xã	Trà Vinh	Trà Cú	Đôn Châu	Xã	6	Nghị quyết sô 934/NQ-UBTVQH13 ngày 15/5/2015
Trà Vinh	Duyên Hải	Đôn Xuân	Xã	Trà Vinh	Trà Cú	Đôn Xuân	Xã	6	Nghị quyết sô 934/NQ-UBTVQH13 ngày 15/5/2015
Đà Nẵng	Hải Châu	Hải Châu I	Phường	Đà Nẵng	Hải Châu	Hải Châu I	Phường	7	
Đắk Lắk	Ea Kar	Cư Yang	Xã	Đắk Lắk	Ea Kar	Cư Jang	Xã	7	
Đồng Tháp	Lấp Vò	Tân Khánh Trung	Xã	Đồng Tháp	Lấp Vò	Tân Khánh Trung	Xã	7	

Annex 2. Résumé en français

Chapitre 1. Introduction

Comprendre pour maîtriser la demande d'électricité des ménages est essentiel pour harmoniser la croissance économique, la sécurité énergétique et la réduction des émissions de gaz à effet de serre, une urgence particulière pour un pays en développement rapide comme le Vietnam au début du 21e siècle.

Premièrement, les outils de gestion axés sur la demande, tels que l'augmentation des tarifs de l'électricité, encouragent les ménages à l'économiser, ce qui réduit les émissions de dioxyde de carbone associées à la génération d'électricité. C'est particulièrement important pour les pays dont une grande partie de l'approvisionnement en électricité repose sur le charbon. C'est le cas de nombreux pays en développement, en particulier en Asie du Sud-Est. Le Vietnam, l'Indonésie et les Philippines utilisent principalement le charbon pour assurer la continuité de leur approvisionnement en énergie.

Deuxièmement, la maîtrise de la demande aide également les pays en croissance rapide à assurer la sécurité énergétique en limitant l'explosion de la demande d'électricité. Ces pays font face à des situations sur les marchés de l'électricité plus chaotiques que les pays développés. Dans les pays développés, le marché est bien établi et la demande est relativement stable. Ils peuvent penser la transition énergétique comme une substitution progressive des sources d'approvisionnement vers une structure comportant une plus grande proportion de sources renouvelables. Dans un pays comme le Vietnam, la demande augmentant en raison de la croissance économique et démographique, le marché connaît une croissance rapide dépassant parfois 10 % par an. La demande peut dépasser la capacité de production électrique et entraîner des pannes de courant. La maîtrise de la demande contribue à éviter de tels problèmes.

La mise en œuvre de la maîtrise de la demande pour les marchés résidentiels de l'électricité nécessite une compréhension approfondie des comportements des clients et de la demande des ménages. La question n'est pas nouvelle, mais dans le passé ce sont surtout des données agrégées qui ont été utilisées pour explorer les facteurs ayant une incidence pour la demande d'électricité résidentielle (Houthakker, Verleger et Sheehan, 1974; Hsing, 1994; Holtedahl et Joutz, 2004; Alberini et Filippini, 2011). Or, deux développements récents posent de nouveaux défis pour l'estimation de la demande d'électricité.

- Premièrement, les études passent des données agrégées vers des données micro au niveau des ménages, mais celles ci sont souvent lacunaires en ce qui concerne les prix (Branch, 1993; Alberini, Gans et Velez-Lopez, 2011), ou alors ramenées au niveau régional plutôt qu'à un niveau national, en raison de l'absence de données nationales au sujet des structures tarifaires (Reiss et White, 2005; Zhou et Teng, 2013).
- Deuxièmement, le changement climatique augmente l'occurrence des vagues de chaleur, un facteur explicatif de la demande d'électricité insuffisamment étudié dans la littérature.

Cette thèse revisite l'histoire de l'estimation de la demande d'électricité des ménages en s'attaquant à ces deux défis. Elle mesure en termes économétriques l'effet des principaux facteurs explicatifs de la demande d'électricité des ménages. Ces facteurs sont : le prix de l'électricité (en tarification progressive par tranches) ; le revenu des ménages ; la démographie (taille et composition du ménage) ; et la température avec les vagues de chaleur.
Les travaux portent sur les ménages du Vietnam pour la période 2012-2016. Quatre traits caractérisent ce contexte.

Premièrement, le Vietnam est un pays tropical connaissant des canicules estivales de plus en plus fréquentes et intenses. La pointe de la demande est en début d'après midi en été. Il est donc pertinent d'y étudier l'impact des vagues de chaleur sur la demande d'électricité.

Deuxièmement, l'enquête nationale sur les conditions de vie des ménages au Vietnam utilise un échantillon tournant, renouvelé par tiers à chaque campagne biennale. Cela permet de construire un ensemble de données de panel à partir de trois séries dans trois années différentes, ainsi que de constituer un ensemble de données groupées à partir des autres. La séparation des données de panel et en données groupées est idéale pour estimer la demande en électricité à court et à long terme.

Troisièmement, le marché de l'électricité résidentielle au Vietnam est un monopole avec un seul vendeur, Vietnam Electricity (EVN). La grille tarifaire de l'électricité – proposée par EVN et finalisée par le gouvernement – est par principe uniforme à l'échelle nationale, même si elle varie dans le temps. Cela permet d'estimer la fonction de demande à partir de données d'enquête micro nationale, en connaissant les prix de l'électricité.

Quatrièmement enfin, le Vietnam présente les caractéristiques d'un pays en développement rapide, où la maîtrise de la demande joue un rôle croissant. En raison du rythme rapide de la croissance économique, la demande en électricité au Vietnam a augmenté, ce qui pose des problèmes pour assurer la sécurité énergétique ainsi que pour développer des sources d'énergie renouvelables. Entre 2006 et 2015, la consommation nationale d'électricité a augmenté à un taux moyen supérieur à 10% (MOIT et DEA, 2017). La demande d'électricité d'ici 2035 devrait croître à un taux annuel moyen de 8% (MOIT et DEA, 2017). Près de la moitié de la nouvelle capacité proposée devrait être alimentée au charbon (MOIT et DEA, 2017). En réponse à cette situation, le gouvernement vietnamien considère l'efficacité énergétique comme un "premier carburant". Les économies d'électricité sont potentiellement estimées à 17% d'ici 2030 (MOIT et DEA, 2017). En pratique, le gouvernement a mis en œuvre des mesures telles que l'augmentation du tarif forfaitaire pour la consommation résidentielle afin d'encourager les personnes à économiser de l'électricité (EVN, 2015).

Les résultats de cette thèse constituent une référence pour les décideurs politiques lorsqu'ils doivent considérer les mesures de gestion de la demande au Vietnam. Mais au-delà, ils élargissent notre compréhension des fonctions de la demande d'électricité résidentielle dans les pays tropicaux en développement rapide, lieu clé de la lutte contre le changement climatique.

La thèse est structurée comme suit. Le chapitre 2 le contexte politique et la structure générale des marchés de l'électricité au Vietnam. Le chapitre 3 examine la littérature. Il passe en revue les modèles théoriques et empiriques existants décrivant les fonctions de demande d'électricité des ménages à court et à long terme. Le chapitre 4 expose comment nous avons construit notre jeux de données. Les quatre chapitres suivants abordent successivement les quatre principaux facteurs explicatifs de la demande d'électricité: prix, revenu, démographie, température.

Chapitre 2. Contexte politique

Pour compléter les informations de base introduites à la section précédentes, précisons les éléments de politique énergétique et sociale déterminant le cadre dans lequel s'exerce la demande d'électricité des ménages, à savoir la structure des prix et des subventions.

En ce qui concerne la structure de prix de l'électricité, il existe deux régimes de prix pour l'électricité résidentielle au Vietnam. Tous deux relèvent de la tarification progressive par

tranches (Increasing Block Tariff, IBT). Le premier régime est une tarification IBT qui s'applique aux ménages achetant de l'électricité directement à EVN. Le second s'applique dans les zones rurales isolées, à faible densité de population et à faible infrastructure. Dans ces régions, EVN vend de l'électricité à des organisations de distribution d'électricité rurales selon un barème de prix de gros. Ces organisations revendent de l'électricité aux ménages avec leur propre barème de prix fondé sur les prix de gros. La fixation des prix de ces organisations est hétérogène. Certaines appliquent un prix unique, tandis que d'autres appliquent une tarification à trois tranches, voire plus.

Cependant, le second régime ne concerne qu'une petite fraction des ménages. En 2014, EVN fournissait directement de l'électricité à 84,57% des communes et à 82,59% des ménages ruraux (Thục Quyên, 2014). En 2015, EVN a fourni directement de l'électricité à 87,88% des ménages ruraux des provinces du sud (Mai Phương, 2015).

En ce qui concerne les subventions à l'électricité résidentielle, le gouvernement vietnamien active à la fois une politique de subvention du prix de l'électricité et un programme de support monétaire direct. La politique de subvention du prix de l'électricité consiste à fixer le prix de la première tranche en dessous du prix de vente moyen (environ 95%). Les bénéficiaires de cette politiques sont tous les ménages à faible revenu et ceux qui bénéficient d'une politique de traitement préférentiel. Le programme de support monétaire direct fixe le montant de la subvention à 30 kWh aux prix en vigueur.

Chapitre 3. Spécification du modèle

Cette thèse étudie le rôle de facteurs via l'estimation d'une forme commune de fonction de demande. Chaque facteur est étudié en détail dans un chapitre distinct. Contrairement aux autres produits consommés par les ménages comme l'eau ou la nourriture, l'électricité en elle-même ne génère pas d'utilité pour les consommateurs. L'utilisation de l'électricité nécessite des appareils tels que des ventilateurs, des climatiseurs (AC) et des téléviseurs (TV). Ainsi, la demande en électricité peut être considérée comme une demande dérivée qui dépend de la demande des appareils ménagers possédés par les ménages (Taylor, 1975).

Étant donné que la demande d'électricité dépend de la disponibilité des appareils, il existe deux types de fonctions de demande: à court et à long terme. Une fonction de demande à court terme est définie par une condition selon laquelle le stock d'appareils est fixe (Taylor, 1975, p. 80), tandis que le stock d'appareils à long terme peut varier (Taylor, 1975, p. 80).

Jusqu'à présent, la plupart des modèles de demande d'électricité au niveau des ménages sont construits sur la base des idées de deux modèles. Fisher et Kaysen (1962) ont publié la première étude qui établit une distinction explicite entre les fonctions de demande d'électricité à court et à long terme. Houthakker et Taylor (1970), à partir de l'idée du stock de capital selon Fisher et Kaysen, ont élaboré un modèle complet de la demande d'électricité résidentielle à court et à long terme.
La fonction de demande à court terme est la suivante :

$$lnE = \alpha_0 + \alpha_1 lny + \alpha_2 lnP + \alpha_3 Z + \alpha_{4j} A_j D_j + \varepsilon \qquad (1)$$

Où
E = consommation d'électricité du ménage
y = revenu du ménage (par habitant)
P = prix
Z = variables connexes telles que CDD / HDD, prix du gaz

A_j = vecteur de l'appareil
D_j = dummy vars = 1 si le ménage est propriétaire de l'appareil j, = 0 sinon

La fonction de demande à long terme est la suivante :

$$lnE_{it} = \alpha + \beta lnp_{it} + \gamma lny_{it} + \theta Z_{it} + \delta lnE_{i,t-1} + \varepsilon_t \quad (2)$$

Où
E_{it} = Consommation d'électricité à la période *t*
$E_{i,t-1}$ = consommation d'électricité à la période *t-1*
p_{it}= prix de l'électricité à la période *t*
y_{it} = revenu du ménage (par habitant) à la période *t*
Z_{it} = un vecteur d'un autre facteur tel que le prix du gaz, le CDD / HDD

Chapitre 4. Données

Les données de consommation utilisées pour cette thèse proviennent des Enquêtes sur le niveau de vie des ménages vietnamiens (VHLSS) 2012, 2014 et 2016. Les prix proviennent des circulaires et décisions successives sur le sujet. Les données de température proviennent du National Centers for Environmental Information (NOOA) sur la période correspondante.

La période 2012-2016 a été choisie parce que 2016 est la campagne la plus récente disponible, et l'enquête de de 2010 est trop différente des trois suivantes. L'échantillonnage tournant du VHLSS permet la construction de données de panel avec trois vagues.

La construction du jeu de données a nécessité de faire trois choix.

Le premier choix concerne une lacune dans l'observation du barème de prix auquel ont fait face les ménages répondant à l'enquête. Comme expliqué au Chapitre 2, le Vietnam a deux barèmes de prix. Le premier est le barème de prix de détail appliqué à environ 85% des ménages achetant de l'électricité directement à EVN. La seconde concerne les prix de gros appliqués à certaines zones rurales isolées, à faible densité de population et à l'infrastructure non organisée. Le prix de vente au détail pour le deuxième groupe est inconnu.

La thèse applique le barème de prix de détail EVN à l'ensemble de l'échantillon. Ce choix de la grille tarifaire est justifié par deux faits. i) Le prix de gros est appliqué à une petite fraction de la population (environ 15%) et la fraction diminue. (ii) Bien que les organisations de distribution d'électricité en milieu rural fixent leurs prix de manière hétérogène, elles adoptent le même schéma de tarification uniforme ou progressive par tranche (IBT). Il n'y a pas de tarification dégressive. Ainsi, les prix sont toujours corrélés positivement avec les prix de détail.

Le second choix est celui de la variable mesurant la consommation d'électricité. En effet les questionnaires VHLSS comprennent à la fois une question sur la quantité consommée (en kWh) et le prix total – la facture d'électricité du mois dernier. L'analyse des réponses brutes montre une cohérences limitée entre ces deux réponses. La différence entre le kWh dans les questionnaires et le kWh calculé à partir des factures d'électricité et le bordereau de prix correspondant nous apparaît importante. La thèse choisit d'utiliser les réponses sur la facture d'électricité et d'en déduire la quantité consommée en kWh. Ce choix est justifié en considérant que la consommation en kWh est plus abstraite pour les répondants que la dépense monétaire. On suppose que les répondants se souviennent du montant de la facture du mois mieux que du nombre de kWh consommés.

Le troisième choix est le rapprochement des séries demande/température, pour associer des données de température à chaque ménage. La température associée à chaque ménage est déterminée en deux étapes: (i) Identifier la station météo disponible la plus proche pendant la période de facturation du ménage. (ii) Assigner la température de la station météorologique à la période de facturation en question. La station la plus proche est déterminée en calculant et en comparant la distance entre les coordonnées du centre du quartier où le ménage se trouve et les coordonnées de toutes les stations météo.

Déterminer les coordonnées géographiques du centre des quartiers, nécessite d'utiliser un fichier d'information géographique (shapefile) décrivant les frontières administratives des quartiers du Vietnam. Le fichier officiel le plus récent disponible datait de 2008, et ne tenait pas compte des modifications administratives depuis cette date. Il ne pouvait donc pas être utilisé pour les données de 2012-2016. Heureusement, des fichiers mis à jour et gratuits sont disponibles, la meilleure source étant le Global Administrative Areas (GADM). Le fichier le plus à jour du Vietnam dans GADM est celui d'avril 2018, qui convient pour notre usage.

Cependant un travail supplémentaire a été requis. Alors que le fichier GADM référence les quartiers par leur noms, les données VHLSS ont uniquement des codes de quartiers. Par conséquent, nous avons dû attribuer des codes de quartier aux noms de quartier correspondants dans le fichier de formes. Nous avons accompli cette tâche en utilisant la liste administrative officielle publiée par le General Statistical Office (GSO). Le processus d'attribution, dont le résultat pourra être utile à tous les chercheurs exploitant sur les facettes géographiques des enquêtes VHLSS, est détaillé dans l'annexe 1.

Compte tenu que les ménages ont été interrogés une, deux ou trois fois entre 2012 et 2016, nous avons composé deux sous-ensembles à partir des données : (i) Un panel composé les ménages apparaissant au cours des trois années et (ii) Un groupe composé des ménages qui n'apparaissent qu'une fois toutes les trois années. Les données de ce groupe sont utilisées pour estimer la fonction de demande de court terme, avec l'équation (1). Il contient des ménages qui n'apparaissent qu'une seule fois dans les données, ce qui correspond à l'hypothèse du niveau d'équipement fixes. Le panel est utilisé pour estimer la fonction à long terme, avec l'équation (2).

Chapitre 5. Élasticité-prix de la demande d'électricité résidentielle

Le chapitre 5 porte sur l'impact de l'augmentation du tarif forfaitaire sur la demande résidentielle. Les deux objectifs de ce chapitre sont les suivants: (i) estimer l'élasticité de la demande par rapport aux prix à court et à long terme; et (ii) déterminer si les ménages répondent au prix marginal ou au prix moyen.

Données et méthodes

Ce chapitre estime la fonction à court terme (1) avec l'ensemble de données en pool et la fonction à long terme (2) avec l'ensemble de données de panel pour répondre à la question (i). Pour la question (ii), la méthode consiste à substituer les variables de prix dans les fonctions par une variable de prix perçue. Le prix perçu est $P \quad = MP\left(\frac{AP}{MP}\right)^{k}$ (Shin 1985). La fonction économétrique correspondante est la suivante

$$lnE = \alpha_0 + \alpha_1 lnZ + \alpha_2 lny + \alpha_3 lnMP + \alpha_4 ln\left(\frac{AP}{MP}\right) \qquad (3)$$

Le paramètre $k = \frac{\alpha_4}{\alpha_3}$. La variance de k est calculée par la méthode d'approximation de la variance d'un échantillon large dans Kmenta (1997, p. 486). Si k = 0, les ménages répondent au prix marginal. Si k = 1 alors les ménages répondent au prix moyen.

Le prix final qu'un ménage paiera dépend de la quantité consommée par le ménage. En conséquence, le prix et la quantité sont encore déterminés conjointement, bien que la structure des prix soit connue auparavant. Le biais du panel dynamique provient de l'inclusion du décalage de la variable dépendante dans la fonction à long terme. Dans le panel dynamique, il peut exister une corrélation entre le décalage et les chocs de demande inobservés en termes d'erreur, ce qui aboutit à l'endogénéité.

Pour remédier à l'endogénéité, ce chapitre applique les variables instrumentales développées par McFadden, Puig et Kirschner (1977). L'approche utilise le prix prévu comme instrument du prix réel. Pour aborder le biais du panel dynamique, ce chapitre applique l'estimateur GMM du système de Blundell et Bond (1998).

Résultats

Les résultats estimés montrent que (i) la demande en électricité des ménages est élastique à court terme. Les élasticités marginale et moyenne des prix sont respectivement de -1,53 et -2,03. Cela contredit les études réalisées dans d'autres pays, mais concorde avec les résultats de Phu (2017) au Vietnam, qui utilisent des données et des méthodes différentes. ii) L'élasticité à long terme est estimée positive. Ce résultat contredisant la rationalité économique nous suggère que l'estimation pour le long terme est peu robuste. Cela s'explique en particulier par la faible variation du niveaux des prix sur la période. (iii) Les ménages vietnamiens réagissent pleinement aux prix marginaux plutôt qu'aux prix moyens. Cela confirme les résultats de Nieswiadomy et Molina (1991), qui concluent que les ménages ont tendance à réagir aux prix marginaux lorsque les tarifs globaux sont en hausse et aux prix moyens aux tarifs globaux décroissants.

Chapitre 6. Pauvreté en revenus et en électricité au Vietnam 2012-2016

Le chapitre 6 porte sur la relation non linéaire entre le revenu et la demande en électricité d'une part et ses implications, d'autre part, pour l'identification du seuil de pauvreté en électricité. L'hypothèse est l'existence d'un seuil de revenu auquel la consommation d'électricité commence à augmenter avec l'augmentation du revenu. Le kWh consommé par habitant des ménages à ce seuil de revenu est un seuil de besoin de base, au sens où même si on est très pauvre, on ne réduit pas sa demande en électricité.

Données et méthodes

Le seuil est testé de deux manières avec le jeu de données en groupe. Tout d'abord, la relation entre le revenu du ménage et la consommation d'électricité est visualisée à l'aide de diagrammes de dispersion de densité avec une série de lignes ajustées OLS pour différents quantiles de revenu. La visualisation indiquerait si les seuils existent. Deuxièmement, le seuil est formalisé par l'approche économétrique mise au point par Barnes et al. (2011).

Le Barnes et al. (2011) est réalisé en deux étapes. La première consiste à diviser le revenu par habitant en n quantiles représentés par n-1 variables nominales reposant sur les quantiles les plus pauvres. La seconde consiste à estimer la fonction à court terme (1) en substituant la variable de revenu aux n-1 variables nominales. Si le seuil existe, les premiers coefficients estimés des variables nominales seront nuls, le reste sera supérieur à zéro. La première variable nominale ayant des coefficients estimés supérieurs à zéro est le groupe qui possède le seuil.

Résultats

La visualisation et les résultats estimés prouvent l'existence d'un seuil de pauvreté en électricité qui correspond à d'autres études (Douglas F Barnes, Khandker et Samad, 2011; He et Reiner, 2016). Pour l'ensemble de l'échantillon national, le seuil de pauvreté en électricité d'un ménage de quatre personnes est de 51 kWh par mois. Le seuil correspond approximativement à la première tranche du tarif progressif (IBT) au Vietnam. Cependant, ce seuil est supérieur au seuil du programme de subvention de transferts monétaires (30 kWh). En outre, les résultats estimés suggèrent également qu'il existe une différence significative dans les seuils de pauvreté en électricité entre les zones urbaines et les zones rurales. Les seuils sont d'environ 118 kWh pour les ménages urbains et de 54 kWh pour les ménages ruraux.

Chapitre 7. Économies d'échelle en consommation d'électricité résidentielle

Le chapitre 7 porte sur le rôle de deux facteurs démographiques: la taille et la composition du ménage. Le chapitre répond à deux questions: (i) si le tarif en bloc croissant annule les économies d'échelle réalisées en matière d'utilisation de l'électricité au Vietnam; et (ii) s'il existe une différence de demande en électricité entre un enfant, un adulte et un aîné au Vietnam.

Données et méthodes

Les approches non-paramétriques et paramétriques sont appliquées à l'ensemble de données groupées. L'approche non-paramétrique est adaptée de Deaton et Paxson (1998). L'approche compare la fraction de revenu dépensée en électricité (w) de deux ménages ayant la même composition mais une différence de taille. Si le ménage le plus grand dépense une fraction w plus petite, alors il existe des économies d'échelle. Le w est estimé par régression polynomiale locale. La moyenne pondérée de w est calculée sur une plage de revenus donnée. Étant donné que la même gamme de revenus est choisie pour les ménages ayant la même composition, la moyenne pondérée permet de prendre en compte la condition selon laquelle les ménages ont le même revenu par habitant.

L'approche paramétrique estime une double fonction logarithmique de la facture totale d'électricité des ménages en fonction de la taille du ménage, du nombre d'enfants et du nombre d'aînés. La proportion d'enfants et d'aînés traduit l'impact composition ainsi que la différence de demande entre un adulte, un enfant et un aîné. Pendant ce temps, le coefficient de la taille du ménage (β_1) contrôle les économies d'échelle. Si $\beta_1 < 1$, la facture d'électricité augmente moins du double lorsque la taille du ménage double. Intuitivement, cela signifie des économies d'échelle.

Résultats

Les estimations tirées de deux méthodes non paramétrique et économétrique, ainsi que l'analyse de sensibilité avec différents ensembles de données, corroborent l'idée selon laquelle, à partir de 2012-2016, les ménages vietnamiens auraient réalisé des économies d'échelle en matière de dépenses d'électricité, indépendamment des IBT du prix de l'électricité. Lorsqu'un ménage double de taille en conservant la même composition et les mêmes ressources, les dépenses en électricité des ménages augmentent de 47% seulement. Les résultats de l'analyse non paramétrique révèlent également que: (i) il existe une convergence des dépenses d'électricité par habitant de tous les ménages ayant la même composition lorsque le revenu par habitant augmente; et (ii) les économies d'échelle dans les dépenses en électricité disparaissent après un certain niveau de revenu élevé par habitant pour certains types de ménages.

Chapitre 8. Vagues de chaleur et demande d'électricité résidentielle

Le chapitre 8 est consacré aux effets des vagues de chaleur sur la demande d'électricité des ménages. Ce chapitre montre que les degrés-jours de refroidissement (CDD), la variable couramment utilisé pour représenter la température dans la fonction de demande d'électricité, sont insuffisants pour saisir pleinement l'impact de la température. CDD néglige la distribution extrême de la température qui sont des vagues de chaleur dans les pays tropicaux.

Données et méthodes

On utilise les données groupées, en introduisant une variable dummy supplémentaire représentant l'occurrence d'une vague de chaleur au cours du mois de facturation. La définition d'une vague de chaleur est fondée sur Russo et al. (2015). Ils définissent une vague de chaleur comme au moins trois jours consécutifs avec des températures maximales supérieures à un seuil quotidiens. Le seuil quotidien est défini comme le 90e percentile de la température maximale du 15 jours avant et 15 jours après pour toutes les années disponibles.

Résultats

Bien que les schémas de réponse de la consommation d'électricité aux vagues de chaleur diffèrent entre les trois échantillons national, urbain et rural, le modèle rural offre des preuves à l'appui de l'argument selon lequel la variable degré-jours de refroidissement ne suffit pas à rendre compte de l'impact de la température sur la demande d'électricité. Une variable qui capture les évènements extrêmes de température améliore la prédiction.

Chapitre 9. Conclusions

Implication politique

Quelles sont les implications de nos résultats pour les décideurs vietnamiens en ce qui concerne la tarification de l'électricité, et les politiques de subvention pour les ménages pauvres ?

EVN (2015) a proposé trois options pour les modifications de la grille tarifaire d'électricité, à savoir une grille tarifaire unique, une grille à trois tranches et une grille en six tranches. Les résultats de l'estimation montrent que les ménages réagissent pleinement au prix marginal et que la demande d'électricité réagit fortement à la variation des prix. Cela implique que l'option d'un tarif unique n'est pas appropriée si le gouvernement poursuit l'objectif de stimuler les économies d'électricité. Conserver le régime de tarification progressive semble plus appropriée pour atteindre l'objectif. En outre, la structure actuelle en six tranches ne semble pas poser de problème pour être perçue, puisque les ménages répondent pleinement au prix marginal.

L'estimation des seuils de pauvreté en électricité apporte une référence objective pour caler la politique de protection sociale. En 2011, le gouvernement vietnamien a mis en place un programme de subvention de l'électricité sous forme de transferts monétaires aux ménages pauvres. Le programme a démarré avec 30 000 VND (environ 1,5 USD) par mois donnés chaque trimestre. Depuis 2014, le montant de la subvention correspond à la valeur de 30 kWh par mois, au prix en vigueur. En juillet 2018, le ministère des Finances a présenté une proposition pour annuler ce programme. Cette mesure d'économie a rencontré une levée de boucliers, et en octobre 2018, le ministère a rétropédalé. Le projet de circulaire maintenant le programme de subventions garde le seuil de 30 kWh. Cependant, nos résultats montrent que le seuil de besoins essentiels est plutôt autour de 50 kWh, en zone rurale. Pour poursuivre l'objectif de soutenir l'accès à l'énergie moderne pour les ménages pauvres, le gouvernement pourrait donc envisager d'augmenter le niveau de subvention. En supplément ou en complément, le gouvernement pourrait envisager de fixer à 50 kWh la première tranche et de fixer un prix bas pour le bloc. Une telle tarification pourrait atteindre l'objectif consistant à induire des

économies d'électricité tout en maintenant un soutien pour les ménages pauvres afin d'accéder à l'énergie moderne.

Nos résultats montrent que la structure de la tarification ne pénalise pas les familles nombreuses. Le tarif progressif fait monter le prix moyen pour les familles nombreuses, mais elles bénéficient des économies d'échelles résultant du partage des services électriques comme l'éclairage. Pour modérer l'impact sur les familles très nombreuses, le gouvernement permet sous conditions de diviser la facture pour diviser la progressivité (MOIT, 2014, 2018). À en juger par les estimations, il semble que la politique de réparation atteigne son objectif. Il semble donc qu'augmenter le prix de l'électricité pourrait être compatible avec un objectif d'équité envers les familles nombreuses, qui sont souvent aussi les ménages moins aisées au Vietnam.

Limites et recherches futures

Bien que cette thèse donne des résultats significatifs (sauf pour l'estimation de la demande de long terme), des recherches futures pourraient permettre de dépasser certaines limitations en termes de données et méthode. Nous en proposons trois.

Il y a 173 météo stations à travers le Vietnam. Or les données de température publiquement disponibles ne concernent que 14 stations, avec une quantité importante de données manquantes à la température maximale posant des difficultés pour identifier les vagues de chaleur. Accéder aux données des 173 stations météorologiques, avec des informations plus complètes pour les températures moyennes et maximales quotidiennes, contribuerait à améliorer la précision de l'estimation.

De plus, nous ne connaissons pas les prix exacts d'une partie des ménages dans les données, ce qui peut causer un certain niveau d'erreurs de mesure. Nous avons appliqué la tarification EVN à l'ensemble de l'échantillon, y compris les 15 % de ménages desservis par des redistributeurs, ce qui entraîne une erreur de mesure. Cependant, la proportion de ménages abonnés directement à EVN augmente d'année en année. Le problème pourrait ne plus exister dans les prochains tours de VHLSS. Nous pouvons appliquer les mêmes méthodes dans cette thèse aux données mises à jour pour obtenir une meilleure estimation.

En termes de méthode, les approches actuelles ne permettent pas d'estimer la fonction de demande à long terme. La fonction de demande à long terme est caractérisée par un processus autorégressif décrit à l'équation (2). L'économétrie du panel fait face à des problèmes qui ne nous ont pas permis de conclure. Les méthodes pour résoudre ces problèmes proposées usuellement comme Anderson et Hsiao (1982), Arellano et Bond (1991) ou Blundell et Bond (1998), se basent sur les données en différence. Mais elles ne sont pas utilisables lorsqu'on ne dispose que de trois périodes. De nouvelles méthodes sont nécessaires pour estimer la fonction de demande à long terme.

Finalement, les résultats surprenants de la thèse ne sont pas tous des limites de méthode ou de données. Certains sont des pistes à explorer pour mieux comprendre les comportements des ménages lors de l'utilisation de l'électricité. Ces explorations pourraient mettre en œuvre des méthodes qualitatives et quantitatives diverses au delà de l'économétrie. Nos interrogations sur le comportement des ménages porte sur trois points: (i) pourquoi la demande en électricité d'un adulte, un aîné ou un enfant au Vietnam ne varie pas; (ii) pourquoi les économies d'échelle dans les dépenses en électricité disparaissent avec l'augmentation du revenu par habitant; et (iii) pourquoi les vagues de chaleur ne semblent pas avoir d'impact sur la demande d'électricité dans les zones urbaines.

ÉCOLE DOCTORALE
Agriculture, alimentation, biologie, environnement, santé (ABIES)

Titre : Déterminants de la demande d'électricité des ménages au Vietnam entre 2012 et 2016

Mots clés : *demande d'électricité résidentielle, blocs de tarif progressifs, pauvreté électrique, économie d'échelle, vague de chaleur*

Résumé : La thèse étudie économétriquement les déterminants de la demande d'électricité à l'aide des données des enquêtes nationales auprès des ménages vietnamiens des années 2012, 2014 et 2016. Quatre facteurs explicatifs de la demande sont étudiés : la tarification, le revenu du foyer, sa composition et la température. Les conclusions principales sont que: (i) Les ménages réagissent aux prix marginaux, la demande est élastique par rapport au prix. (ii) Il existe un seuil de revenu à partir duquel la consommation d'électricité des foyers augmente quand le revenu augmente : la consommation d'électricité des foyers ayant ce revenu peut être considérée comme le niveau de besoin fondamental, un seuil de pauvreté pour l'électricité. (iii) La progressivité de la tarification ne pénalise pas les familles nombreuses : le tarif progresse moins vite que les d'économies d'échelle des dépenses d'électricité. (iv) Nous n'observons pas d'effet de la composition du foyer en termes enfants / adultes / personnes âgés sur la dépense d'électricité. (v) Les vagues de chaleur - un phénomène lié au changement climatique - impactent la demande d'électricité et devraient être mieux prises en compte dans l'estimation de la demande.

Title : Exploring the determinants of household electricity demand in Vietnam in the period 2012–16

Keywords : *residential electricity demand, increasing block tariffs, electricity poverty, economies of scale, heatwaves.*

Abstract : The thesis employs a micro survey in the form of pooled and panel data sets to estimate electricity demand in Vietnam in the period 2012–16, focusing on four main factors: increasing block tariffs, income, demographics, and heatwaves. The estimate results lead to five main conclusions. (i) Households do respond to marginal prices and demand is elastic to price. (ii) There exists an income threshold from which household electricity consumption increases as income increases. The electricity consumption of households in the income group is the reference level of electricity poverty threshold. (iii) The increasing block tariff does not cancel out economies of scale in electricity expenditure of households. (iv) There is no difference in electricity expenditure across children, adults and elders. (v) Heatwaves – a climate change-related phenomenon – do have impacts on electricity demand and need to be addressed carefully in estimating electricity demand in the future.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France