เครื่องมือสำหรับการวัดและการควบคุมเงินอุดหนุนพลังงานในประเทศรายได้ปานกลาง

นายแดเนียล เรย์ ลูวิส

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเศรษฐศาสตรดุษฎีบัณฑิต
สาขาวิชาเศรษฐศาสตร์
คณะเศรษฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559

Tools for Measuring and Controlling Energy Subsidies in Middle Income Countries

Mr. Daniel Ray Lewis

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Economics
Faculty of Economics
Chulalongkorn University
Academic Year 2016

Thesis Title	Tools for Measuring and Controlling Energy Subsidies in Middle Income Countries
By	Mr. Daniel Ray Lewis
Field of Study	Economics
Thesis Advisor	Associate Professor Pongsa Pornchaiwiseskul, Ph.D.

Accepted by the Faculty of Economics, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

....Dean of the Faculty of Economics
(Professor Worawet Suwanrada, Ph.D.)

THESIS COMMITTEE

....Chairman
(San Sampattavanija, Ph.D.)

....Thesis Advisor
(Associate Professor Pongsa Pornchaiwiseskul, Ph.D.)

....Examiner
(Associate Professor Tanapong Potipiti, Ph.D.)

....Examiner
(Assistant Professor Yong Yoon, Ph.D.)

....External Examiner
(Associate Professor Chaiyuth Punyasavatsut, Ph.D.)

แดเนียล เรย์ ลูวิส : เครื่องมือสำหรับการวัดและการควบคุมเงินอุดหนุนพลังงานในประเทศรายได้ปานกลาง (Tools for Measuring and Controlling Energy Subsidies in Middle Income Countries) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ.พงศา พรชัยวิเศษกุลดร, 165 หน้า.

วัตถุประสงค์ของวิทยานิพนธ์ฉบับนี้คือการพัฒนาเครื่องมือทางเศรษฐมิติที่เป็นประโยชน์ในการออกแบบการให้เงินอุดหนุนด้านพลังงาน การศึกษานี้แบ่งออกเป็นสามบท บทที่หนึ่งเกี่ยวกับความต้องการพลังงานและระดับเงินอุดหนุนที่เหมาะสม บทที่สองเกี่ยวกับการประเมินเงินอุดหนุนด้านพลังงานเมื่อเวลาผ่านไป และ บทที่สามเกี่ยวกับการกำหนดเป้าหมายและการกระจายเงินอุดหนุนด้านพลังงาน ถึงแม้ว่าเป้าหมายหลักของการให้เงินอุดหนุนในงานนี้คือครัวเรือนที่มีรายได้น้อย แต่ก็ยังได้มีการกล่าวถึงการอุดหนุนกลุ่มเป้าหมายโดยแบ่งตามภูมิภาคหรืออาชีพด้วยเช่นกันเครื่องมือที่กล่าวถึงในวิทยานิพนธ์ ได้แก่ เครื่องมือที่ 1: Elasticities with Quantile Regressions - ใช้เพื่อทำความเข้าใจความต้องการพลังงาน ของประชาชนที่อยู่ชายขอบของการกระจายรายได้ เครื่องมือที่ 2: Spline Regressions – เพื่อแสดงระดับการใช้พลังงานที่เพิ่มขึ้นมื่อรายได้เพิ่มขึ้น นอกจากนี้การศึกษายังตอบปัญหาว่า เมื่อไหร่ควรให้เงินอุดหนุนแก่ครอบครัว เมื่อไหร่ควรให้เงินอุดหนุนแก่รายหัว เครื่องมือที่ 3: Engel Curves – เพื่อหาปริมาณพลังงานที่ควรได้รับเงินอุดหนุน เครื่องมือที่ 4: เส้นความยากจนในการใช้พลังงาน – เพื่อช่วยในการพิจารณาว่าใครควรจะมีสิทธิ์ได้รับเงินอุดหนุนด้านพลังงาน เครื่องมือที่ 5: ชุดข้อมูลค่าใช้จ่ายด้านพลังงาน – เพื่อช่วยในการกำหนดว่าใครควรจะได้รับผลประโยชน์จากเงินอุดหนุนและกำหนดเงินอุดหนุนที่ควรสนับสนุน เครื่องมือที่ 6: งบประมาณและ Pie Charts – เพื่อช่วยในการกำหนดต้นทุนและผลประโยชน์ของการให้เงินอุดหนุนและเพื่อทำความเข้าใจผู้ได้รับผลประโยชน์ เครื่องมือที่ 7: ลูปข้อมูลขนาดใหญ่ – ใช้เพื่อทำความเข้าใจที่ดีขึ้นเกี่ยวกับ ลักษณะของผู้ใช้พลังงานและผู้ได้รับเงินอุดหนุนในปัจจุบัน เครื่องมือที่ 8: การแสดงข้อมูลทางภูมิศาสตร์ - ใช้เพื่อวิเคราะห์รูปแบบทางภูมิศาสตร์ของการใช้เงินอุดหนุนในประเทศ นอกจากนี้ยังใช้เพื่อสนับสนุนว่าทำไมเงินอุดหนุนอาจมีความสำคัญต่อประเทศรายได้ปานกลางต่อจากนั้น

ข้อเสนอแนะเฉพาะสำหรับการให้เงินอุดหนุนของประเทศไทยได้แก่ระดับการใช้ไฟฟ้าฟรี 80-140 KwH ต่อครัวเรือนต่อเดือน และเงินอุดหนุนสำหรับก๊าซหุงต้ม 3-5 กิโลกรัมต่อครัวเรือนต่อเดือน หรือเงินอุดหนุน 20 ลิตรเชื้อเพลิงการขนส่งต่อครัวเรือนต่อเดือนเป็นต้น

สาขาวิชา เศรษฐศาสตร์ ลายมือชื่อนิสิต ....

ปีการศึกษา 2559 ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

5485901229 : MAJOR ECONOMICS
KEYWORDS: ENERGY SUBSIDY POLICY ECONOMETRICS THAILAND

DANIEL RAY LEWIS: Tools for Measuring and Controlling Energy Subsidies in Middle Income Countries. ADVISOR: ASSOC. PROF. PONGSA PORNCHAIWISESKUL, Ph.D., 165 pp.

The purpose of this dissertation is the development of econometric tools useful in the design of energy subsidies. The tools are divided between three chapters, a chapter on Energy Demand and the Optimal Level of Subsidies, a chapter on Evaluating Energy Subsidies over Time and a chapter on Targeting and Distribution of Energy Subsidies. Although the primary target for subsidies in this paper are low income households, subsidies designed to target groups by region or by profession are also discussed. The Tools discussed in the Dissertation include: Tool 1: Elasticities with Quantile Regressions - *Used for understanding energy demand at the edge of the income distribution* Tool 2: Spline Regressions - *To show the marginal degree to which energy use increases with income and to study when household versus per capita subsidies are preferable.* Tool 3: Engel Curves - *To determine the quantity of energy that should be subsidized.* Tool 4: Energy Poverty Lines - *To help determine who should be eligible for energy subsidies* Tool 5: Energy Expenditure Time Series - *To help to determine who receives benefits from subsidies and to determine which subsidies are worth supporting.* Tool 6: Budget and Pie Charts - *To help to determine the cost-benefit of providing a subsidy and to better understand who benefits.* Tool 7: Big Data Loops - *Used to understand the characteristics of current energy and subsidy users of energy.* Tool 8: Geographical Data Visualization - *Used to analyze geographical patterns of subsidy use in the country.* In addition a justification of why subsidies may continue to be relevant for Middle Income Countries is provided. Four different approaches to providing subsidies are discussed with pros and cons of each given. These are Subsidies that are 1) Targeted , 2) Self-Selecting, 3) Give the Same Quantity to All, and 4) Cash-Transfers. In juxtaposition to much of the current literature, the author makes a case for the first three options. While cash transfers are economically efficient, they are a successful political solution and they require a high degree of information and control that middle income countries not want or be able to provide.

While most of this dissertation is generally focused on tools for targeted subsidies, self-selecting subsidies can also be an elegantly simple solution that does not categorize people, and in terms of fairness, giving the same to all is likely the most just solution. Specific suggestions for Thailand subsidies are proposed, including an 80-140 KwH level of free electricity per household and IF NEEDED, a 3-5 kg level of subsidy for LPG or a 20 liter subsidy for transportation fuel per household.

Field of Study: Economics	Student's Signature ....
Academic Year: 2016	Advisor's Signature ....

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Pongsa Pornchaiwiseskul, for his patience and consideration throughout the process of the Ph.D. and dissertation. He has been very generous in allowing an old student to follow his own lead as far as possible.

I would like to thank my friends, and the members of my committee, for listening to my ideas, for long discussions on data and the future of economics, for working together on projects, for giving suggestions, and basically for being great friends.

I would like to thank my students, both for listening to my ideas in class, and making suggestions, and for my senior research students, for helping to explore new techniques and ideas about energy policy.

I would like to thank the EBA program to allow me time to devote to my Ph.D. and for employing me steadily for the past 15 years.

I would like to thank the Faculty of Economics, Chulalongkorn University for being a home for much of the working part of my life. There are many great people and projects here.

I would like to thank my wife, Sukanda, for being a companion and for listening and sharing hopes and dreams during the process of working on this Ph.D. as well as before and after.

Sincerely,

Daniel Ray Lewis

Faculty of Economics, Chulalongkorn University

CONTENTS

Page

Page

Page

Page

Chapter 1 – Introduction and Motivation

The purpose of this dissertation is the development of econometric tools useful in the design of energy subsidies. The tools are divided between three chapters, a chapter on Energy Demand and the Optimal Level of Subsidies, a chapter on Evaluating Energy Subsidies over Time and a chapter on Targeting and Distribution of Energy Subsidies. Although the primary target for subsidies in this paper are low income households, subsidies designed to target groups by region or by profession are also discussed especially in the Targeting chapter.

1.1 Overview of the Econometric Tools used in the Dissertation

(Included in Chapter 3 (Paper 1) Energy Demand and the Optimal Level of Subsidies)

Tool 1: Elasticities with Quantile Regressions

Used for understanding energy demand at the edge of the income distribution

Elasticities are a common tool for understanding demand for energy products. Although elasticities can be calculated using OLS for different income groups, quantile regressions are better for capturing the marginal behavior of different income households at the edges of the distribution. The paper first presents the basic OLS results for elasticities of each energy type. Quantile regressions are then used to calculate elasticities all along the distribution. This is useful for subsidy work as we are often interested in persons at the bottom edge of the income distribution.

Figure 1 Benzene91 OLS Regression Elasticities

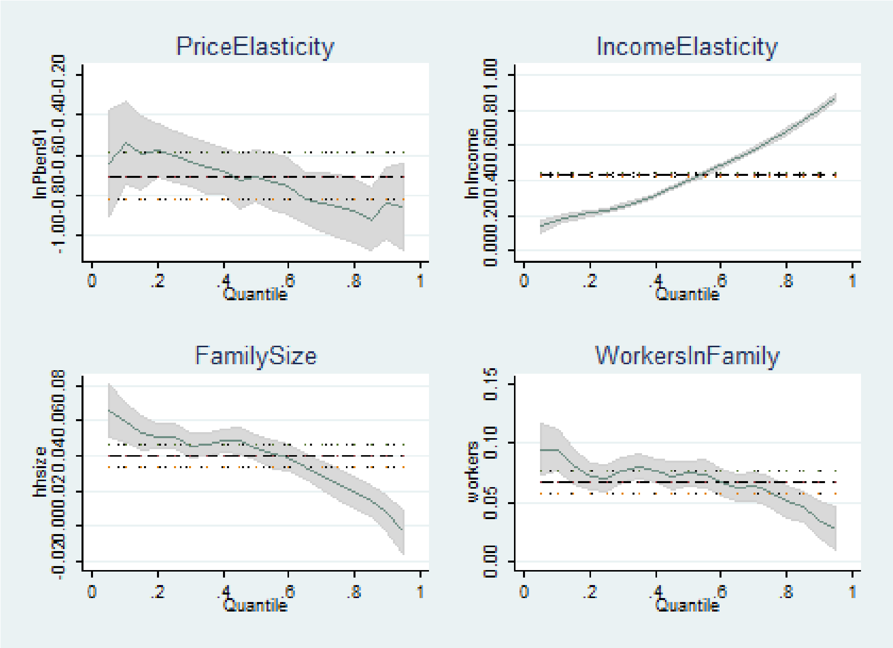
Source	SS	df	MS		
				Number of obs =	28948
				F(5, 28942) =	1501.72
Model	3284.44649	5	656.889299	Prob > F =	0.0000
Residual	12659.9185	28942	.437423763	R-squared =	0.2060
				Adj R-squared =	0.2059
Total	15944.365	28947	.550812348	Root MSE =	.66138

lnQben91	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnPben91	-.7846325	.0588774	-13.33	0.000	-.9000348	-.6692301
lnIncome	.4379173	.006697	65.39	0.000	.4247909	.4510438
hhsize	.0415487	.003224	12.89	0.000	.0352294	.0478679
workers	.0574572	.0048833	11.77	0.000	.0478857	.0670287
rural	-.0339226	.0089322	-3.80	0.000	-.0514301	-.016415
_cons	1.189538	.2292817	5.19	0.000	.7401351	1.63894

SOURCE SES 2012

The OLS regression of log quantity on log price and log income above shows that the price elasticity of demand for benzene 91 is -0.78 (slightly elastic) while the income elasticity is 0.44 (necessity). These numbers are useful summary statistics, but they simply and may falsely represent the entire distribution. Quantile regression allows us to look at the edges of our distribution and to see how elasticities change over a range of energy use, all in an easy to understand graphical representation. In the figures below, the dark dashed line shows the OLS estimate of elasticities, while the lighter dashed lines show the 95% probability band for OLS. The solid wandering line shows the estimate of the elasticity at different places along the distribution with quantile .1 representing decile 1, quantile .2 representing decile 2 etc., and the shaded area represents the 95% probability band for each place in the distribution. For the lowest decile, both price elasticity and income elasticity seem to be more inelastic for this low usage group, while family size seems to matter more as does the number of workers.

Figure 2 Benzene91 Quantile Regression Elasticities



SOURCE SES 2012

Tool 2: Spline Regressions
To show the marginal degree to which energy use increases with income and to study when household versus per capita subsidies are preferable.

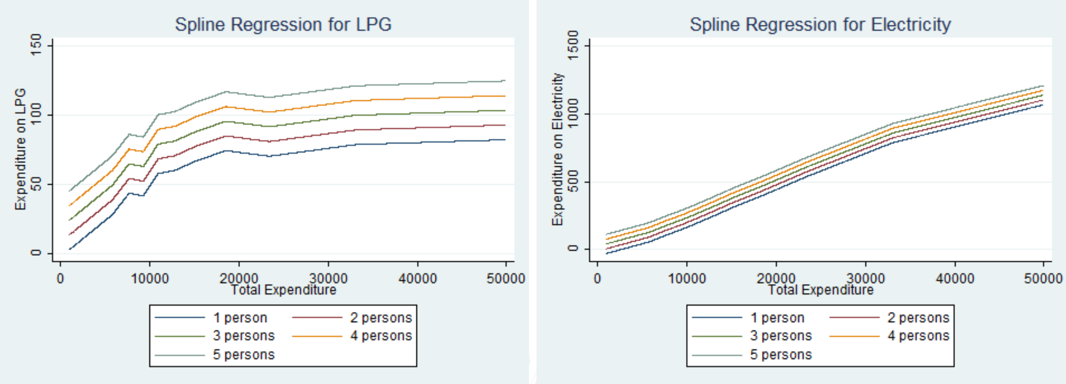
Should subsidies be at the household level or at the per capita level?
A light in the middle of a room can provide light to all, but every person needs a separate bowl of rice.

Spline regressions are introduced to help determine the appropriate unit for subsidy regressions. Non-parametric spline regressions, although often used for forecasting, also provide two other functions in this paper.
Spline regressions employ a true marginal concept designed to look at how an addition of one variable will affect the use of another variable without recourse to the previous part of the regression. This is not true for elasticities. Elasticities are not marginal as they rely on dividing by the entire energy use.
By dividing the regression into parts separated by "knots" a picture of how energy use changes with income can be modeled. Each segment shows the relationship between income and the energy good under consideration.

In addition, spline regressions are an easy way to model whether subsidies should be at the household or per capita level, as can be shown in the spread of the lines below.

Per person level preferred ***Per household level preferred***

Figure 3 Spline Regressions



SOURCE: SES 2011

For LPG, used for cooking, usage goes up quickly at low levels of income, but upon reaching a certain level, higher income does not result in much increase in usage. Electricity, on the other hand, shows increasing usage throughout the range of incomes. LPG usage depends on the number of persons in the household as shown by the dispersion of regression lines, while for electricity this does not seem to be very true.

These tools give an indicator of how changes in price will affect energy use for different income levels. What if we want to have a specific idea of what is the appropriate level of energy to subsidize?

Tool 3: Engel Curves

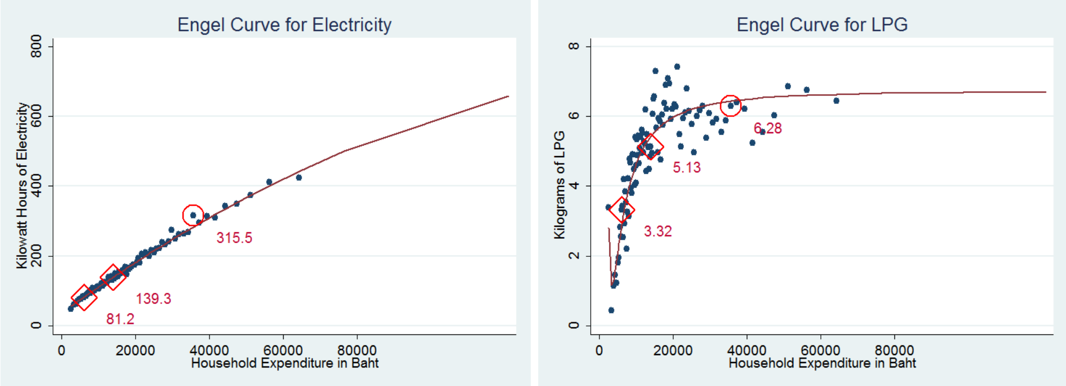
To determine the quantity of energy that should be subsidized.

It is often difficult to determine the level of energy to subsidize. This paper takes the approach of adopting both a philosophy and a tool. The philosophy postulates that there will always be poor people, at least relative to the rest of society, and it is our objective to always help the bottom 10% to achieve the standards and living conditions of the rest of society. Thus we never solve the problem of poverty, but we always reach down to the ones on the bottom to pull them up, so that as a society we can all grow together.

In practice, this means we want to assist the bottom 10% of households in terms of income to achieve energy use similar to those immediately above them in terms of income.

The tool we use for this purpose is based on the familiar Engel curve, which plots income versus energy use. In our version, the sample is divided into one hundred income group centiles, with each centile plotted on a graph as below.

Figure 4 Engel Curves for Electricity and LPG



NOTE: OPEN DIAMONDS ◊ DEMARCATE THE 10% AND 50% INCOME LEVEL, WHILE AN OPEN CIRCLE ○ DEMARCATES THE 90% INCOME LEVEL. ONE APPROACH TO SUBSIDIES IS TO HELP THOSE BELOW THE 10% INCOME LEVEL TO USE ENERGY AT A TARGET LEVEL BETWEEN THE 10% AND 50% INCOME LEVELS, OR BETWEEN THE OPEN DIAMONDS. FOR INSTANCE, IN THE ELECTRICITY ENGEL CURVE ABOVE, WE WOULD TRY TO ENSURE THAT EVEN THE POOREST HOUSEHOLDS COULD USE BETWEEN 81.2 TO 139.3 KWH OF ELECTRICITY.

SOURCE: SES 2013 [1]

This Engel curve serves several purposes. First, as is standard with Engel curves, it shows how demand for the product changes as household income increases.
In the Engel curves above, three points are marked. The lowest diamond shows the 10^{th} centile in household income. Points to the left are poorer and use less energy. This is the group we would hope to subsidize. The second diamond represents the median income. The range of energy use between the 10% and 50% we consider to be normal energy use, or target use for our country. The final point is the 90% centile above which additional energy use should perhaps be taxed to support the poorer group. Again, we can find an actual quantity to give us an indicator range to tax.

Included in Chapter 4 (Paper 2) Evaluating Energy Subsidies over Time

Tool 4: Energy Poverty Lines

To help determine who should be eligible for energy subsidies

Another related problem is determining the income level at which subsidies should be made available and setting eligibility requirements. Energy poverty lines are a tool that can be used to determine eligibility for energy subsidies. There are several approaches that are outlined in the paper.
Poverty lines were originally envisioned to represent a fixed number of calories and nutrients to sustain life, and an energy equivalent exists, supplying an amount of energy to ensure light, heat, cooking and transportation.

---

[1] In textbooks Engel curves are often plotted with Income on the Y axis and Quantity on the X axis. The axes are switched in empirical economics because of potential difficulties in plotting backward bending portions of the curve associated with inferior goods.

Figure 5 Minimum requirements for energy

Lighting	*300 lumens at household level*
Cooking	*1 kg wood, or 0.3 kg charcoal, or 0.4 kg LPG or 0.2 liters ethanol per person* *40% more efficient stove than 3 rocks* *Particulate matter less than 10 $\mu g/m^3$*
Heating &Cooling	*Minimum indoor temperature of 12^0C* *Maximum indoor temperature of 30^0C*
Information	*Receive/send information to the outside world*
Livelihood	*Access to energy needed for livelihood*

Source: Practical Action 2010

However, as starvation decreased due to the green revolution, the focus of poverty lines changed to measure other statistics, such as a fixed share of income used for food (or energy), or an attempt to identify standards below which persons should not fall, or an attempt to assist the poorest share of the population. Energy poverty lines had a brief period of popularity when energy prices peaked in 2007-2011 and likely will be adopted again if humans lose the race between new energy sources and the inevitable decline in current sources. The options for a poverty line are discussed in depth in the paper, but eventually this paper will focus on a poverty line focused on helping the poorest decile of households.

Tool 5: Energy Expenditure Time Series

To help to determine who receives benefits from subsidies and to determine which subsidies are worth supporting.

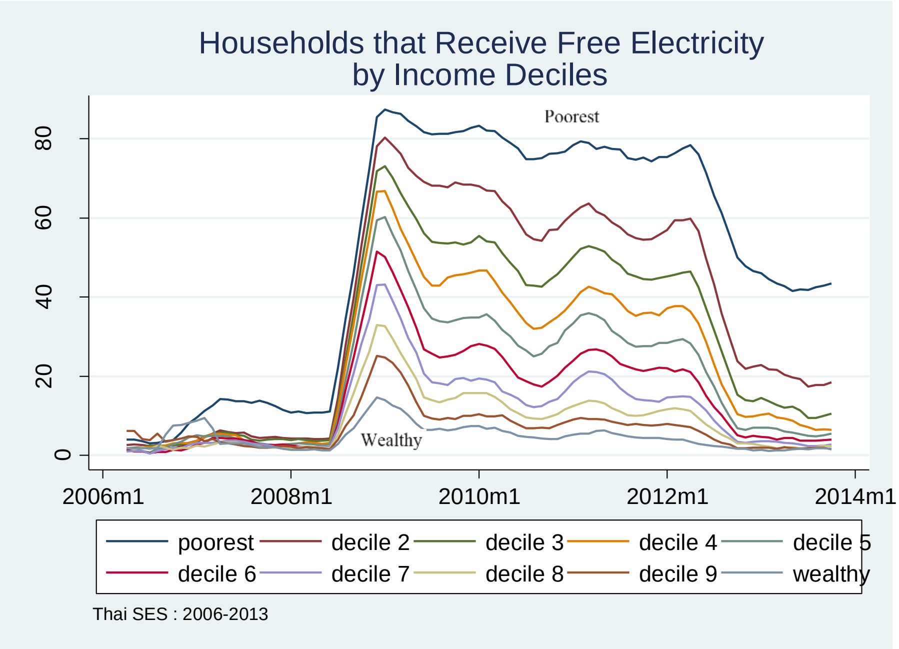
Subsidies should benefit those who need them most, and not waste money subsidizing those who have sufficient resources already. Many subsidy programs suffer because as wealthy people spend more on all goods, they also spend more on subsidized goods. Therefore, a larger share of subsidies go to the wealthy than go to the poor. Designing programs that reach the poor group while excluding wealthier groups is a challenge that can be aided by a clear understanding of who is making use of the subsidies.

The expenditure time series used in this paper take advantage of the fact that annual social economic surveys usually also record the month the survey was conducted allowing the reconstruction of a monthly time series. By dividing the time series into different income deciles, then dividing by month we can reconstruct the benefit paths of the subsidy program for different deciles over time. It is possible to model either coverage or the amount used by each group over time.

Thai socio-economic survey data is collected year round with the survey month coded into the data. Households are divided by the income decile to which they belonged resulting in 10 parallel time series of expenditure data. To give a more consistent time series, averages were smoothed using a 6 month smoothing framework.

The Free electricity program (lifeline levels) was initiated in 2008. The purpose was to help households that used exceptionally low levels of electricity under the assumption that they were the poorest. Compared to other aid programs this one was surprisingly effective. Using 330,000 households from 8 years of the SES expenditure survey, we can see the pattern of benefits from the program using a coverage chart.

Figure 6 Timeline of Share of Households Using Free Electricity



SOURCE: SES (2006-2013) NOTE: IN THIS CHART, EACH INCOME DECILE HAS ITS OWN LINE SHOWING THE SHARE OF THAT DECILE THAT BENEFITS FROM THE PROGRAM. THE HORIZONTAL AXIS IS TIME IN MONTHS. IN 2006, THE PROGRAM DIDN'T EXIST, SO NO ONE BENEFITED. IN MID 2008 THE POLICY WAS INTRODUCED , WITH FREE ELECTRICITY FOR THOSE WHO USED BELOW 90 KWH. IN 2013, THE CUTOFF WAS REDUCED TO 50 KWH.

A small trial program was floated in 2007, and the full program began in the middle of 2008. This graph shows what share of each decile received free electricity, with the line on the top being the poorest decile. Eighty percent of the poorest group took advantage of the program. About 60% of the second poorest group took advantage of the program., 50% of the 3rd poorest group, etc. The problem comes when we get down to the 6th poorest group, or households that are slightly wealthier than the average. In this group, about 20% of households took advantage of the program.
This kind of data about who uses what is rare and hard to obtain in a country in which 50% of citizens do not receive a formal paycheck and 80% do not pay income tax.
It is possible to create a composite coverage index which shows the median coverage of the program over time. If this index were 3, for instance, as many people above the 3rd decile receive coverage as below the 3rd decile. The objective is to design policies with the lowest coverage index as possible.

Tool 6: Budget and Share Pies

To help to determine the cost-benefit of providing a subsidy and to better understand who benefits.

Another tool that is useful to evaluate the value and effectiveness of government policies is to use SES survey data to estimate the share of final benefits of a subsidy that go to each income group.

Although administrative costs are necessarily excluded, the share of final benefits that go to each income group can be estimated both in terms of coverage and in terms of the amount received for each type of energy. This data makes use of survey data, appropriately weighted by the number of households, and is most easily presented in a pie or bar chart format.

In the charts below it can be seen that free electricity is a subsidy program that reaches its target audience quite well, while LPG mostly misses its mark.

Figure 7 Electricity (in-kind) Budget Share and Use Share per Quartile

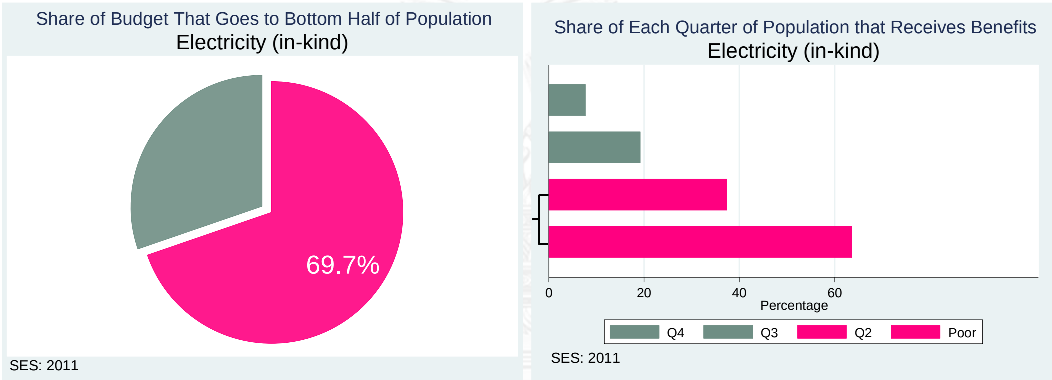
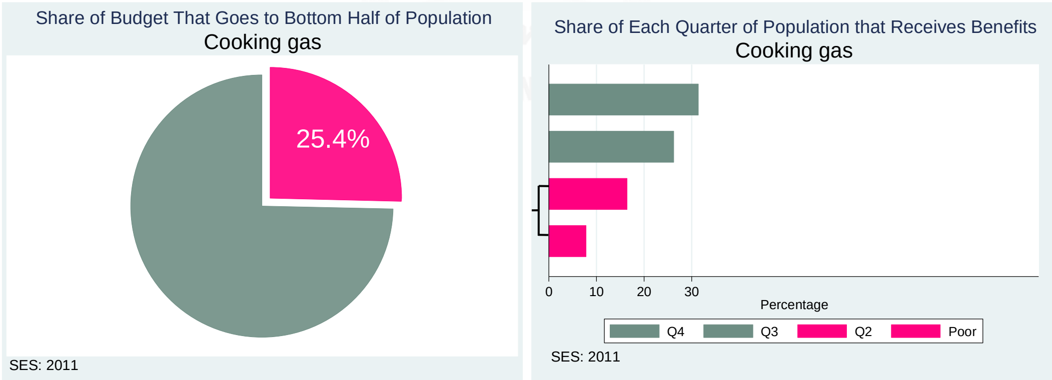


Figure 8 LPG Budget Share and Use Share per Quartile



The pie chart on the left shows the total amount received in subsidies for electricity and for LPG by each half of the population while the bar chart on the right addresses the issue of coverage and is in quartiles. More than 60% of the poorest income quartile took advantage of the free electricity program, while less than 10% of the poorest decile took advantage of LPG cooking gas. The cooking gas subsidy is poorly targeted.

There is usually a tradeoff between more complete coverage and the share of the subsidy that goes to the target group. If conditions to enter the program are stringent, only some of the target group can make use of the subsidy. If conditions to enter are lax then most of the target group can enter, but we also have many of the non-target group that enter. We can design an optimal level of stringency by designing a penalty function for false positives, while rewarding true positives. This will be the work of a future paper.

Chapter 5 (Paper 3) The Targeting and Distribution of Energy Subsidies

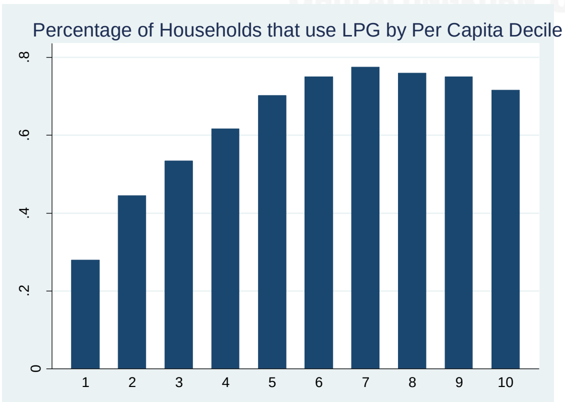
Tool 7: Big Data Loops

Used to understand the characteristics of current energy and subsidy users of energy. Big data techniques offer many opportunities to better understand the government's clients, the subsidy beneficiaries. Just as big business can understand customers and design advertising or benefit campaigns tailored to them, so the government can appropriately target subsidies by understanding their customers better. Although data about income and behavior is often limited in middle income countries, it is possible to design programs around "types" of users based on characteristic from the SES, without ever knowing the actual identity of the user. The technique discussed here is to use loops that accrete R^2 (or correlation which is the same thing) to build understanding and predictive power in identifying a subsidy user. This paper illustrates an example of a household that receives an LPG subsidy.

Loops

Loops are first used to identify what factors are most significantly related to LPG uses in the SES survey. A loop regresses each of the 500 variables in the SES separately against LPG use, excluding only variables with small sample sizes. Income is positively related to LPG use, although not in a linear way, as shown below. In the below table we can see that LPG use increases as income deciles increase, until income decile 7, while thereafter the percentage of households using LPG decrease for the highest income deciles.

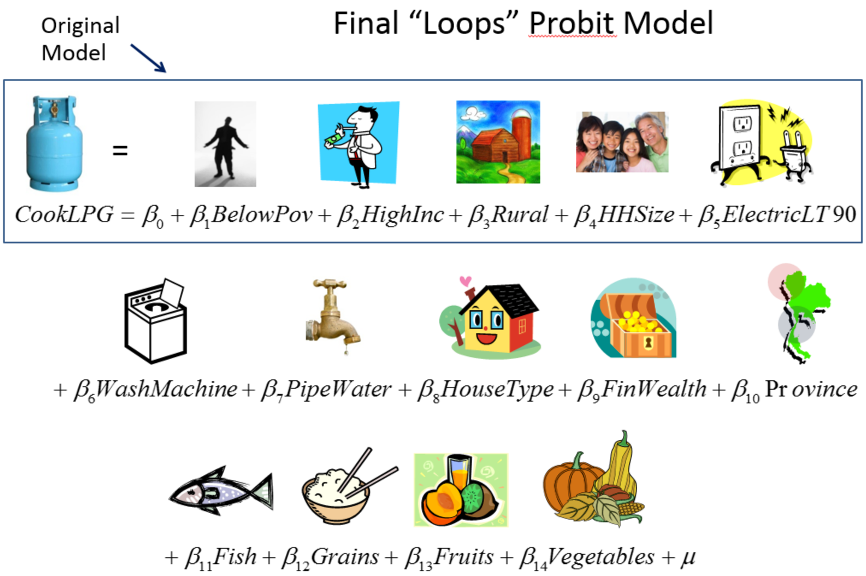
Figure 9 Percent Usage of LPG for Cooking by Income Decile



SOURCE SES 2009

Thereafter a measure of income is added to the regression and the model becomes Y = income + X . Regressions are again run looking for significant increases in predictions of LPG use. The results are added to the original equation. The process continues building in a forward stepwise fashion. Categorical variables are searched for separately. A loop using deciles is added to look for non-linear behavior. In the end the best predictive model uses the following factors. This model can identify ways to deliver subsidies and tie subsidies better to appropriate users.

Figure 10 Example of Predictive Model Generated with Loops



Other big data tools used in predictive analysis such as nearest neighbor techniques and categorical trees are discussed briefly, although a complete analysis is beyond the scope of this paper.

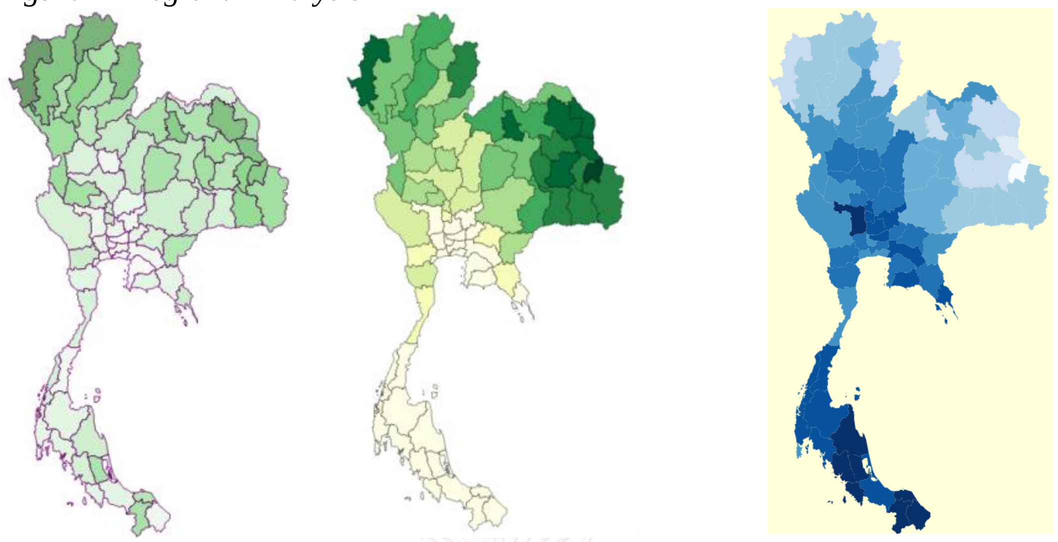
Tool 8: Geographical Data Visualization

Used to analyze geographical patterns of subsidy use in the country.

The last tool described here is to use data visualization to point out patterns in the data that OLS and other techniques tend to obscure. The following maps show that free electricity is available most readily in mountainous environs, presumably since refrigeration and fans require less energy there. Inevitably focusing on mountainous regions was not part of the objective of the program to subsidize poor households. Mountainous regions are also mostly forested, so households use charcoal and firewood for cooking. Supplying subsidized LPG to those who also receive free electricity is likely to fail as mountain people are not in need of LPG. [2]

[2] Actually LPG is used most consistently in the South, so LPG subsidies might be part of a platform for the Democrat party.

Figure 11 Regional Analysis



Free Electricity Use Firewood or Charcoal Use LPG for Cooking

SES: 2013

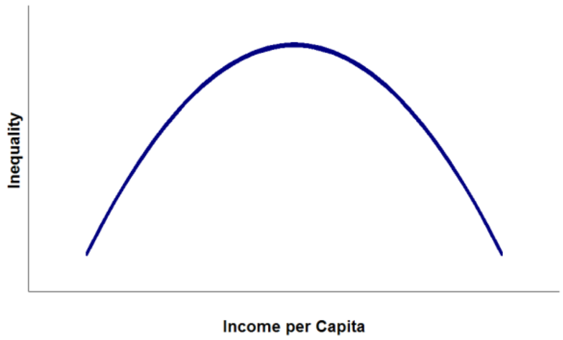
These eight tools will be developed in much more greater depth in the papers of the dissertation, but as they make up the main body of work they are introduced here.

1.2 Development Level and Energy Policy

1.2.1 Middle Income Countries

Middle income countries are different from wealthy countries and different from poor countries. In a middle income country, income inequality tends to be high, with a large aspiring middle class that wishes to rise to the level of wealthy countries. Meanwhile, a poorer traditional sector may be held back as income is tied to the value of unskilled labor and the production of commodities. Middle Income countries are at the center of the Kuznets curve where inequality is at its highest

Figure 12 Hypothetical Kuznets



KUZNETS (1955)

Income inequality increases the need and political demand for redistributive policies, and many middle income countries are led by populist governments. Populist countries, following the interests and demands of their constituents, promote subsidies.
Subsidies are unpopular in both the economics and development world at this juncture. (They are perceived as the genie that has been let out of the bottle and cannot be coaxed back in.) However, it is the contention of this paper these subsidy programs are valuable to the poor, and can be successfully administered, especially with the big data tools now available to us, and that the behavioral economic understanding that allows us to target subsidies to needy groups will allow them to be used efficiently in the future.

1.2.2 Eight Energy Policy Objectives

Middle income countries need to stress different aspects of Energy Policy than either wealthy countries or poor countries. Popular objectives of energy policy include the following:

Table 1 Energy Policy Objectives for Country Groups

	Poor	Middle Income	Wealthy	Exporter	Importer
Basic Needs	X				
Environment			X		
Energy Integration					X
Energy Security					X
Conservation			X		X
Energy Investment	X		X	X	
Alternative Energy			X		X
Energy Subsidies		X			

SOURCE: LEWIS (2014)

Although every category can be of importance to every country, there is a necessary urgency about certain categories depending on income level and whether the country is an importer or an exporter that supersede even the culture and values that often determine policy.

1.2.2.1 Basic Needs

Basic Needs is ensuring that every citizen has equal access to modern energy such as electricity for light, usually gas for cooking, and gasoline for transport.

Table 2 Proposed Minimum Energy Requirements per Household
The non-profit organization Practical Action proposes the following guidelines as minimum standards for access to energy at the household level. These goals could be achievable by the year 2030 and should be adopted as an addition to the UN millennium goals.

Lighting	*300 lumens at household level*
Cooking	*1 kg wood, or 0.3 kg charcoal, or 0.4 kg LPG or 0.2 liters ethanol per person* *40% more efficient stove than 3 rocks* *Particulate matter less than 10 $\mu g/m^3$*
Heating &Cooling	*Minimum indoor temperature of 12^0C* *Maximum indoor temperature of 30^0C*
Information	*Receive/send information to the outside world*
Livelihood	*Access to energy needed for livelihood*

SOURCE: PRACTICAL ACTION (2010)

1.2.2.2 Environmental Concerns

Environmental Concerns are a result of industrialization so might exist in both middle income and developed countries. Mitigation is expensive and so is more likely to be undertaken in wealthy countries. It is likely to be a second order concern in middle income countries due to even stronger interest in growth and development. ***Thailand is a middle income energy importer, and so the issues that are most salient here are energy subsidies, alternative energy, energy integration, energy security and it should be conservation.***

1.2.2.3 Energy Integration

Energy Integration refers to sharing infrastructure between countries so as 1) to facilitate sales between exporters and importers, 2) to increase efficiency by using energy where and when it is needed, and 3) as a risk reduction strategy in case of catastrophic failure of the power supply in one country. In ASEAN, energy integration would benefit those in the center of the region most as transportation of energy is expensive and often involves large "wheeling" losses in power as is the case with electricity or pipelines. Importers would benefit more at present as world energy is currently in short supply.

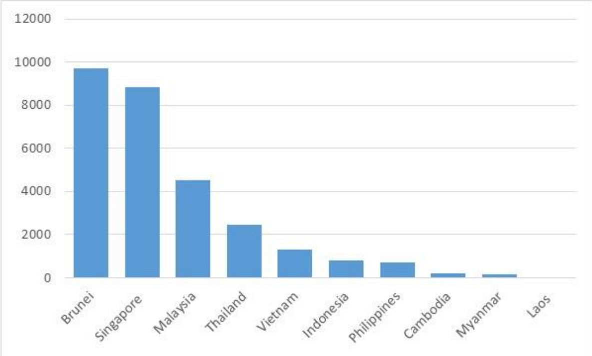
1.2.2.4 Energy Security

Energy Security refers to ensuring sufficient supplies of energy by varying sources of supply, types of energy used as for instance in the generation of electricity, building redundant infrastructure to protect against failures, and storing sufficient supplies of critical fuels to protect against any interruption in supply chains.

1.2.2.5 Conservation

Conservation is most important in wealthy countries because of their very high usage rates. Poor countries are generally trying to increase energy usage to improve quality of life, while wealthy countries tend to use energy in wasteful ways. Middle income countries may have a mix of wealthy and poor income groups requiring conservation by the wealthy groups, and increases to meet basic needs by the poor.

Figure 13 Kilowatt Hours Electricity Use in ASEAN per Capita per Year



WORLD DEVELOPMENT INDICATORS (2010B)

1.2.2.6 Energy Investment

Energy Investment is often cross-border with wealthy countries and energy exporters investing in poor countries which lack sufficient capital and technology to invest themselves. Middle income countries tend to be cautious about foreign investment, as the terms are generally more favorable to the investor in terms of both profits and rights of use of the final product. Middle income countries would prefer to self finance investment if possible.

1.2.2.7 Alternative Energy

Alternative Energy may serve as security for the future when fossil fuels are in short supply, as security for the present as a domestically produced energy source, or as a potential saver of foreign exchange. Most alternative energy is still more expensive than conventional energy sources, but for energy importers there really is no alternative to moving towards alternative energies, and all countries have some interest. Most governments tend to subsidize alternative energy as an investment towards the future. Setting the right level of subsidy (often Feed-in Tariffs) is a very challenging and interesting project, unfortunately outside the scope of this current work, which focuses only on consumption subsidies.

1.2.2.8 Energy Subsidies

Energy (Consumption) Subsidies are used to keep the price of necessities affordable for poor families. Because of the high price of energy and it's importance for quality of life, it is a frequent category to be subsidized along with medical care and in some countries, basic types of food. Another reason energy is often subsidized is that both inelastic supply and inelastic demand make world energy prices prone to wild fluctuations. Governments have the job of solving the country's problems and respond with short run subsidies which turn into long run subsidies. Subsidies have at times resulted in unreasonable long term expenses for governments. Therefore transparently short run methods are preferable.

1.2.3 Why Subsidies are a Middle Income Country Problem

Wealthy countries do not subsidize energy – in actuality it is usually taxed, and can be a substantial source of revenue for the state. The many negative externalities associated with energy use provide good justification for taxing energy heavily. Even with high levels of taxation, in wealthy countries energy is generally affordable at basic levels to even poor households. The exception is winter heating for the poorest households and occasionally developed countries may subsidized fuel oil for this purpose.

Energy policy for developed countries tends to emphasize technology to improve alternative energies – since rapid improvements can lead to energy supply in the future and especially a dominant place in the supply chain and trade of this new technology. Another focus is on conservation of energy – since everyone is already using so much. In addition energy security is an important issue as are environmental issues.

The poorest countries recognize the importance of energy for quality of life, but cannot afford to subsidize energy because of extreme budget constraints. The focus is usually to make energy available at all. Poor countries try to attract foreign investment to develop energy sources in exchange for foreign exchange, and their own use of the energy. They also try to build infrastructure such as electricity grids to make bring basic needs to as many people as possible. Finally small scale alternative energy can be a substitute for expensive infrastructure in isolated areas. In short, there is no money for subsidies.

Middle income countries often have a high level of income inequality partly due to some persons living in the modern economy and some left behind in the traditional economy. Another reason for inequality is that the country is not yet at the level that it can adequately meet the needs of those who cannot be productive members of society, such as the old, crippled, and even the young.

However, because of the modern economy there are enough resources to subsidize the poor through some form of energy subsidy. This is done rather than providing income transfers because there is little information about citizens and income levels at the household level because most of the economy is informal, because there may be a lack of administrative capacity or trust, and thirdly because subsidies are a useful political tool especially in democratic countries. Suffice it to say that energy subsidizes are a common characteristic of ,middle income countries.

The following table was produced by the (IMF, 2013) and shows the close relationship between energy subsidies and middle income countries. The table shows the share of energy subsidies relative to the government budget in each country. The author has extracted only countries in South and Southeast Asia. In this table, it is the middle income countries that use subsidies, while wealthy countries such as Singapore and Brunei do not, and poor countries such as Cambodia and Laos also do not.

Table 3 Asian Subsidies for Energy as a Percent of Government Revenue

Asian Subsidies for Energy (as a % of total government revenues)

Country	Petroleum	Electricity	Natural gas	Coal	Total
Bangladesh	7.56	22.12	13.45	0	43.13
Bhutan	1.39	n.a.	n.a.	n.a.	1.39
Brunei Darussalam	3.77	1.57	0	0	5.34
Cambodia	0	n.a.	n.a.	n.a.	0
China	0	0.68	n.a.	n.a.	0.68
India	6.75	1.72	0.9	0	9.37
Indonesia	14.51	3.69	0	0	18.2
Laos P.D.R.	0	n.a.	n.a.	n.a.	0
Malaysia	5.67	1.49	1.41	0	8.57
Myanmar	9.35	n.a.	n.a.	n.a.	9.35
Pakistan	1.02	10.23	19.89	0	31.14
Philippines	0	0	0	0	0
Singapore	0	n.a.	n.a.	n.a.	0
Sri Lanka	7.99	3.26	0	0	11.25
Thailand	0.66	7.24	0.61	1.08	9.59

Source: IMF 2013 (data generally from 2011)

1.3 In Defense of Energy Subsidies

1.3.1 Introduction

The years 2006-2013 saw increasing energy prices for almost all energy classes. During this era of high energy prices, the government of Thailand initiated a variety of programs to alleviate the adverse effects of these prices on the quality of life of those with low incomes. Energy policies controlling price were enacted across all sources of energy including electricity, benzene, Liquefied Petroleum Gas (LPG), and diesel. Electricity for households using less than 90 KwH was given for free. The price of benzene was fixed for some period as the world price continued to rise. Diesel was fixed at a maximum retail price of 30 baht per liter, Gasohol (Benzene with 10% ethanol) was sold at fixed or reduced prices. LPG (propane) gas was sold at a fixed discounted price for all of this decade.

Alleviating poverty was only one facet of Thai energy policy during this very active period. Another primary objective was ensuring sufficient future energy supplies through subsidizing alternative energy, including high subsidies for electricity produced using solar energy, lower subsidies for energy from other alternative energy sources such as biomass, wind, biogas, etc., and an ultimately unsuccessful campaigns to institute the use of nuclear power and increase the use of coal.

Additional energy policies were addressed at investing in energy sources in other countries, energy security, climate change, and to a lesser degree energy conservation. Finally some policies were designed to support industry, either directly related to energy, or with significant implications for energy use such as automobiles. Although the focus of this paper is primarily on those policies designed to alleviate poverty, it could be instructive to look at the distributional effects of other energy policies as well. Supporting alternative energy has put significant upward pressure on overall electricity prices in Thailand, with distributional effects. Furthermore, the support that the government has given to the automobile industry also has had a strong effect on distribution of benefits and energy use. The effects of the controversial first-car policy in which taxes were not required on the first car a family bought will be discussed briefly below.

Large companies have excellent information about their customers/clients but governments often know very little about their customers/clients. Although government policies are designed for particular purposes, such as to help the poor, or to help a region or group, it is often difficult to determine the effects of such policies since it is not clear who is receiving benefits. Little information is available about individual poor families, because most income does not flow though formal channels. Meanwhile, profiling is in its infancy in Thailand. Recent big data approaches allow for a better estimation of the effects of polices, and suggest improvements in policy designs in the future.

1.3.2 Review of Literature

In recent years, there has been much discussion of the expense and misalignment of subsidy programs around the world. As the price of energy rose to a peak in 2008, and remained high thereafter, programs that subsidized energy became very expensive. The International Energy Agency (IEA) estimated the total cost of energy subsidies to national governments at about 500 Billion dollars a year, or approaching 1% of global GDP at the peak of this period (IEA OPEC World Bank, 2010). In some countries, such as Indonesia, Bangladesh, and Pakistan, the subsidy bill was much higher in the range of 15% to 43% of government expenditure.(IMF, 2013) p. 68. This money, it was argued, could be better spent on other government programs. The program reason given for the expense of government subsidies was poor targeting. (World Bank, 2010a) If the subsidies reached only the poorest persons, the cost of the programs would be low. Therefore the problem could be thought of as an inability to target, rather than the subsidy programs themselves.

According to Arze del Granado, Coady, and Gillingham (2012), fuel subsidies are a costly approach to protecting the poor due to substantial benefit leakage to higher income groups. The top income quintile receives six times more in subsidies than the bottom quintile. Furthermore, according to Dartanto (2013), fuel subsidies in Indonesia leads to a severe budget deficit and worsened income distribution. Almost 72% of fuel subsidies are enjoyed by the 30% of the richest income groups.

Rao (2012) suggests that subsidies targeted only to kerosene-dependent urban areas would have a higher efficacy than broad-based subsidies. In urban areas, subsidies are progressive, and provide benefits of up to 5 to 10% of household expenditure among poorer households which lack affordable access to LPG and biomass. On the other hand, kerosene subsidies are regressive and of minimal financial value to poor rural households because household quotas are based on cooking needs, but kerosene is used predominantly for lighting.

Several studies in Thailand found that universal price subsidies benefited the rich. (Amranand, 2014; Lewis, 2011) Anand, Coady, Mohommad, Thakoor, and Walsh (2013) also propose that developing a system to better target subsidies will be a major factor in reducing subsidies over the medium run. Therefore, in the energy sector reform process, instead of blanket energy subsidies, targeted subsidies were proposed. For example life-line tariffs for electricity, which allow consumers to receive a subsidized tariff rate or free electricity for consumption below a certain level. In the Philippines, the Electric Power Industry Reform act introduced a life-line tariff schedule at a subsidized rate for poor households in 2006. (DFID, 2002) Indonesia has a similar plan. In Malaysia, targeted subsidies were in the form of giving smart cards to owners of fishing boats and public transport vehicles to buy a limited amount of certain fuels at a subsidized prices. (IMF, 2013) (IMF 2013) (IMF 2013) (IMF 2013) (IMF 2013) (IMF 2013)

(Nikomborirak, 2014) recommended targeted subsidies of LPG for the poor in Thailand, with eligibility for subsidies tied to those who receive free electricity. Unfortunately the connection which is conceptually strong, is difficult in practice because of the strong mismatch between those who receive free electricity (north, mountains, poor, small family) and those who use LPG (south, municipal, well off, large family). This is discussed further in section 10 of this paper.

1.3.3 Further Justifications for Subsidy Programs

1.3.3.1 A Small Subsidy can Make a Big Difference

Although energy subsidies have received a bad reputation over time, it is mostly because they are so poorly targeted. Many subsidy programs benefit the middle or upper classes with the majority of government funds going to these groups. If energy subsidies are well targeted, the share of a country's income that goes to pay for them can be extremely low. That is because with high income inequality as in Thailand, increasing the income of the poorest citizens by 10% would cost very little. Doing so would cost less than 1% of GDP. The challenge is to target the energy subsidies sufficiently to those who need them to make subsidy programs economical and efficient enough to be practical. Thus small income transfers can increase the utility or happiness of the country by a very large margin.

1.3.3.2 Big Data Provides a Solution for Better Targeting

This raises the question: Can energy subsidies be better targeted? Although governments have been trying to target subsidies for a long time, they have had only poor tools to work with. The advent of big data and increased computer power, allows the government to improve the targeting of subsidies through profiling much in the way that big business has been able to move towards individualized marketing. That is one of the primary purposes of this paper – to highlight the role that big data and improved econometrics makes possible.

1.3.3.3 Theory vs. Practice - Subsidies versus Cash Transfers

It is a well-established tenet of Economic Theory that it is always at least as good to provide a cash transfer as it is to change the price of a good. A cash transfer would allow individuals to select only goods they themselves preferred which could indeed be, but need not be energy, and with welfare effects at least as high as with the subsidy. This is summarized in the Second Fundamental Welfare theorem (Arrow & Debreu, 1954) "under the assumptions that every production set Yj is convex and every preference relation <i is convex and locally non-satiated any desired Pareto-efficient allocation can be supported as a price quasi-equilibrium with transfers." (Mas-Colell, Whinston, & Green, 1995). Therefore it is generally proposed that the poor should be identified and a cash transfer should be made to them in lieu of any subsidies on goods.

Evidence on the other side of this debate comes from three directions. The first is the requirement for perfect information. Strangely enough it is not easy to identify who is poor. Most income and salaries do not go through a formal accounting framework. Only about 20% of the people in Thailand file income tax, and only about 30% have any paper trail of their income at all (Interview, Thai Tax Official).

The second argument is about transactions costs. Although, in theory, it is possible to go to every house in the country every 5 years, and find out if they are poor, the costs to do so would be prohibitive, nor does the government have the capacity to take on such a project. Nor is there a centralized way to transfer funds, although the PromptPay program currently under review in the cabinet may alleviate this problem. Subsidies, if well designed, can be a cheaper and easier solution.

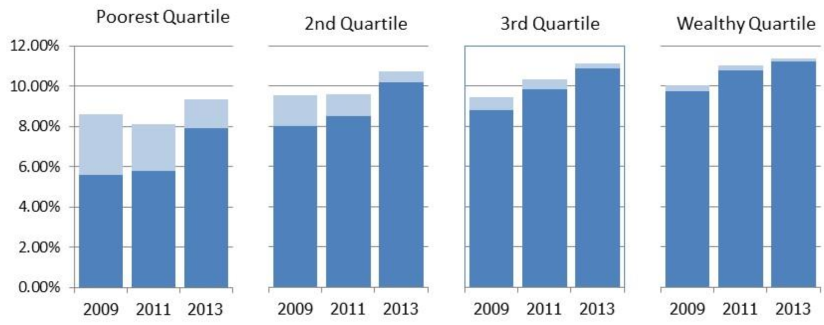
The final issue is that subsidies are not usually an economic decision, but a political one. A one time decision to pay cash to the poor does not work for political campaigns, and so when they have been used before in Iran and Latin America, it just leads to new subsidies being introduced on top of the cash transfers. (Salehi-Isfahani, 2011)

In fact, if Economists want to work with politicians, the author feels they should design short-run and high profile beneficial programs, that can be completed by one government, otherwise there a surfeit of overlapping programs with increased chance for corruption.

1.3.3.4 Rising Share of Expenditure Used on Energy

The final reason to support energy subsidies is an increasing burden and expenditure on energy. For all income groups and for all ages, the share of expenditure going to energy has been rising over time. In 2009, share of expenditure on energy ranged from 8-10% while four years later, share of expenditure ranged from 9 to 11.5%. This suggests that energy subsidies should be of increasingly concern to the government. Figure 1 looks at energy use by quartile, with wealthier quartiles using a greater share of income on energy. Dark bars show cash expenditure, grey bars show energy received for free. For all income groups, the share of the budget used for energy has increased between 2009 and 2013. The gray part of each bar is the energy received for free, and suggests that the country is doing a worse job of aiding the poorest and oldest citizens as shown by the large increase in cash used on energy, and the decrease in the energy for free. For the poorest decile, cash spent on energy increased from 5.9% of expenditure to 8.5%. A large part of this expenditure comes from tighter rules about free electricity. A second cause is a reduction in free firewood.

Figure 14 Total Share of Expenditure Spent on Energy (Expenditure Quartiles)



NOTE: IN THE TABLES, THE DARK BAR SHOWS SHARES OF CASH INCOME SPENT ON ENERGY WHILE THE GREY SHOWS FREE ENERGY. FREE ENERGY CONSISTS OF SUBSIDIES ON ENERGY OR IS ENERGY COLLECTED AS FREE FIREWOOD OR HOME-PRODUCED CHARCOAL. TOTAL EXPENDITURE HAS BEEN RISING FOR ALL GROUPS WHILE FREE ENERGY HAS BEEN RAPIDLY FALLING SUGGESTING STRESS ON POORER HOUSEHOLDS. SOURCE: SES 2009, 2011, 2013

Figure 2 shows the relationship between age of household heads, and energy use. Adults in the age range 40-59 use the highest share of expenditure on energy. If the household head is past retirement, their energy use drops. Again, the country seems to be doing a worse job of caring for its old people as cash spent for energy has increased significantly for the beyond retirement group.

Figure 15 Total Share of Expenditure Spent on Energy (Age Groups)

SOURCE: SES 2009, 2013

1.3.4 Conclusion

This chapter gave a brief introduction and overview to the tools that would be used in the rest of the dissertation. The second section introduces eight main types of energy policy, and after explaining each briefly, justifies why energy subsidies are of particular concern to middle income countries. The third and last section gives a brief defense of the use of energy subsidies. Mainstream economics is strongly against commodity subsidies as they distort prices and demand, and can lead to a less than optimal maximization of welfare relative to giving cash instead. This is usually due to poor targeting of benefits. The reasons why middle income countries continue to use subsidies is explained as a result of lack of perfect information, political incentives to provide subsidies, and high transaction costs. This paper argues that subsidies can be designed and used in ways that are good for the poor without putting undue strain on the economy. The tools in this dissertation can help policy makers do so.

Chapter Two – Methodological Issues and Introduction to Energy and Poverty in Thailand

2.1 Introduction

This chapter addresses two background issues relative to the rest of the dissertation. Most of the data used in the dissertation are from the Thai Socio-Economic Survey. Section 2.2 looks at some of the methodological decisions and issues that pervade all other chapters. These issues include expenditure as a proxy for income, using household rather than per capita income, including goods received for free as well as purchased goods. There is often no perfect decision on these issues, but the reasoning for the choices made are explained. Section 2.3 reviews Energy Issues in Thailand giving some background to existing policies and their rational, as well as more general issues that affect energy and poverty.

2.2 Methodological Issues

We use household data, made up of the National Statistics Office Socio-Economic Survey data for 8 years, 2006-2013. In each year about 0.4% of all households in are sampled. This a sample of 42,000-44,000 households per year throughout the kingdom. A representative multistage sampling frame was used in data collection. Stages include region, province and administrative area (urban – rural) Provinces are divided into primary sample units based around amphoe and then housing blocks are identified. Fifteen households are then chosen at random.
Three attempts are made to visit each previously identified household. Out of 52,000 chosen households, the teams are generally able to contact about 80% which gives the sample size of 42,000-44,000 households per year.
Survey weights are used to adjust for frequency of sampling issues. Households in rural areas are harder to visit and have somewhat lower representation. The other difficult area is Bangkok, where households live in apartments and are perhaps not often home. On average, each visited household represents about 250 total households, but this may vary from 150 to 700 depending on the area, with sampling weights being highest in the aforementioned Bangkok and rural areas. Total number of observations over the 8 years is approximately 330,000 households. The questions in the survey consist of about 400-600 questions depending on the year.

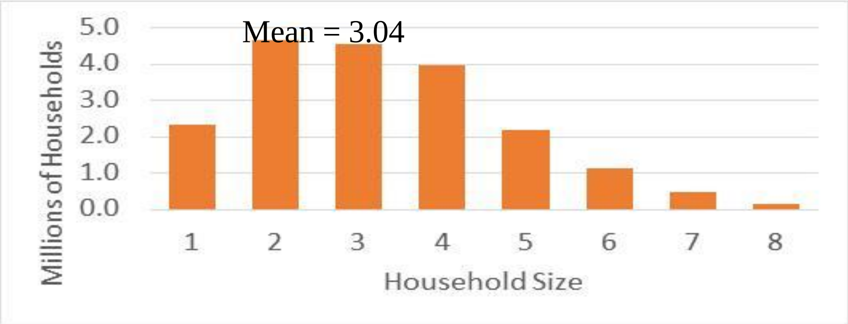
2.2.1 Household versus Per Capita or Equivalency Data

To consider distributional issues, we use household income as a dependent variable. There were a number of candidate measures. Most important is whether to use household data or per capita data.

Household data was chosen because:

1) The data is collected at the household level.
2) Many energy products make sense at the household level.
3) It is easy to add family size as a variable in regressions, but the converse is not true.
4) Family size has small standard deviation with 2/3 of families being 2, 3 or 4 persons.

Figure 16 Family Size in Thailand



SOURCE: SES 2013

This issue is taken up again in more detail in Chapter 4 (Paper2) in the section on energy poverty lines. Unless otherwise specified data is used at the household level.

2.2.2 Expenditure rather than Income as a Proxy for Income Level.

Another issue is whether to use the series for Income or the series for Expenditure when determining how well-off a family is. The two series should be similar, Most studies use expenditure, or quoting the World Bank

> *"Consumption is conventionally viewed as the preferred welfare indicator, for practical reasons of reliability and because consumption is thought to better capture long-run welfare levels than current income." (World Bank, 2001)*

Income can be much more volatile as "Income in the previous month" depends on the time of year, e.g. for farmers. Others referenced in this regard are (Deaton, 1997), and (Meyer & Sullivan, 2003). In short, monthly expenditure was chosen since:

1) Expenditure is a more stable series. For instance, a farming family receive most of their income in a single month which might or might not be the survey month.
2) Income was not collected for some of the years of the survey.
3) It is less likely that people will prevaricate about expenditure.
4) Unlike income, expenditure cannot be negative.

Generally when income is referred to in the dissertation, what is actually meant is expenditure as a proxy for income.

2.2.3 Cash and Non-Cash Expenditure

The Social Economic Survey includes both 1) goods and services that are bought and 2) goods and services that are collected, gathered, given or received for free. This is done to better account for quality of life in Thai households. If money income is a constraint it may be decided to use only purchased goods when doing some types of research. In this paper we use both purchased and "non-purchased items.
Combined cash and non-purchased expenditure were chosen because:

1) Free goods can enter regressions as a dummy variable
2) For inelastic goods such as energy, demand will be similar even if some of the good is not purchased.
3) Non-purchased goods make up only a small proportion of energy related goods, except electricity and firewood.

Altogether about 22% of "expenditure" was not purchased. Most of the non-purchased value is rent for owned housing. For the poorest households, non-purchased goods account for a higher 47% of the total, but not too much of that is energy.

2.2.4 Missing and Zero Values

There are no zero values in the SES survey, except where they are used occasionally in dummy or categorical variables, and even these are rare. A missing value in an expenditure category could in theory be construed as either a true missing value or as being zero expenditure on this item. The SES asks a long list of expenditure questions, and so for convenience all zero expenditures are denoted as missing values. To avoid confusion the National Statistics Office ensures that all questions are answered. (Interview, anonymous SES survey enumerator).
Sometimes in the research, we would like to know the average of persons who spent money on an item. For instance, in chapter 5 we see that people spent 300 baht for a tank of LPG if they bought one. At other times it is useful to also average in zero values for households that don't purchase that category, so for Engel curves, regular users of LPG use about 140 baht a month on LPG –they don't buy a tank every month. Therefore the research often converts missing values temporarily to zeros to answer specific questions. The data is stored in its original form with missing values.

2.2.5 Survey Weights

The SES multistage survey contains survey weights. Although econometric programs have the capacity to specify survey weights in the beginning of a dofile or program, the weights then become very "black box" as it is not clear when weights are being used in what form. The author chose to use survey weights that are specified explicitly in every command.
Survey weights represent the number of households that the sampled household represents. So if the weight is 342.321 this household represents 342.321 similar households. Adding up the weights for all sampled households will give you the total number of households in Thailand – about 20 million households.

The weights can be used in several ways – i) analytical weights, ii) probability weights, iv) frequency weights and iv) user defined weights. The weight variable a52 is the same in all cases, but for analytical weights the weights are squared for use in variance or other squared equations. Probability weights are used as is, whereas frequency weights need to be integers. Since some commands require frequency weights such as for tabulation, the author created a rounded version of a52 (called a52r) to be used where necessary. Stata seems to use analytical weights in regressions by default when calculating standard errors which made for very high t-statistics. I reasoned that variation came from the number of sampled households, not from the 20,000,000 households in the country, so that probability weights would be more appropriate. I used probability weights in my regressions and received lower t-statistic values but still strongly significantly different from zero.

There were times especially early in the process of the dissertation when I was working with many per capita variables that I really needed a per capita weight. For instance it might be useful to know what percent of people lived in a house with a color TV. I could find the percent of households with a color TV by using the household weights, but without knowing how many people were in each house I could not know the number of people who could watch. I created another weight a88 = a52*a04 where a52 is the household weight and a04 is the number of people in the household. Summing up all the values of a88 gives you the 65 million people in Thailand. I discussed this formulation with several different statistics people from NSO and they could find no problem with using it.

2.2.6 Data Frequency of Thailand's Socio-Economic Survey

The data used in this study are mostly from the Thai Social Economic Survey, conducted annually. However a marker is included in the SES indicators denoting in which month the data is collected. The data is for the prior month, so that actually January 2011 survey results are tabulating information for December 2010, so there is a one month lag that must be corrected for when matching survey data with price data. Although the survey has been conducted every year since 2006, income variables are only collected in odd years (2007, 2009, 2011, 2013), and there have been some changes made to variables. In general the analysis below either uses all practically useful data or selects a few years to give a feel for the issue in question.

The following summary statistics for 2013 give an indication of critical values used in the study. Households are divided into quartiles (25%), and median values are supplied for each quartile.

Table 4 Median Expenditure Education and Age by Quartile

	HH Expenditure Baht/Month	All Energy Baht/Month	Share of Expenditure	Median Highest Education	Age of HH Head
Wealthy	33069	3800	11.5%	9.6 (2-yr college)	51
Quartile 3	17687	1700	9.6%	7.3(upper second)	50
Quartile 2	11414	1095	9.6%	5.8 (lower second)	52
Poor	6783	593	8.7%	4.3 (primary)	58

SOURCE: SES 2013 NOTE: IN 2013 THE EXCHANGE RATE WAS APPROXIMATELY 30 THB: 1 USD

The median household in the wealthy quartile had monthly expenditure of 33,000 baht or about 1,100 USD a month which is 13,000 USD a year. The second quartile spent just over half of that at 17,687 or about 600 USD. The poorest households used a median of 6,783 baht or about 220 USD a month. Since households had an average of 3 persons, per capita numbers were approximately 1/3 of this.
Households spent 8.7-11.5% of their income on energy, with poorer households on the lower end. Education level reflects the person in the household with the highest education level. The wealthiest household had a median educational level of almost 2 years of college or vocational school, while the poorest had completed only primary school. The age of the household head was nearer 50 reflecting the tendency for several generations of Thais to live together. The exception is the poorest quartile where the age of the household head was older at age 58.

Table 5 Median Expenditure and Share of Households Using Electricity and LPG by Quartile

	Electricity			LPG		
	Baht/Month	KwH	Share Using	Baht/Month	Kilograms	Share Using
Wealthy	820	241	99.9	107	5.2	83.7
Quartile 3	450	149	99.7	95	4.6	71.8
Quartile 2	320	108	99.2	64	3.1	58.6
Poor	220	80	97.8	29	1.4	42.8

SOURCE: SES 2013

Electricity: The most interesting takeaway is that usage does not change very much for the bottom ¾ of the population. The median usage for most of the population of 80-149 KwH a month is not nearly enough to run an air conditioner, but should be enough for a refrigerator, TV, lights and some small accessories like cooking pots and fans.
LPG: Less than half of poorer households use LPG, and for those who do, the monthly expenditure on LPG is low. Usage refers to those who reported they use LPG as the main source for cooking. For some households it may be that little cooking occurs in the house besides cooking rice in the universal electric cooker. For those who use LPG as the primary cooking fuel as indicated in the survey the purchase of a 15 kg LPG tank generally does not occur every month, resulting in a 300 baht cost at 2-3 month intervals. It is unlikely that this average 100 baht per month is a significant expense, especially as poorer families are more likely to cook with collected firewood or their own charcoal. It is not obvious that LPG needs to be subsidized. Furthermore, most of the subsidy will go to wealthier families, as shown in later sections.

Table 6 Median Expenditure and Share of Households Using Gasohol 91 and Diesel by Quartile

	Gasohol 91			Diesel		
	Baht/Month	Liters	Share Using	Baht/Month	Liters	Share Using
Wealthy	1000	27	42.2	2000	66.6	49.6
Quartile 3	700	19	42.2	1500	49.3	29.4
Quartile 2	500	14	38.5	1000	33.0	14.7
Poor	400	11	27.6	600	19.7	3.1

SOURCE: SES 2013

Table 7 Median Expenditure and Share of Households Using Benzene 91 and Benzene 95 by Quartile

	Benzene 91			Benzene 95		
	Baht/Month	Liters	Share Using	Baht/Month	Liters	Share Using
Wealthy	800	18	21.8	1000	21.2	8.4
Quartile 3	650	14	23.8	600	13.3	6.4
Quartile 2	500	11	25.7	500	11.0	6.6
Poor	400	9	20.4	400	8.5	4.4

SOURCE: SES 2013

As a support for alternative energy, the Thai government has subsidized the use of gasohol (gasoline with 10% ethanol) with a variety of subsidies and promotions including a price reduction for gasohol 91 relative to its substitute, benzene 91. However, in 2013 the government tired of its attempts to persuade citizens to use gasohol through lower prices, and forbid the sale of the alternative, benzene 91. Before 2013, pricing policies resulted in low uptake as there were concerns that gasohol would hurt motorcycle engines. The government resolved the issue in 2013 by stopping the sale of benzene 91 in Thailand, meaning that most motorcycle drivers were not given a choice and switched by necessity to gasohol 91. As 81% of households outside of Bangkok own a motorcycle (SES 2013) the use of gasohol jumped in the beginning of 2013.

By looking at households who own a motorcycle but no other vehicles, it can be observed that about 6% of motorcycle drivers decided to switch to benzene 95 rather than switching to gasohol 91, presumably due to concerns about their engines. The percent of motorcycle drivers using benzene 95 prior to 2013 was about 2%. That figure jumped to about 8% after the switch to gasohol.

In data for 2013 there are about 20% of people reporting that they were still using benzene 91, even though it was no longer for sale. Since benzene 91 was not available (except for January surveys which record data from December.) it is assumed that these people were mistaken. Inattentive users probably would not notice the switch between benezene91 and gasohol 91 as they are essentially the same fuel used in the same way. One explanation is that some motorcycle drivers buy benzene/gasohol from brown beer bottles or other repackaging by the side of the road.

Diesel is used primarily as a fuel for pickup trucks and farm trucks. In rural areas, pickup trucks outnumber cars with a ratio of 2:1. Although diesel was slightly subsidized (the price was capped at 30 baht/liter), many rural users were not even aware of the subsidy program, so it is not clear why the subsidy program was needed. It appears that the subsidy was aimed at the transport sector rather than at farmers. (Chuangwilai, 2014)

2.2.7 Education, Region, Age

Income is not the only relevant measure of distribution. From a political point of view, the region in which benefits accumulate is important in that it may sway a political party to select policies which support its political constituents. In this paper geographical distribution is supported through the use of regressions at the regional level and by provincial level mapping for more detailed analysis of geographical distribution.

Educated households may use energy in ways that differ from less educated households, especially when it comes to alternative energy or conservation. This is the thesis of another forthcoming paper (Chankrajang and Lewis 2017). Policies may benefit one education level more than another.

A household is made up of several individuals with differing educational backgrounds. Taking an average of these backgrounds is difficult since anyone under the age of 20 has likely not reached their final educational attainment. The head of the household is likely older and not the most educated person in the household. In this paper we used the highest educational attainment of anyone within the household as a measure of education.

Elderly households have different expenditure patterns from younger households, and may benefit from different subsidies. For anyone working with income distribution in Thailand, it is obvious that the poor in Thailand are often those beyond retirement age. There is very inadequate support for those above the retirement age at this time. A household is made up of many individuals and an average of ages might give very different results if there is a newborn or an elderly person in the household. This paper uses the age of the household head as a measure of age. Generally the oldest working person will be considered the head of the household. Elderly person who are cared for by the family would not be considered “head of household”.

2.2.7.1 Deciles

To calculate per capita deciles, households are multiplied by weights and by number of members of households. For instance, if a household stands for 350 similar households, and if it has 3 members, that person stands for 3*350 = 1050 persons. This process is a bit awkward, and it is best done through an automatic do file. Sizes of per capita deciles (decile 1 is poorest) are as follows:

Table 8 Per Capita deciles

Decile PC	Households	Avg. Weight	Avg. Family Size	Population
1	2,426	756	3.53	6,471,161
2	2,782	672	3.46	6,471,161
3	3,277	583	3.39	6,471,161
4	3,655	532	3.33	6,471,161
5	4,167	479	3.24	6,471,161
6	4,457	465	3.12	6,471,161
7	4,699	447	3.08	6,471,161
8	5,349	403	3.00	6,471,161
9	5,694	393	2.89	6,471,161
10	7,338	305	2.89	6,471,161
Total	43,844	472	3.13	64,711,610

SOURCE: SES 2009

With per capita deciles, the number of households included in each decile increases with income as lower per capita income is tied to many family members in the household. We will see that Household deciles are much more evenly spaced. Household deciles are as follows:

Table 9 Household deciles

Decile HH	Households	Avg. Weight	Representing # Households
1	3,975	493	1,957,922
2	3,801	515	1,957,922
3	3,982	492	1,957,922
4	4,108	477	1,957,922
5	4,255	460	1,957,922
6	4,492	436	1,957,922
7	4,581	427	1,957,922
8	4,908	399	1,957,922
9	4,785	409	1,957,922
10	4,960	395	1,957,922
Total	43,844	447	19,579,220

SOURCE: SES 2009

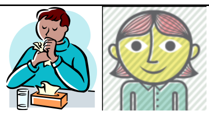
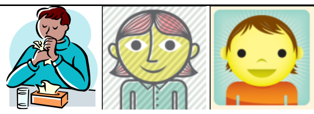
Besides these two types of deciles, we also tried using adult equivalency deciles whereby children under the age of 15 were counted as half a person, and adults as one person, (following (Sarntisart, 2011)). It was hypothesized that this would make a good intermediate index between household and per capita measures of energy use. Although this approach is still appealing theoretically since children use less energy than adults, in practice the results were quite similar to using per capita data, and judged to be not worth the extra trouble and confusion of calculating and explaining them.

Another alternative would be to measure only the adult members in a household in computing deciles, as some trial simple OLS regressions of energy use vs. adults & children suggest that children use very little energy relative to adults. 21% of the approximately 110,000 people included in the survey were children under 15.

Equivalency Scales - Finally, although there is not time to include it in this version of the paper, there are several equivalency scales that are being used in various parts of the world. Statistics Canada uses a 40:30 scale whereby the first adult counts for 1, each additional adult counts for 0.4, and each child counts for 0.3. Through empirical studies they have found that this best fits characteristics of poor houses found in Canada, so that it is possible to compare a single adult making $25,000, with a family of four making $50,000. The OECD has adopted an equivalency scale that counts the first adult as 1, additional adults as 0.5 and children as 0.3.

Recently, a somewhat simpler alternative has become popular whereby household income is divided by the square root of the number of household members. This provides a very simple measure that approximates the two equivalency scales above. After experimentation with both per capita and household deciles, it was decided that if an equivalency scale were desired, this last criteria was the best, as it meets both the criteria for simplicity and the criteria for transferability between countries. It could be desirable to use an equivalency scale in some cases as it resolves issues of household size that make both the per capita and household deciles undesirable as explained in the following table. However, interpretation becomes very muddled.

Figure 17 Comparing household and per capita electricity use

Family Members	House Uses	Each Person Uses 2
		
		
		

SOURCE: AUTHOR

Suppose that we have three households, and that each person in each household uses two units of electricity, BUT there is a fixed cost of running a house that requires one additional unit. Now suppose we were to rank these households in terms of household usage. In creating household deciles, the bottom household would be the richest, and the top one, the poorest. Now what if we were to use per capita deciles? Then precisely the opposite would be true - the top household would be the richest, and the bottom household would be the poorest. But in fact, neither of these two cases is true. All three households use the same electricity per person.

2.3 Energy and the Poor In Thailand

2.3.1 Overview

Thailand is a middle income country which finds itself in the position of strongly supporting energy policy for the poor in the short run, while ignoring or damaging energy policy for the poor in the long run. This cycle is driven by political objectives which are by necessity short-run. The focus of this section is about how to better design the long term energy policy for the poor, not only through classic solutions such as conservation and appropriate technology, but also by building financial and legal structures that will support the equitable distribution of energy in the long run. Public private ventures into alternative energy would give the government a stake into future energy production, Improved property tax regulation would give the government more options in terms of taxing and controlling energy production in the future to better redistribute the benefits. Independence for the Ministry of Energy would allow for more rational energy pricing. More targeted approaches to subsidizing at risk groups, would leave more money available for supporting energy research, and for improving public infrastructure. An improved rail system, especially for the transport of goods, is essential to keep prices of goods low in the future. Finally, rational energy pricing which encourages all of us to use energy in careful and responsible ways will help prepare us for the future.

Thailand is a Middle Income Energy Importer

Thailand is a middle income country with gross per capita income of 5815 USD or about 16340 USD on a PPP basis. Income inequality is still high but decreasing with a Gini coefficient of 0.38. (World Bank, 2015) Income inequality results from a large poor rural population twinned with a modern developed city. Thailand is surrounded by countries which have lower levels of per capita income and is a magnet for international migration, so that some of the poorest groups are migrant workers from surrounding countries.

With a middle income level, poor persons do have access to modern forms of energy. 99% of houses have access to electricity, 64% of rural households use LPG for cooking, and there an estimated 23 million motorcycles in the country, the chosen transportation for the poor, for a total population of 67 million (SES 2013). Heating is not a problem as the average daytime temperature is 33 C and the average low is 23 C with only minor variations between seasons (Thai Meteorological Department).

Which Currently has a Pro-Poor Energy Policy

The government of Thailand has sufficient resources and political motivation to subsidize the use of energy for its poor people and has arguably done too much rather than too little in manipulating energy prices in favor of the poor. Electricity is subject to a rising tariff based on quantity used, which subsidizes poor households, while taxing larger industrial users.

Table 10 Electricity Block Pricing in Thailand

Consumption	*Baht*	*US$*
First 15 kWh (0 – 15th)	1.8632	0.05
Next 10 kWh (16th – 25th)	2.5026	0.07
Next 10 kWh (26th – 35th)	2.7549	0.08
Next 65 kWh (36th – 100th)	3.1381	0.09
Next 50 kWh (101st – 150th)	3.2315	0.09
Next 250 kWh (151st – 400th)	3.7362	0.11
Over 400 kWh (401st – up)	3.9361	0.11

SOURCE: THAI BOI (2017)

The most common cooking fuel used by the poor is Liquid Petroleum Gas (LPG), which has enjoyed price supports for 25 years. Due to recent drop in world energy prices it has recently been possible to reduce the subsidies on LPG, but prices are rising again on the world market. During the entire period of data used in this dissertation, LPG has been heavily subsidized. Currently (Jan 2017) subsidies on LPG are about 4.70 baht/kg or about 20% of the retail price. (Bangkok Post, 2017) It is unusual and expensive for a net importer to subsidize energy use. When the subsidy was first instigated, and for the ensuing 15 years, Thailand was a net exporter of LPG, but the continued subsidy combined with very high energy prices, attracted energy users from a variety of sectors particularly transportation and industry. Thailand is currently a large net importer of LPG. The low price of LPG has retarded development of new production capacity since any new LPG produced would be sold below cost. Compressed Natural Gas (CNG) faces similar problems but has a smaller share of the retail market.

Table 11 Subsidized prices for period of study.

	Price Baht	Price USD	World Price USD	Appropriate Price/kg based on World Price
LPG	20.29/kg	0.58/kg	510/tonne	0.70/kg
CNG	12.55/kg	0.36/kg	300/tonne	0.50/kg

SOURCE: PTT (2017)

Even Diesel, a popular fuel for the transportation of farm produce, has been subsidized in recent years. The subsidy has come and gone, at times putting extreme pressures on government finances. Historically the government has capped the price at 30 baht/liter for diesel, subsidizing when the world price pushes diesel prices above this level. With current low global prices for crude oil, the diesel price is 26.44 baht/liter and does not require subsidization.

Thailand has been so conscientious in promoting energy for the poor over the years partly due to idealism, but also partly due to populist politics. With more than half of its population living in rural areas, and an electoral system that results in very frequent elections, the needs of rural areas dominate each election period. Energy has an impact on the life of rural people directly, and is an important contributor to the price of food as well as to fertilizer. Even as many successive governments have advocated raising the price of LPG to more reasonable levels, none have dared to do so.

In recent history, only one energy minister, Piyasvasti Amranand, an Economics Ph.D. from LSE, has been able to buck the trend of populism and attempt to rationalize the use of energy in Thailand by promoting competition, rational pricing, alternative energies and conservation policies. Active in energy policy in the 1990s, he returned briefly as Minister of Energy and instituted many changes. Notably, he was the energy minister installed by army generals after the much condemned coup of 2006, and after an elected democratic government returned, so did short sighted populist policies.

But is finding subsidizing energy to be expensive

The cost of the government's policies is oppressive. In the first four months of this year, the government has spent about 40 Billion Baht (USD 1.3 Billion) subsidizing energy, only half of which recouped from energy taxes on other forms of energy. At this rate, energy subsidies will make up about 3% of the government budget this year. Furthermore, much of the money leaves the country, as the government must buy LPG for domestic use on the world market. Total imports of energy made up about 18% of total Thai imports, and energy imports represented about 12% of Thai GDP. Another problem with current policy is that energy cannot be used as a form of general tax revenue which puts an additional burden on other taxes, and reduces possible government services. In other countries, energy taxes make up a significant percent of tax revenue, but Thailand redistributes nearly all of its

And is Ill-Prepared for Future Higher Prices.

Although the government has been assiduous in protecting the poor from short term shocks to energy prices, they have done little to prepare the country for long term high energy prices or possible shortages that are likely to be a part of life in the next couple of decades.

Pro-poor policies need to developed and in put in place to help prepare Thailand for a high energy price future. The possibility exists to try to design policies in a way that would protect the poor from the worst effects of long term high prices. Part of that process is to allow everyone to adjust to high energy prices, not by sheltering the country from the reality of world. But much can be done to prepare the country through long term macroeconomic policies as addressed below.

2.3.2 Poor and Energy Worldwide

2.3.2.1 Basic Human Needs

Humans even when they are the poorest of the poor, have basic needs of food and shelter and will find some way to meet them. These needs are not optional and need to be protected and made available especially in a way that does not require the person to spend all their time and resources to acquire them.

The UN recognizes the importance of these basic needs and have included eradication of poverty and hunger as part of Millennium Goals development goals. In terms of energy, basic human needs are as follows:

- Energy is needed for Cooking
- Energy is needed for Heat

- Energy is needed for Light
- Energy is needed for Transportation
- Energy is needed for Communication

2.3.3 Overview - What is Happening in Poorer Countries

2.3.3.1 Introduction

Before addressing the situation in Thailand in more detail, the following is a quick review of the serious issues being faced by other developing countries. For further information on this topic, the reader is referred to the excellent cited references. Thirty eight percent of the world's population still uses traditional biomass for cooking, and sixteen percent of the world's population has no access to electricity. (International Energy Agency, 2016) Therefore, the problems being faced by Thailand as a middle income country with access to modern forms of energy are very different from those faced by the poorest countries in the world. These areas include most of Sub-Saharan Africa and South Asia, and many other developing areas both in the north and in the south which face a lack of access to modern energy.

Access to energy is a prerequisite for poverty reduction and human development, and a lack of access to modern forms of energy limit the potential for poor people to better themselves. The time burden of collecting fuel for cooking and heat, the negative health effects of localized pollution, the lack of adequate sanitation, and a lack of refrigeration, the difficulty of studying without access to light, and extreme conditions faced by those without access to heating and cooling can all undermine the health and ability to develop of the poor.

The non-profit organization Practical Action proposes the following guidelines as minimum standards for access to energy at the household level. These goals could be achievable by the year 2030 and should be adopted as an addition to the UN millennium goals.

Table 12 Energy Requirements per Household per Year

Lighting	360 KwH per year (or 30 KwH month)
Cooking	88 kg wood, or 28 kg charcoal, or 35 kg LPG or 18 liters ethanol per person 40% more efficient stove than 3 rocks Particulate matter less than 10 μg/m^3
XXX Heating &Cooling	Minimum indoor temperature of 12^{0}C Maximum indoor temperature of 30^{0}C
Information	Receive/send information to the outside world
Livelihood	Access to energy needed for livelihood

SOURCE: TENNAKOON (2009) ASSUMES FAMILY SIZE OF 3

2.3.3.2 Health

Poorer families in lower income countries make extensive use of biomass for cooking and heating and use kerosene lamps for light. These energy sources result in high levels of particulate pollution in homes, resulting in high levels of pulmonary infection and exposure to carcinogens. As woman and children are often responsible for the cooking and gathering the firewood, they are especially at risk.

Woman and children are also responsible for collecting the biomass which is arduous and can take up to three hours a day. The time spent collecting firewood and dung as well as a lack of adequate lighting later in the evening restricts access to education for poor children.

In colder climates, health problems are compounded as open fires inside confined spaces results in poisonous gases as well as very high levels of particulate matter. Health facilities for treating diseases are often remote and poorly equipped.

2.3.3.3 Appropriate Technology

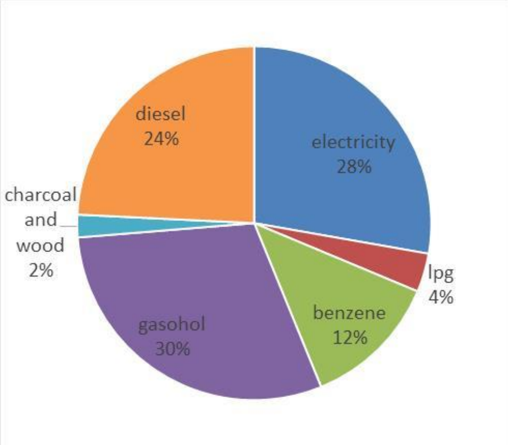
Poor persons are often isolated from centralized sources of electricity and power, and multifarious efforts have been made to design appropriate technology that can be used at a local level in isolation from the grid. Appropriate technology efforts have been an objective for helping the poor for at least half a century, and are focused around – better cooking stoves, better designed lamps, better sanitation, better management of organic systems and localized energy sources, such as localized solar, gas, or biomass production.

Better stoves which reduce the amount of required biomass, as well as reducing airborne emissions have been one approach to mitigating the problem of insufficient and biomass and high pollution levels. They have been successful in some places at reducing pollution. However the efficiency of the stove reduces the heat given off since the firewood heat energy is being used purely for cooking. This may not be acceptable in countries in which part of the purpose of the stove is for generated heat. Solar heaters for water are another easy way to improve the quality of life of rural dwellers. Small scale biogas production can use available material to create cooking gas for the household improving sanitation and air quality. Further information about appropriate technology is widely available and beyond the reach of this paper.

2.3.3.4 Energy and Expenditure by the Poor

According to the 2013 Social Economic Survey the direct use of energy by all groups in Thailand is about 10% of their total budget, but the particular energy form used varied by region and income group. In Bangkok use of gasoline and electricity were higher than other regions, while in the Northeast, LPG, charcoal and diesel were relatively larger. Expenditure on energy per household in 2013 for the whole country was 2034 baht (58 USD) a month. Total expenditure on all goods was 19,061 (544 USD) per household.

Figure 18 2013 Expenditure on Energy in Thailand



SOURCE: 2013 SOCIAL ECONOMIC SURVEY

In a 2010 summary report of energy studies from developing countries, the World Bank (2010a) found direct expenditure on energy varied from country to country ranging from 4 to 20% of total expenditure in the sampled countries. The following chart shows some sample countries. Share of expenditure on energy may be lower in poorer countries since persons may obtain fuel by collecting it. At lower incomes, expenditure on biomass becomes more important. Kerosene is also relatively important in very poor countries. In Thailand, expenditure is high on gasoline for motorcycles.

Table 13 Expenditure Share on Energy in Selected Countries

	Bangladesh	Cambodia	Indonesia	Thailand
Kerosene	1.1	1.1	2.2	0
LPG	ND	0.1	0.1	0.6
Gasoline, diesel	0.1	ND	0.8	6.3
Electricity	0.7	0.5	3	2.9
Biomass	5.3	5	2.5	0.8
Total energy	7.3	6.8	8.8	11
Total food	64	72	67	46
Transport	2.4	0.2	1.8	1.3

SOURCE: (BACON, BHATTACHARYA, & KOJIMA, 2010), P. 41

2.3.4 Energy and Food Prices

Food is a sizable portion of the typical poor person's expenditure, making up 40 to 80 percent of total expenditure. The same World Bank (2010a) study cited above found that in some countries an increase in energy price increased food prices and that the indirect effect of increases in food prices had a larger impact on poor people than the direct impact of higher energy expenditure.

This is unlikely to be true in Thailand, which is on the lower end of share of income spent on food, and on the higher end of income spent on energy. Nevertheless, any increase in energy prices will have a substantial impact on food prices - through increases in transportation costs, cooking costs for prepared food, and fertilizer. Poor people spend a higher percentage of their income on food, and a slightly higher proportion of their income on energy according to the Household Energy Consumption Survey (Ministry of Information and Communication Technology, 2006), so they would be more affected by energy driven inflation than other groups. Fear of inflation is one of the reasons why the government is cautious about relaxing energy subsidies. Inflation, especially in terms of food, is a very visible and disliked measure of the lack of success of the government. Inflation is inevitable in an era of rising energy prices but it certainly can be delayed to be attributed to a future government. The government is joined in its fight against inflation by the central bank, which has a mandate to target a certain inflation rate.

Several recent studies look at the effect of energy prices on inflation. The Bank of Thailand in their 2008 Inflation Report (2008) uses the BOT Macroeconomic model and found that a one percent increase in the price of crude oil would lead to an increase in inflation of 0.02 percent in the first quarter, and an increase of 0.6% in the fourth quarter following a price shock. Starting from a base of 85-86 USD/barrel in 2008 and 2009, energy prices rose to a high of 120 USD/barrel in late 2009, a 41% increase, so the model would predict that inflation would increase from energy alone by 0.02*41=0.84 percent in the first quarter and about three percent in the fourth quarter. These numbers seem surprisingly low, but the main concept that some of the effect on inflation will be delayed by as much as a year is interesting. Another study found inflation elasticities for changes in crude oil prices after one year of about 0.25 which seem closer to what was observed.

According to research by Krungsri Bank (2011) based on Thai input-output data, sectors that should be most affected by an energy shock include all forms of transportation, where up to 50% of the cost is energy, followed by mining, agriculture, fisheries and construction materials. Indirectly, energy shocks have the largest impact on tourism, construction, real estate, and automobiles.

Higher energy and food prices may have some positive effects on the rural poor if the value of their output increases, but for day laborers, migrants, and particularly the urban poor, higher prices can only mean hardship.

Given the hardships entailed by inflation on certain groups, isn’t it better to continue subsidizing energy? The hard reality is that energy prices will go up, but isn’t there a moderate approach that can be taken? For instance, could Thailand continue to subsidize LPG, but not for all groups and for all purposes? It should be possible to buffer the poor from world prices a little by giving a fixed subsidy and letting prices adjust upwards but at a level lower than world prices, or by subsidizing a certain percentage of the world price. This approach was adopted in 2012-2013.

Another option would be to adjust the minimum wage upwards to cover the cost of inflation. The minimum wage would affect generally the poorest landless groups, since that is the group which receive minimum wage. In Thailand, some groups may be paid lower than minimum wage, such as illegal migrants who might get 70-80% of the minimum wage, or rural workers who are beyond the reach of government regulations. Given the tight labor market in Thailand at this time these wages would likely adjust upward as well. The net effect could be an improvement in income distribution if wage increases did not carry through to all groups. (Lewis, Lekfuangfu, & al., 2010)

2.3.5 Energy Prices and Ministry Autonomy

The Thai energy situation would be improved if there was increased independence for the Energy Ministry. Prices of basic commodities are fixed in Thailand for about 200 products in all. These include energy. Prices are controlled by the government and do not adjust rapidly to market forces. The original intention of this policy was price stability and protection of the poor from concentration in the production sector. As a country with a small economy, especially 20 years ago, it would be possible for a few producers to monopolize the market.

For most products, the controlled prices work well. Either consumers or producers can appeal current price levels, but generally it will be industry complaining when production costs rise. The government usually delays but eventually responds by raising the price similar to a "dirty float" system for commodity prices, that is reasonably fair to both parties.

Energy is an exception because it is such an important commodity. Petroleum prices are not set at the Ministry of Commerce, as are other prices, but are set through coordination between the prime minister's office, the Petroleum Trading Company of Thailand (PTT) which is a monopoly provider of some products, and the Ministry of Education. PTT, although being a private company since its privatization in 2001, works in close coordination with the government, often working in the interests of the government rather than its own shareholders. For example, PTT sold LPG and CNG at a loss for the past 10 years representing a loss of about 1 billion USD which had to be bad for shareholders. Although PTT is a public company freely traded on the SET stock exchange, the Ministry of Finance is its biggest shareholder.

Because of the close political ties between these organizations, and because the prime minister appoints the minister of energy, it may not be possible for the Ministry of Energy to raise energy prices. Because the elected government may benefit from short run subsidies such as for benzene that harm the long run financial stability of the country, the Ministry of Energy should be free to set prices in the long run interest of the kingdom. In much the same way as the Bank of Thailand needs to be free of political influence of the government for the good of the country, the Ministry of Energy should be free to set energy policy. Given the importance of energy to the overall economy and the poor track record of choosing political advantage over rational policy, the government should cede autonomy to the ministry. The ministry might then still control prices in using a managed float system if desired.

2.3.6 The Poor and the Fight for Land

For the past 100 years or so, land has been cheap. Returns to agriculture and land owners have been low, and because of this, rural land is often in the hands of the poor. It is likely that the next century will see some reversal of this trend. Over the course of history, land has been the most valuable of assets. If the value of land starts to increase again, protecting communities and small holder ownership of land will become an important issue again. One way to prepare for this would be to clear up the large backlog of cases in Thailand in which land tenure is under dispute. Right of possession after 20 years or even 50 years would be a good start in this direction.

2.3.6.1 Local Development versus Exports

Another well-tread story revolves around what is the best use of land from a national perspective? – Is it better to produce energy for the city or for export, or to produce energy as a local fuel. Local needs are great, but terms of trade mean that energy could be traded for something even more valuable. If so, would production be controlled locally, or by outsiders?

2.3.6.2 Food versus Energy – Competition for the Efficient Use of Land

The world's population is still growing rapidly and use of land for energy supplants the use of land for food. Perhaps more visible is when food crops such as corn, oil, and sugar are converted to energy, increasing the price of basic commodities. Some alternative crops which grow on marginal land such as Jatropha, grown for its oil seeds, or algae, grown for its energy value, replace food crops.

Increased efficiency of land use offers one solution. Poly-generation crops which can meet many needs at the same time, such as sugar (sugar, ethanol, electricity), rubber (rubber, wood), corn (grain, sugar, oil, ethanol, silage) offer hope for intensive use of scarce areas of land. Future hopes for cotton – seeds could be bred for oil, cassava – starch and ethanol, and for oil palm – remains could be used for biogas, are hopeful, among others.

Biofuels present opportunities for the poor as well as a risk. It will lead to higher returns for agricultural production of all kinds, and in many cases energy shortages may benefit the rural poor by increasing the value of their food and energy crops. More at risk, are the urban poor. Increased use of resources for energy production, with mechanized production will mean the urban poor are the most likely to be squeezed by high energy prices.

There are also many attempts to turn basic biomass or cellulose into energy. To date, all of these approaches, including the most developed of them, Jatropha, has not been shown to be adequately productive on marginal soils. Besides which, what is marginal land for humans may be the last refuge of increasingly beleaguered animal plant and insect species.

2.3.7 Tax Policy

2.3.7.1 Land Taxes

At present, the property tax system in Thailand is very weak, with almost no property tax at all. This has made property a favorite vehicle of speculation for the wealthy, even more than in other countries. Because there are no taxes on property, much property is sitting idle, as the owner is primarily interested in capital gain. A property tax would help bring land back into productive use.

2.3.7.2 Focused Subsidies

Energy subsidies should be much more targeted to the poor. The primary difficulty with energy subsidies is how to get the subsidy to the right person. Subsidizing a type of fuel because a certain group uses it, causes huge distortionary forces over time. In the short run, with low price differentials, there is little substitution between fuel types, but over time better technology is adopted that allows substitution of the cheap fuel for the fuel most naturally suited for a given purpose, so that eventually any fuel can imperfectly substitute for other types. If that sort of innovation effort were used to substitute to alternative or renewable fuels, we would all be better off and better prepared for the future. Prices send strong signals, but if the signals are distorted and lead us in the wrong direction it can waste time and be destructive.

In addition, price differentials drive the rate of switching between fuels so that as a fuel becomes much cheaper than its alternatives it becomes price efficient for almost everyone to use it. Thus the great switch from liquid gasoline and diesel in the transport sector in Thailand to less convenient gas.

Finally, large subsidies on a fuel in one country lead to large scale smuggling along its border so that in Thailand subsidized LPG finds its way across borders to be sold in Malaysia, Cambodia or Burma where LPG is almost twice as expensive.

When we impose a tax or give a subsidy the buyer and seller of the product are both affected. If it is a tax, the buyers see the higher price and buy less. The seller are concerned with the level of sales and reduce price somewhat to attract the buyers to return, thereby reducing their profits. Both absorb some of the effect of the tax. Depending on the elasticities of each group and their ability to absorb price changes, the tax or subsidy can have a somewhat larger effect on either the buyers or sellers. Economists refer to this as the incidence of the tax which depends on the relative elasticities of supply and demand.

If the government wants to tax a product, both buyer and seller will be negatively affected, but if a government wants to subsidy a product both buyers and sellers will be positively affected. The best product to subsidize is one in which both buyers and sellers are poor.

For many years LPG was the perfect product to subsidize since it was primarily used by sidewalk food vendors who include it in their sales price, and the food was bought by the urban poor working class. At that time, biomass was the main fuel in the countryside. Subsidizing LPG kept production costs lower for food sellers, and kept food costs low for poor workers. Over the years this system has largely fallen apart as the primary users of LPG have shifted. Now fleets of taxis in Bangkok use subsidized LPG, by some estimates, at a cost to the government of almost 500 USD a month – well over the monthly income of the taxi drivers themselves. This benefits poor sellers - certainly cheap LPG keeps taxi fares cheap generating sales for the poor taxi drivers, but if half of the effect of the subsidy goes to the customer, certainly it would be hard to argue that taxi customers are poor.

In addition cheap LPG as a fuel distorts the market by increasing the number of taxis on the road, increasing traffic congestion, pollution, and capital expenditure on automobiles. With an extremely low unemployment rate of less than one percent, there are other jobs that taxi drivers could be doing as well,

The issue of where the subsidy goes becomes even worse when we consider that a considerable proportion of the scarce government revenue goes to wealthy oil producers in other countries. This scarce government revenue would be better spent on more appropriate subsidies for the poor.

2.3.7.3 Energy Taxes

The Thai government taxes some kinds of energy, and uses the revenue to cross-subsidize other forms of energy use. The function of the tax is not to compensate for environmental damage or to raise revenue for the government; rather it is for income redistribution across energy users.

Table 14 Current energy prices and taxes

Oil Prices and Taxes Thailand April 2011	Price Baht	Includes Tax Baht	Price USD
Unleaded Gasoline 95	48 /liter	7.5/liter	1.60
Unleaded Gasoline 91	44 /liter	6.7/liter	1.47
Gasohol95 10% Ethanol	39 /liter	2.4/liter	1.30
Gasohol91 10% Ethanol	36 /liter	0.1/liter	1.20
Biodiesel 2% Palm	30 /liter	0.01/liter	1.00
LPG	18.13/kg	-12/kg	0.60
CNG	8.5/kg	-6/liter	0.28
Kerosene	36 /liter	0.13/liter	1.20

SOURCE: EPPO (2017)

2.3.7.4 Cross Price Subsidy

According to a recent IMF report (IMF, 2013), taxes on gasoline of 7 baht a liter (0.20 at 2008 prices) implied that more than 75% of the countries of the world charge a higher gasoline tax than in Thailand. Countries impose taxes on energy for a panoply of reasons, foremost among them for generating revenue, but also to compensate for negative externalities from carbon emissions, other pollution and congestion, and also as a source of research and infrastructure money for future energy needs. In Thailand, rather than acting as a source of revenue, in recent years the oil fund system has often been a net expense for the government. Considering the considerable costs that energy use places on the country, as well as the obligation that all sectors of the economy should contribute to the tax base for the social good, this seems less than appropriate. The logic of the cross-subsidy has been that gasoline users are generally better off than LPG users, so the tax system acts as a form of income transfer. Diesel was taxed at 5 baht a liter, less than gasoline, as it was considered to be less of a luxury product given its key role in the transport of goods.

The cross-substitution system works by taxing gasoline and diesel to underwrite the cost of subsidizing LPG. Generally this system has tended to be revenue neutral. However, the presence of large flows of money has been tempting and various governments have used the oil fund money to fund other populist campaigns.

One area in which this system has not worked well as a subsidy for the poor is for transport by motorcycle, the vehicle of necessity by the poor, since motorcycle engines all use gasoline.

2.3.8 Conservation and Environment

Conservation has not been a primary concern for recent governments. As a policy that emphasizes frugality it does not do well with voters. Although some incentives for conservation do exist, they are neither promoted nor observed. An earlier attempt to reduce the use of energy by closing malls early was an abject failure. As Thailand is a very hot country, malls, convenience shops, and restaurants all use air conditioning to draw people into their shops. A meal in an air conditioned restaurant may be twice as expensive as one outside. Part of the margin is built into selling air conditioning. Therefore it is very difficult to get Thai businesses to reduce its conspicuous use of air conditioning no matter how wasteful it may appear.

More hopeful would be to find ways to use electricity more carefully in a myriad of different ways. Conservation is probably best pursued by higher energy prices overall. Rather than promote particular plans, the use of higher electricity prices would encourage consumers to use electricity in efficient ways by their own motivation.

Although several campaigns to promote conservation have been evident in recent years, they have sponsored by the business community and focused on saving businesses money for items that the business community had been providing free. So, for instance, there was an active campaign not to use plastic shopping bags to help protect us from global warming, while no mention was made of conserving fuel or other petroleum products.

2.3.8.1 Specific Taxes

Specific taxes are one way a country can encourage conservation on goods that are particularly significant energy hogs. Several specific taxes were initiated during the Piyasvasti tenure, but some of these taxes, although still on the books, have had their tax rates reduced to zero. Air conditioners were taxed at 15% ad valorum, as were certain types of electronic equipment, and these goods now have zero tax rates. Automobiles still pay steep excise taxes - 30% on the typical small car - but those taxes predate the conservation efforts and were put in place for a variety of other reasons. More focus on conservation, if it can be done in a positive way, would be very useful to help prepare Thailand for an expensive energy future.

2.3.9 Poor-Friendly Investment

Enormous sums will be needed for energy infrastructure in coming years as the world adjusts from a complete reliance on fossil fuels to a more sustainable energy mix. New alternative energy infrastructure will be required. At present, the Thai government provides incentives for alternative energy projects, especially ethanol and palm oil. These incentives take the form of tax breaks and shared responsibility for infrastructure. Incentives are needed, because when oil is at 80 dollars a barrel, ethanol production is not very profitable, and construction costs are prohibitively high. It is most likely that this situation will change in the near future, and that these ethanol plants will be very profitable for a very long time. This raises the question of whether it is appropriate to use government money to fund them? Will Thai tax payers benefit from cheaper energy in the future? If they will not benefit then perhaps an alternative structure is needed so that energy will not be priced out of the reach of future generations of the poor. A related issue is about what instruments are available to control and tax energy producers in the future when the price of energy is high and the population must buy it at high prices from private suppliers.

2.3.9.1 Public Private Partnerships

Public Private Partnerships (PPP) offer one approach to supply much needed venture capital for energy infrastructure, while at the same time keeping some control and share of the eventual profits in the public sector. Ideally, the government would be a minority shareholder with the private sector managing the company for profit. Risks and profits would be shared between the private investor and the government.

2.3.9.2 Foreign Investment

Laws should also be in place to allow and in fact encourage foreign direct investment in the alternative energy sector. The primary purposes would be to bring technical knowledge about alternative energies to Thailand in addition to capital investment. Much of the research about alternative energies is taking place in developed and larger countries that have strong research institutions, and large capital resources. Foreign direct investment in energy is obstructed in the present time, since foreigners are not allowed to own land and energy is land intensive. One possible solution would be to allow PPPs between an outside firm and the government to own land jointly. Another approach would be to allow long term 99 year leases such as Vietnam does at present. If these solutions are not possible politically, it should at least be possible to facilitate investment in the sector to bring together cooperation between foreign expertise and Thai investors, which would allow the opportunity to be future energy producers for Thai persons.

2.4 Conclusion

Thailand is actively trying to shield its populations from the worst of the energy shocks resulting from volatile and expensive energy. However, the shocks are not temporary, but are likely to be with us for a long time to come. Long term strategies and policies need to be in place to help the poor survive future high energy prices, not simply to hide from them. Simply keeping prices low is not sufficient, although it may be attractive to a populist government trying to get reelected.
This summary has tried to give some background about the problems faced by poor who do not have access to energy, a critical assessment of energy policy in a middle income country, Thailand, and to discuss some of the ways in which policy can address and protect the long term interest of the poor.
This short overview cannot hope to be complete in addressing any of these topics, but several insights hopefully were brought up in the brief overview.
In particular,

- The poor are much affected by higher energy prices both directly, and through food prices.
- Subsidies for transportation fuel are not appropriate on a permanent basis. Any such subsidy will eventually lose its effectiveness as it distorts behavior towards overuse and over-substitution.
- Energy policy should not be in the hands of populist governments with short term goals, but rather a dedicated professional ministry. Another option would be the design of subsidy programs that are short-term by nature.
- Land tenure, land taxation and land policy will all be important in insuring that costs and benefits from expensive energy are shared.
- Improved Rail infrastructure is important.

- Subsidizing energy investment should be done is a way that gives long term benefits such as through Public Private Partnerships.
- Conservation is an important part of preparing for the future, including for the poor, and one tool is specific taxes on energy hogs.
- Subsidies should continue but need to be more narrowly focused on target groups with tight controls and assessments of eligibility in order to sustainable in the future.

Chapter Three (Paper 1) Energy Demand and the Optimal Level of Subsidies

3.1 Introduction

The focus of this paper is to describe some econometric tools that can be used to quantify and justify a certain level of subsidy for each of the primary energy needs of electricity, cooking fuel, and transportation. The research relies on a social justice framework that assumes that societies will endeavor to be somewhat equal in their distribution of benefits, and we should always try to help the poorest amongst us. Subsidies are generally designed to assist the poor or other needy group with basic necessities of living, acting as a form of social safety net. Typically subsidies are provided for goods that are considered to be necessities. However, the same good may be a necessity at a low quantity, yet act as a luxury at a high level. For example, this is a characteristic of electricity – required for light, refrigeration, and communication at a basic level, yet used for luxuries such as air conditioning, entertainment equipment at a high level. Therefor it may be appropriate to subsidize only to certain level.

3.2 Data and Methodology

Data for this study comes from the 2009-2013 Socio-Economic Survey (SES) performed annually by the Thai National Statistics Office (NSO). Price data came from several sources – the EPPO website for transport fuel and LPG Prices, and Electricity prices in a condensed form from the Board of Investments, summarized from circulars of (EGAT, 2017) (Thai BOI, 2017) (EPPO, 2017)
There were two primary data challenges for this paper. The first challenge was "getting the prices right" as the appropriate price is the retail price that households pay which can vary both by location and by quantity purchased. Electricity, in particular, is complicated as it uses 2-part pricing with a fixed and a variable component, and the variable cost per Kilowatt Hour (KwH) uses multiple block pricing where the first units are cheaper than later units.

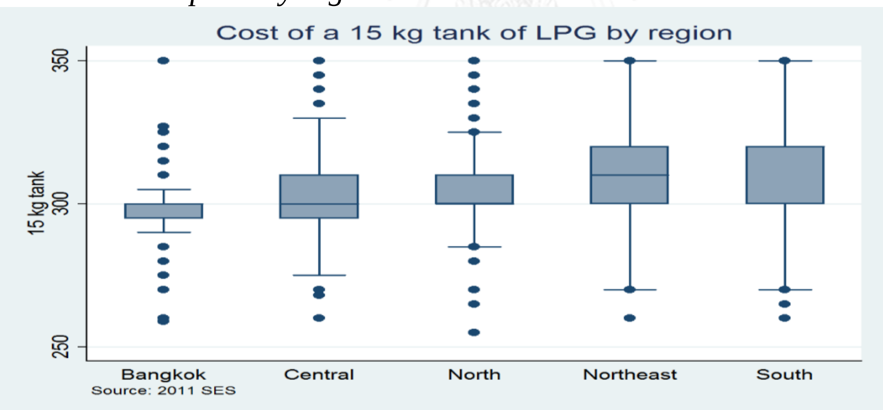
Table 15 Electricity Block Pricing in Thailand

Consumption	*Baht*	*US$*
First 15 kWh (0 – 15th)	1.8632	0.05
Next 10 kWh (16th – 25th)	2.5026	0.07
Next 10 kWh (26th – 35th)	2.7549	0.08
Next 65 kWh (36th – 100th)	3.1381	0.09
Next 50 kWh (101st – 150th)	3.2315	0.09
Next 250 kWh (151st – 400th)	3.7362	0.11
Over 400 kWh (401st – up)	3.9361	0.11

SOURCE: THAI BOI (2017)

Transport fuel prices are set weekly by the Petroleum Authority of Thailand (PTT, 2017) and pricing varies slightly by province calculated on a Sri Racha/Bangkok center plus a surcharge based on the cost of transporting the fuel to the point of sale. Monthly average fuel prices were collected from the EPPO website and transportation costs were added by province to obtain province level retail prices for transport fuel. LPG comes in standard 15 KG tanks. The basic subsidized price per KG is fixed, but the tanks are priced slightly differently in different regions. In this case, SES data was used directly to find the retail price for LPG by generating statistics on the prices paid for LPG per household that bought a single tank, which all clustered somewhere around 300 baht a month. In this case, median price per region was used.

Table 16 LPG price by region



SOURCE: SES 2011

The second data challenge was that data in the SES is all expenditure data, when what is needed is quantity. All expenditure data needed to be divided through by prices to find quantities of fuel used. Pricing for electricity was exceptionally difficult because of the complicated pricing structure, including a large proportion of poor households who receive free electricity. Free electricity is still "priced" in the SES, and as a rule of thumb, SES uses a price of 2 baht per KwH for free electricity (Anonymous SES Data Collector). A separate dofile was written to convert expenditure data from electricity back into quantities.

As in other related papers I have used household expenditure as proxy for household income.

3.3 Elasticities and Quantile Regressions

One of the workhouses of the applied economist is elasticities
Elasticities change relative to the income of the household. To try to capture that effect quantile regressions were used to show elasticities over income.
Elasticities are useful to study the effects of changes on demand. The economy is constantly changing and the Ministry of Energy has to adjust supply and demand in line with these changes. The commonly studied elasticities are price (demand) elasticity, income elasticity, and cross-price elasticity.

Price elasticities

If a business knows its price elasticity of demand, it can calculate a profit-maximizing or social welfare maximizing price.

Income Elasticities

If a business knows its income elasticity of demand it can forecast the quantity it needs to supply in the next period. The most commonly used measure of income is the widely forecast GDP.

Cross-Price Elasticities

The cross-price elasticity helps to define the market by showing what products are closely tied to their own. It also can be used to predict the effects of, for instance, the restaurant across the street raising their prices.
The Ministry of Energy or the Energy Policy & Planning Office can estimate the effect price changes due to tariff of tax changes will affect the demand for electricity. Generally energy has inelastic demand so it would be profit maximizing to increase prices. Therefore the price is generally set by the government, in either passing the price to a publically held utility, or mandating a price to a privately held but publically regulated utility. The objective function is to maximize social welfare given sufficient return to producers
Income elasticities are likely used to forecast residential supply requirements as GDP increases each year, and may be disaggregated by region or income group.
Cross-Price elasticities are not used but may be valuable in forecasting the demand for alternative vehicle fuels. Most other energy sources are difficult to use as substitutes.
Data requirements for elasticities include usage of a sufficiently large representative group of households along with the price they paid and their income levels. SES data provide household income and expenditure on energy, but not prices so that some sort of transformation process needs to take place to convert expenditure into quantities. Publically available price series are generally used, but block prices, and two part pricing may complicate the process.
If panel data is available it is best to look at the same households on time as it is possible to control for all of the unmeasurable variability in the sample. However, if the sample size is large,, e.g. greater than 10,000 households individual differences between households will tend to average out.
In many cases, only cross sectional data are available. This data may still be used for finding elasticities with the caveat that the elasticities may be slightly more elastic than with panel data, in the short run. In the long run they should be similar. Short run elasticities differ because households are resistant to change so may take time to

adjust to price and income fluctuations in the short run. Cross-sectional data captures long run effects as households have already adjusted to their current level of income. The most common way to fit elasticities is to take the natural logs of both sides of the equation. Taking the regression of the two sides will then give an approximation of the elasticity, which is the percent change in quantity due to a percent change in another variable. If a time series component is available (panel data again) it is slightly preferable to calculate their relationship directly by first finding the percent change in each variable. However for the normal values of elasticities near one, it shouldn't matter much.

Some sort of explanatory variables should also be included in the regression. These variables may include the number of family members, the number of working persons, whether the household is rural or urban, and other variables that are of interest.

Finally it is possible, in fact desirable, to calculate the income elasticity for different income groups. In the case of calculating subsidies this may be a critical part, although spline regressions, below, may do a better job of capturing marginal effects.

3.3.1 Elasticity Result Section

For each of the following energy products, the following approach is used. First an OLS regression is shown of the lnQ lnP lnI relationship. The coefficients are the respective elasticities. Several explanatory variables are included, notably

hhsize - household size or number of family members,

workers - number of working persons in family (energy may be used for consumption or for production – this is to account for persons driving to work in the morning etc.)

The constant term was not suppressed, although it should not have much meaning in a log-log regression. The regression therefore becomes

$$\ln Qelectric = \beta_0 + \beta_1 \ln Pelectric + \beta_2 \ln Income + + \beta_3 HHSize + \beta_4 Workors + \beta_5 Rural + \mu$$

After presenting OLS data for each energy good, quantile regressions are using to give more detail about the changes in elasticity over energy use or income/expenditure. Although it would be possible to calculate elasticities for a subsection of the data, e.g. the poorest quartile, in practice this gives poor results since 1) the sample size is smaller, thus giving less consistent results, and 2) it becomes unclear what to do with households that are on the edge of the break between quartile 1 and quartile 2.

Quantile regressions are useful for looking at the value of coefficients all along the distribution, so that, for instance, we can estimate elasticities for someone at the 63% of energy use or the 78%. In the figures below, first quantile regression results are reported for some critical benchmark values at q10 q25 q50 q75 q90. If we were to sort energy use from least to most, q10 would be 10% of the way along this continuum.

Quantile regressions do not discard any data points, but instead weight data points below the interval in question equal to data points above the interval. For example, the 10% of data points below q10 could have a weight of 1 each, while the 90% of data points above / to the right of q10 would have a weight of 1/9 each. The result is that points above the interval all count, but each counts for less than points below. If

we looked at q25, points below would count for 1 and points above for 1/3. If we looked at q50, points above and below count equally, making it almost the same as OLS. However, OLS is based on the mean relationship and q50 would be based on the median relationship.
The quantile regression looks like this

$$\min_{\beta \in \Re^p} \sum \rho_\tau(y_i - \xi(x_i, \beta)).$$

3.3.2 Electricity Elasticities

Figure 19 Electricity: OLS Regression

Source	SS	df	MS		
Model	6518.91013	4	1629.72753	Number of obs =	38222
Residual	7518.24117	38217	.196725048	F(4, 38217) =	8284.29
				Prob > F =	0.0000
				R-squared =	0.4644
				Adj R-squared =	0.4643
Total	14037.1513	38221	.367262796	Root MSE =	.44354

lnQelectric	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnIncome	.5638376	.0038724	145.60	0.000	.5562476	.5714276
hhsize	.0660085	.0019141	34.49	0.000	.0622568	.0697601
workers	-.0683576	.0028648	-23.86	0.000	-.0739727	-.0627425
rural	-.1508405	.0049455	-30.50	0.000	-.1605339	-.1411471
_cons	-.4641769	.0367681	-12.62	0.000	-.5362434	-.3921104

DATA SOURCE: SES 2013

The elasticity of income for the OLS equation is .56, a necessity. There are no units for elasticities. As household size increases, for each family member, the percent increase in electricity is 6.6%. Actually this is not very large, and implies that most electricity use is at the household level. The coefficient on workers is unexpectedly negative – supposedly it is because they are spending a lot of time outside of the household, and therefore not using electricity at home. Rural households use 15% less electricity than urban households. For electricity, unfortunately we do not have price elasticities due to a lack of variation in price in this year. We will now see how these numbers vary along the distribution using quantile regressions.

Figure 20Electricity: Quantile Regressions

Quantile Reg Electricity Income Elasticities (dependent=lnQelectric)

	q10	q25	q50	q75	q90
lnIncome	0.38***	0.45***	0.56***	0.64***	0.69***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
hhsize	0.06***	0.06***	0.06***	0.05***	0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
workers	-0.02***	-0.05***	-0.07***	-0.08***	-0.08***
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
rural	-0.09***	-0.10***	-0.14***	-0.17***	-0.19***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.66***	0.26***	-0.46***	-0.88***	-1.06***
	(0.08)	(0.06)	(0.06)	(0.06)	(0.09)
R2					
N	38222	38222	38222	38222	38222

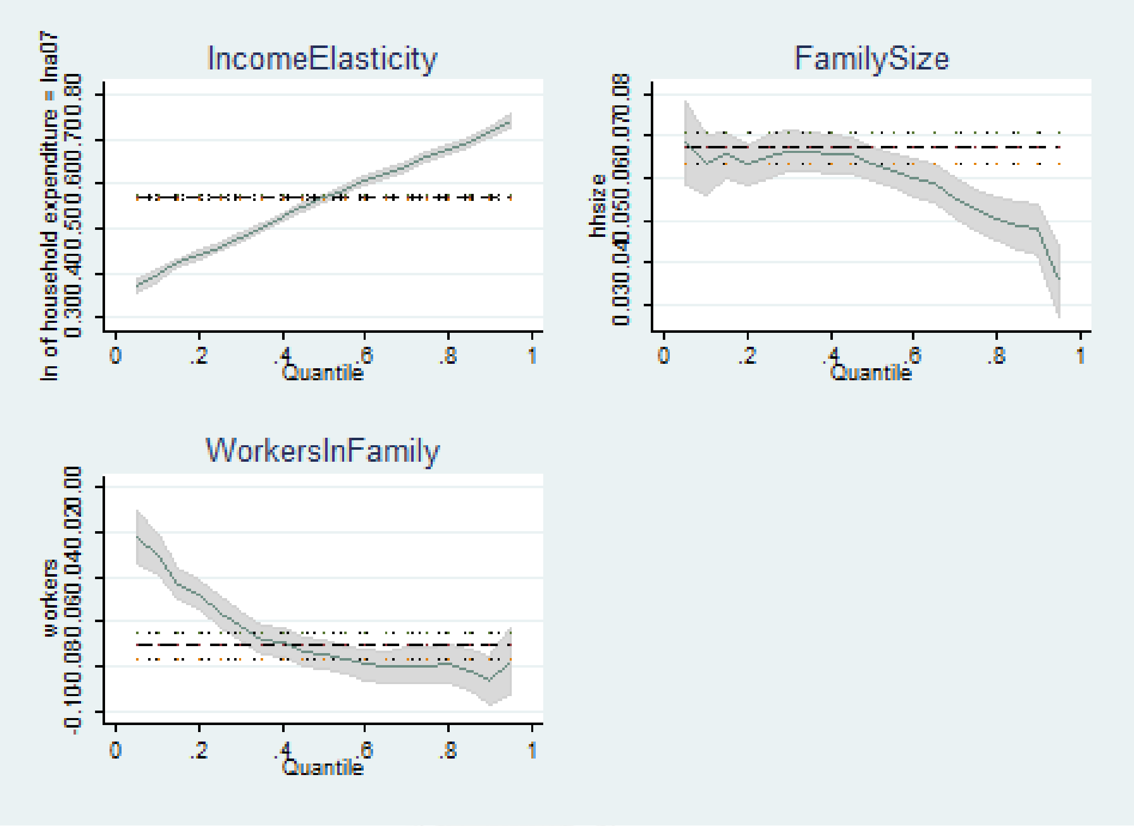
* p<0.05, ** p<0.01, *** p<0.001

DATA SOURCE: SES 2013

The first interesting feature is that demand for electricity becomes more elastic as households become wealthier and use more energy. Household size has a similar effect along the entire distribution with about a 5-6% increase per family member. Workers have a negative effect on electricity use which is increasing negative as we move up the distribution. Electricity use also is less in rural areas, and the effect is more pronounced at the top end of the distribution. Please note that that all the coefficients are significantly different than zero at the 0.001 significance level. Numbers in parentheses are standard errors.

We can get a better view of the entire distribution of the quantile regression by plotting values for the entire range of the distribution as shown below. In the diagrams below, the heavy dashed line shows the OLS value for this regression. The lighter dashed lines show the 5% confidence interval for the OLS regression. The quantile regression is shown with a solid blue line which tends to trend up, down or sideways as we move up the distribution. The grey area around that blue line is the 5% confidence interval along the curve. The OLS value will be similar to the q50 value in the middle of the distribution as in a quantile regression the middle point counts all data points equally. However, the values are likely to be a little bit different as OLS uses means and quantile regressions use median values.

Figure 21 Electricity: Quantile Regressions Graphical Representation



****

3.3.3 LPG Elasticities

Figure 22 LPG: OLS Regression

Source	SS	df	MS		
Model	40.2213052	5	8.04426105	Number of obs =	10140
Residual	895.003377	10134	.088316891	F(5, 10134) =	91.08
				Prob > F =	0.0000
				R-squared =	0.0430
				Adj R-squared =	0.0425
Total	935.224682	10139	.092240328	Root MSE =	.29718

lnQLPG	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnPLPG	.5142781	.0776818	6.62	0.000	.3620063	.6665499
lnIncome	.0200778	.0053137	3.78	0.000	.009662	.0304936
hhsize	.028893	.0023585	12.25	0.000	.0242698	.0335161
workers	-.0130934	.0035445	-3.69	0.000	-.0200413	-.0061454
rural	.0770032	.0064746	11.89	0.000	.0643118	.0896946
_cons	.7608607	.2435693	3.12	0.002	.2834165	1.238305

DATA SOURCE: SES 2013

The first question that is why is the price elasticity positive? Unfortunately this is probably an artifact of the data and the elasticity is probably useless. The price of LPG is subsidized and regulated and did not change over the period of the year.

Therefore the variation in price is a result of small changes in price in different regions. It happens that LPG is priced a bit higher in the south (many energy prices are set from Sri Racha / Bangkok with a surcharge for transportation cost). LPG is also used much more frequently in the South. It is possible that adding region as a variable to the regression would eliminate this anomaly, but since LPG did not change price during the year, we are unlikely to learn much. The next interesting result is that the income elasticity is close to zero. It appears that the cost of LPG is low relative to other expenses and even poorer households can buy as much as they need. Meanwhile there is little to use the LPG for besides cooking, and so as the family income goes up, there is little or no change in purchases. Household size and workers are similar to what we saw with electricity. Workers are outside the house during the day so use less LPG. Additional family members add a small amount to LPG purchases. R-squared is small suggesting that using LPG is determined by unmeasured factors.

Figure 23 LPG: Quantile Regressions

Quantile Reg LPG Price & Income Elasticities (dependent=lnQLPG)

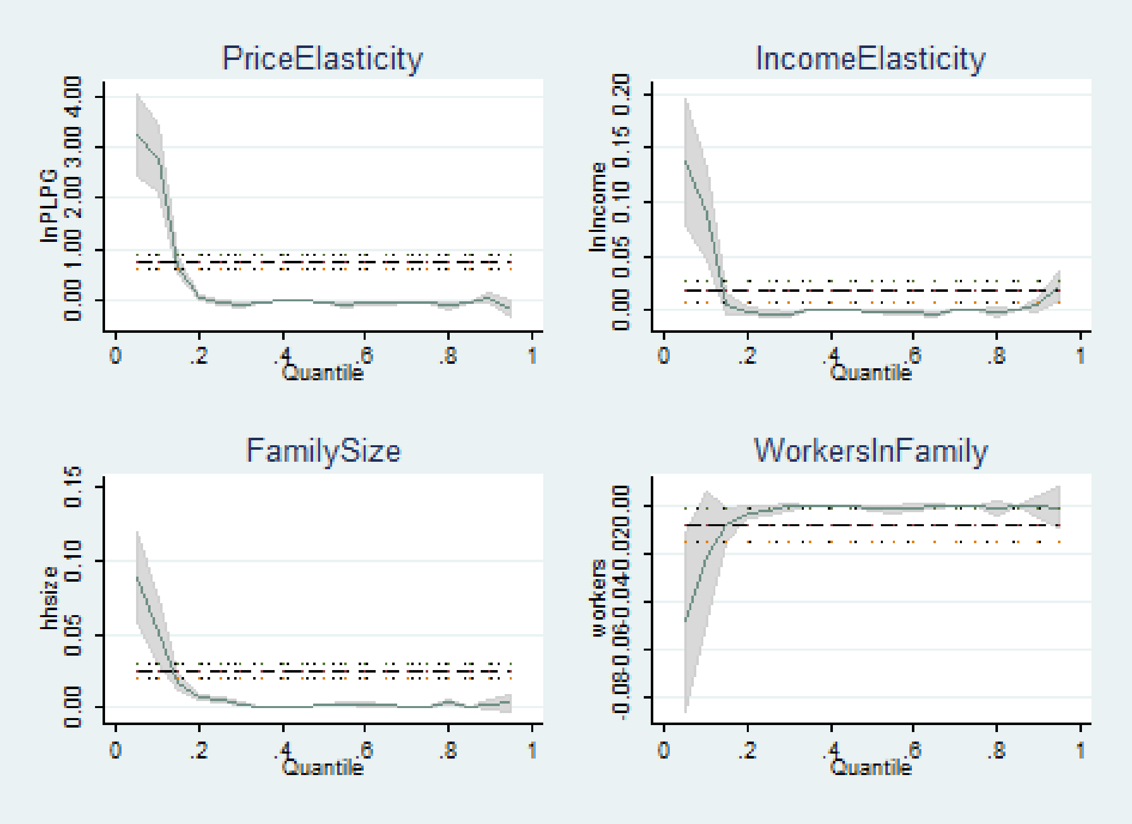
	q10	q25	q50	q75	q90
	none	none	none	none	none
lnPLPG	1.45***	-0.07*	-0.13***	-0.19***	-0.21***
	(0.28)	(0.04)	(0.02)	(0.03)	(0.05)
lnIncome	0.06***	-0.00	0.00	-0.00	-0.00
	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)
hhsize	0.04***	0.01***	0.00***	0.00***	0.00**
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
workers	-0.03*	-0.00	-0.00*	-0.00	0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
rural	0.52***	0.02***	0.01***	0.01***	0.02***
	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	-3.13***	2.89***	3.11***	3.31***	3.41***
	(0.89)	(0.11)	(0.05)	(0.10)	(0.16)
R2					
N	10140	10140	10140	10140	10140

* p<0.05, ** p<0.01, *** p<0.001

DATA SOURCE: SES 2013

The quantile regressions present several interesting results. The price data from OLS seems to have been driven by outliers near the bottom of the distribution, and for most quantiles, price elasticity is negative and quite inelastic as we would expect. Income elasticity is only significant for the poorest households. There may be an income level poor enough that it is hard to purchase LPG, but it must be quite low. Workers do not give us much useful information, but rural is positive and significant but small for all energy / income levels.

Figure 24 LPG: Quantile Regressions Graphical Representation



The data suggests that something is very different for the poorest / lowest LPG households. Because some of the results are unusual it would be best to repeat this exercise with another year's dataset. However, it is possible that the poorest households would switch to LPG if the price were lower or their income increased.
****

3.3.4 Benzene91 Elasticities

Figure 25 Benzene 91: OLS Regression

Source	SS	df	MS
Model	3284.44649	5	656.889299
Residual	12659.9185	28942	.437423763
Total	15944.365	28947	.550812348

Number of obs =	28948
F(5, 28942) =	1501.72
Prob > F =	0.0000
R-squared =	0.2060
Adj R-squared =	0.2059
Root MSE =	.66138

lnQben91	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
lnPben91	-.7846325	.0588774	-13.33	0.000	-.9000348	-.6692301
lnIncome	.4379173	.006697	65.39	0.000	.4247909	.4510438
hhsize	.0415487	.003224	12.89	0.000	.0352294	.0478679
workers	.0574572	.0048833	11.77	0.000	.0478857	.0670287
rural	-.0339226	.0089322	-3.80	0.000	-.0514301	-.016415
_cons	1.189538	.2292817	5.19	0.000	.7401351	1.63894

DATA SOURCE: SES 2012

The OLS regression shows a weakly inelastic price elasticity of benzene91. This is surprising as it would have been suspected that demand for transport fuel would be strongly price inelastic as people need to get to work, the market etc. This may be a result of substitution with other transportation fuels. The income elasticity describes a necessity as we would expect with fairly low increases in use as income increases. Demand increases by 4% for each additional household member and increases 6% for each working person in the house as we would expect as workers will need to travel to their work place. Incidentally, if we increase the number of workers by one, then that also means the number of household members increases by one, so this person increases benzene91 use by 4% + 6%=10%. Rural households use slightly less benezene91, although we will see that this effect will mostly see this effect disappear in the quantile regressions below.

Figure 26 Benzene 91: Quantile Regressions

Quantile Reg Benzene91 Price & Income Elasticities (dependent=lnQben91)

	q10	q25	q50	q75	q90
	none	none	none	none	none
lnPben91	-0.66***	-0.61***	-0.76***	-0.99***	-0.94***
	(0.15)	(0.10)	(0.09)	(0.11)	(0.12)
lnIncome	0.19***	0.24***	0.40***	0.62***	0.78***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
hhsize	0.06***	0.05***	0.05***	0.03***	0.01
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
workers	0.08***	0.07***	0.06***	0.05***	0.03*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
rural	-0.01	-0.01	-0.02*	-0.03*	-0.05***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	2.17***	1.97***	1.38***	0.72	-0.55
	(0.60)	(0.40)	(0.34)	(0.44)	(0.46)
R2					
N	28948	28948	28948	28948	28948

* p<0.05, ** p<0.01, *** p<0.001

DATA SOURCE: SES 2012

This quantile regression begins to show us the value of using a quantile regression approach because it allows different energy use / income levels to react differently. The most important point to note is that poorer households are often using transport fuel as a necessity to get to work or the market, while wealthier households are using it as a luxury to travel to relations or to vacation spots. Another difference, is poor households use motorcycles while wealthier households use cars.

In the data above we see that poorer households are more inelastic with respect to price than are wealthy households– they don't have a choice about using the fuel. They are also much less income elastic, so that increasing the income of the poorest households doesn't increase usage much, while wealthier households use a much greater share of additional income on fuel. Household size is more important for poorer household demand as it is not possible to put too many people on each motorcycle. The same is true for workers – each worker likely goes to work at a different location. The rural effect on demand is only significant and important for wealthy / high energy households.

Figure 27 Benzene 91: Quantile Regressions Graphical Representation

The graphical version of this data shows similar results, but with increased detail and with a better sense of the confidence we have in each coefficient as shown by the width of the confidence bands. One caution is in order – the dark dashed line is not where the coefficient is equal to zero. Rather it is the OLS value of the coefficient, so that we can see where the quantile estimate is above the average and where it is below. The reader must take care to study the vertical intercept to see what the value of the coefficients are .

****

3.3.5 Benzene95 Elasticities

Figure 28 Benzene95: OLS Regression

Source	SS	df	MS
Model	474.893026	5	94.9786052
Residual	1463.30138	2732	.535615439
Total	1938.19441	2737	.708145563

Number of obs	=	2738
F(5, 2732)	=	177.33
Prob > F	=	0.0000
R-squared	=	0.2450
Adj R-squared	=	0.2436
Root MSE	=	.73186

lnQben95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnPben95	-2.277033	.6110792	-3.73	0.000	-3.475257	-1.078809
lnIncome	.5069223	.0230795	21.96	0.000	.4616673	.5521773
hhsize	.0370251	.0113141	3.27	0.001	.0148402	.0592101
workers	.0590982	.017752	3.33	0.001	.0242895	.0939069
rural	-.1175399	.0317019	-3.71	0.000	-.179702	-.0553777
_cons	6.343541	2.370172	2.68	0.007	1.69603	10.99105

DATA SOURCE: SES 2013

Benzene95 is used almost exclusively in passenger cars and since only about 20% of Thai households own a car, that means the sample size is much smaller than for benzene91 and highly skewed towards wealthy families. The t-stats in this regression, although all significantly different from zero, are much smaller than in previous regressions because of the relatively small sample size of about 3000 households. Nevertheless the signs and size of all the coefficients are reasonable with the exception of the price elasticity which seems excessively elastic. R-squared is reasonable at 24% - there is still a great deal of variation in benzene95 use due to variations in life styles nationwide as we would expect, but there seems to be enough data to answer the questions we want to ask about elasticities.

Figure 29 Benzene95: Quantile Regressions

Quantile Reg Benzene95 Price & Income Elasticities (dependent=lnQben95)

	q10	q25	q50	q75	q90
	none	none	none	none	none
lnPben95	0.70	-1.49*	-3.83***	-2.67*	-1.70
	(1.26)	(0.71)	(1.06)	(1.20)	(1.24)
lnIncome	0.24***	0.24***	0.46***	0.79***	0.88***
	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)
hhsize	0.04	0.06***	0.06**	0.02	-0.00
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
workers	0.10**	0.08***	0.04	0.04	-0.00
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
rural	0.03	-0.04	-0.10*	-0.14**	-0.08
	(0.05)	(0.03)	(0.05)	(0.05)	(0.05)
Constant	-3.51	5.32	12.71**	5.77	1.66
	(4.80)	(2.75)	(4.11)	(4.66)	(4.79)
R2					
N	2738	2738	2738	2738	2738

* p<0.05, ** p<0.01, *** p<0.001

DATA SOURCE: SES 2013

When looking at the price elasticities, our concerns about the price elasticity do not diminish. Results are clearly being driven by the middle of the distribution. One possibility is that the high elasticity is driven by households switching between benzene91 and benzene95. As the price of all gasoline diminishes, households may feel that they can again afford to put benzene95 in their car.

The income elasticity duplicates the results we found for benzene91, with poor households acting as though benzene is a necessity, while wealthier households are moving towards it being almost a luxury. Number of workers has a positive coefficient as expected as people need to travel to work, but becomes insignificant for wealthier households. The rural variable does not seem to tell us much.

Figure 30 Benzene95: Quantile Regressions Graphical Representation

Again results seem to transition in ways that would be expected. The most dramatic transition here is the change in income elasticity across the distribution.
****

3.3.6 Gasohol91 Elasticities

(data from 2012 to avoid confusion of transition to only Gasohol in 2013)
In 2012, Gasohol was available at a price discount, but was not widely adopted as there was a story that it was bad for motorcycle engines. The government addressed this concern by simply eliminating the option and only selling gasohol starting in 2013 so that motorcycle drivers had no choice but to switch. The primary users of gasohol in 2012 may have been liberal and wealthy car owners. Gasohol is included for completeness in our study. The sample size of 5000 households seems adequate for good reliable statistics.

Figure 31 Gasohol91: OLS Regression

Source	SS	df	MS
Model	1418.75473	5	283.750947
Residual	2733.58708	5033	.543132739
Total	4152.34181	5038	.824204408

Number of obs =	5039
F(5, 5033) =	522.43
Prob > F =	0.0000
R-squared =	0.3417
Adj R-squared =	0.3410
Root MSE =	.73698

lnQgasohol91	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnPgasohol91	-1.259823	.2084855	-6.04	0.000	-1.668545	-.8511005
lnIncome	.7665699	.0169855	45.13	0.000	.733271	.7998688
hhsize	-.0265898	.0088361	-3.01	0.003	-.0439124	-.0092673
workers	.0022626	.0127694	0.18	0.859	-.0227709	.0272961
rural	-.0963861	.0218518	-4.41	0.000	-.1392251	-.0535471
_cons	.1915316	.769441	0.25	0.803	-1.316908	1.699971

DATA SOURCE: SES 2012

Demand for gasohol has a negative and elastic price elasticity. As an almost exact substitute for benzene91, having elastic demand is not too surprising since as the price increases, drivers can easily switch back to benzene. The income elasticity is fairly high for a fuel suggesting that as income goes up gasohol use goes up quickly. As most gasohol users were fairly well off, we would expect the OLS results to reflect typical fuel use patterns of wealthy people, i.e. fuel is almost a luxury product. Interestingly, smaller households seemed to use gasohol more than big ones.

Figure 32 Gasohol91: Quantile Regressions

Quantile Reg Gasohol91 Price & Income Elasticities (dependent=lnQgasohol91)

	q10	q25	q50	q75	q90
	none	none	none	none	none
lnPgasohol91	-0.86	-1.12**	-1.18***	-1.61***	-0.84**
	(0.54)	(0.43)	(0.35)	(0.38)	(0.31)
lnIncome	0.34***	0.54***	0.83***	0.96***	0.98***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
hhsize	0.00	-0.00	-0.03*	-0.04*	-0.04*
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
workers	0.06	0.01	0.00	-0.02	-0.03
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
rural	-0.15*	-0.06	-0.07	-0.10**	-0.09*
	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)
Constant	1.88	1.31	-0.72	0.15	-2.38*
	(1.89)	(1.56)	(1.28)	(1.43)	(1.17)
R2					
N	5039	5039	5039	5039	5039

* p<0.05, ** p<0.01, *** p<0.001

DATA SOURCE: SES 2012

The quantile regressions for gasohol help to clarify several issues. Price elasticities are only significant at the top end of the distribution where much of the use was. Income elasticities show the same pattern as with other fuels – poorer households treat fuel more as a necessity. Data for rural is mildly negative, and other variables doesn't seem to tell us much.

Figure 33 Gasohol91: Quantile Regressions Graphical Representation

The price elasticity has a very wide dispersion throughout the range of income /energy use and may not be too reliable. Only income elasticity changes in any large way across the distribution. Perhaps OLS would be enough in most cases for this fuel since the sample size may be inadequate to make quantile regressions effective.

****

3.3.7 Diesel Elasticities

Figure 34 Diesel: OLS Regression

Source	SS	df	MS
Model	1295.99509	5	259.199017
Residual	4086.52377	10662	.383279288
Total	5382.51886	10667	.504595374

Number of obs	=	10668
F(5, 10662)	=	676.27
Prob > F	=	0.0000
R-squared	=	0.2408
Adj R-squared	=	0.2404
Root MSE	=	.6191

lnQdiesel	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnPdiesel	-2.575778	.7170957	-3.59	0.000	-3.981419	-1.170137
lnIncome	.6097136	.0113248	53.84	0.000	.587515	.6319122
hhsize	-.0061028	.0049344	-1.24	0.216	-.0157751	.0035695
workers	.0075181	.0077567	0.97	0.332	-.0076864	.0227226
rural	-.000728	.0133232	-0.05	0.956	-.026844	.025388
_cons	6.60827	2.464125	2.68	0.007	1.778125	11.43841

DATA SOURCE: SES 2013

The final energy that we will look it is diesel. Again the price elasticity is negative as expected, but a lot larger than we would expect as diesel has fewer substitutes. (There may have been a few low speed diesel or biodiesel users but not many) The income elasticity is consistent with other transport fuels we have investigated. Data for other variables are insignificant so let's see what the quantile regression can tell us.

Figure 35 Diesel: Quantile Regressions

Quantile Reg Diesel Price & Income Elasticities (dependent=lnQdiesel)

	q10	q25	q50	q75	q90
	none	none	none	none	none
lnPdiesel	-4.88*	-3.24*	-3.49**	-1.03	2.03
	(2.07)	(1.52)	(1.29)	(1.21)	(1.17)
lnIncome	0.59***	0.61***	0.64***	0.61***	0.63***
	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)
hhsize	-0.00	0.00	-0.01	-0.01	-0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
workers	-0.01	-0.00	0.02	0.02	0.01
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
rural	-0.03	0.01	-0.01	-0.01	0.00
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	13.86	8.50	9.46*	1.79	-8.54*
	(7.19)	(5.24)	(4.46)	(4.17)	(4.03)
R2					
N	10668	10668	10668	10668	10668

* p<0.05, ** p<0.01, *** p<0.001

DATA SOURCE: SES 2013

Despite a large sample size of about 10,000 households, and a fairly high R-square of 24%, the regression does not give us many meaningful results. The income elasticity is about .6 and is significant and consistent across the distribution. It is to be imagined that diesel is used more as a production fuel (rather than consumption) relative to other fuel types, but it appears diesel use is not tied clearly to the number of workers in the family. It is to be recommended that other variables than those included here be used to help determine diesel use.

Figure 36 Diesel: Quantile Regressions Graphical Representation

The graphical presentation is helpful as it shows the high dispersion of results and broad confidence bands for all variables. Actually the values of the coefficients are quite consistent over the distribution so the dispersion is likely due to missing variables in the regression.

3.3.8 Conclusion and the Path Forward

This concludes our analysis of elasticities. We will return to look at predicting fuel use several more times. The Spline regressions presented next help us evaluate whether fuel use is at the household level or the per capita level, and again look at additional fuel use from a marginal perspective. Engel curves will help clarify the appropriate level of energy to subsidize if we want to help the poorest households in our country. Energy poverty lines and energy time series will help show us usage patterns of fuel over time. Data visualization will point out the clear relationship between altitude above sea level and both free electricity and LPG use. Finally big data techniques will illustrate an inductive approach to building predictive models of energy use.

3.4 Spline Regressions

3.4.1 Introduction to Spline Regressions

Spline regressions are a non-parametric technique that allows for flexibility in describing and modeling data. They work by setting “knots” in the regression line, and allowing separate parameters on each side of the knot. This is useful for describing marginal effects or behaviors, that other techniques such as elasticities do not supply. For instance, as household income grows, electricity might first be used for lights and refrigeration. Later the household might buy electrical equipment such as a washing machine and microwave. As income increases further the household begins to make regular use of air conditioning. There is no reason that the demand for electricity should follow a parametric trajectory. Nonetheless here is a requirement that there is a large amount of data so that results will not be affected by outliers. To some extent this can be controlled by reducing the number of knots to keep sample size large enough. Therefore, it is better to use traditional parametric regressions when the sample size is small as parametrics can forecast into regions with only sparse data available.

In our regressions, ten knots were chosen, which for electricity breaks the sample size into 43000/10 = approximately 4000 data points for each segment. For some other fuel types, the number of households using that fuel is less, so the number of data points per segment is less making the lines a bit unstable (For instance, note the little squiggle in the LPG regression which has only about 1000 households per segment). The ten knots reflect the household income deciles used elsewhere in the paper. Although spline regressions can contain quadratic terms and cubic terms, and often do when they are used for forecasting, in our model we only use linear splines. There are several reasons why it would ***not*** be wise to add higher order terms. First, although our data set is large, it is not huge. In order to use higher order terms it would be better to have much more data, probably a minimum of an order of magnitude more so that random variation will not play havoc with local parameter estimates. Second, the final objective of our exercise is to design energy pricing and subsidies, and it would be much more useful to have a block pricing model based on certain quantities usage, then to have a different price for everyone in the economy. In fact, we want to end up with some simple rules to help us set price rules in the real world.

3.4.2 Data

As before, the data we use is the Socio Economic survey for a recent year, in this case 2011. Because we are only using linear splines, data preparation for our model is very simple. Actually the following steps are done automatically by the program, but it is informative to see what is happening as it is quite simple and intuitive.

Let us suppose the knots are determined by 2000 baht intervals in household expenditure (Actually they are set in our paper by ten equal intervals of rank of household expenditure, but the below is shown for expositional clarity) Then spline regressions are generated by treating marginal expenditure for each new expenditure decile as a separate variable.

Table 17 Example of Data Preparation for Spline Regressions

HH Expenditure	expend1	margexp2	margexp3	margexp4	margexp5
4000	4000	0	0	0	0
6000	5805	195	0	0	0
8000	5805	1871	324	0	0
10000	5805	1871	1603	721	0
12000	5805	1871	1603	1650	1071

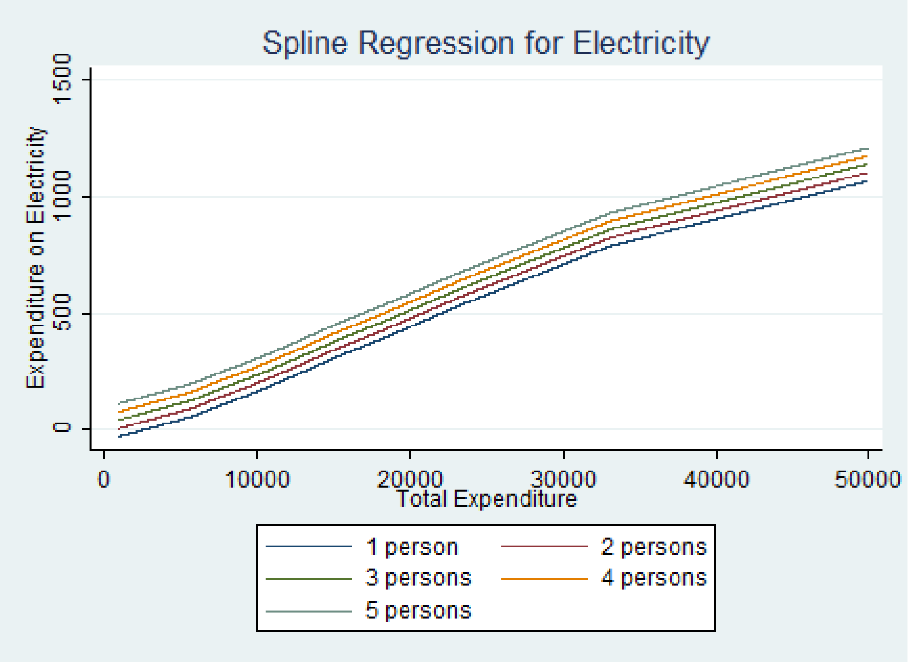
Equation 1

$$\mathbf{y} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} expend1 + \boldsymbol{\beta_2} margexp2 + \boldsymbol{\beta_3} margexp3 + \boldsymbol{\beta_4} margexp4 + \boldsymbol{\beta_5} margexp + \boldsymbol{\beta_6} margexp6 + \boldsymbol{\beta_7} margexp7 + \boldsymbol{\beta_8} margexp8 + \boldsymbol{\beta_9} margexp9 + \boldsymbol{\beta_{10}} margexp10 + \boldsymbol{\beta_{11}} familysize + \boldsymbol{\epsilon_i}$$

After the regression was completed, an additional technique was used to get the five lines shown in the diagrams below, which is handy for determining the extent to which energy is a household or per capita good. Family size was replaced by a new constant variable (either 1,2,3,4, or 5) and the **predict** command was used to run regressions on the new data. The resulting regression lines show the effect of the expenditure on fuel use depending on family size.

3.4.3 Results of Spline Regressions

Figure 37 Spline Regression for Electricity

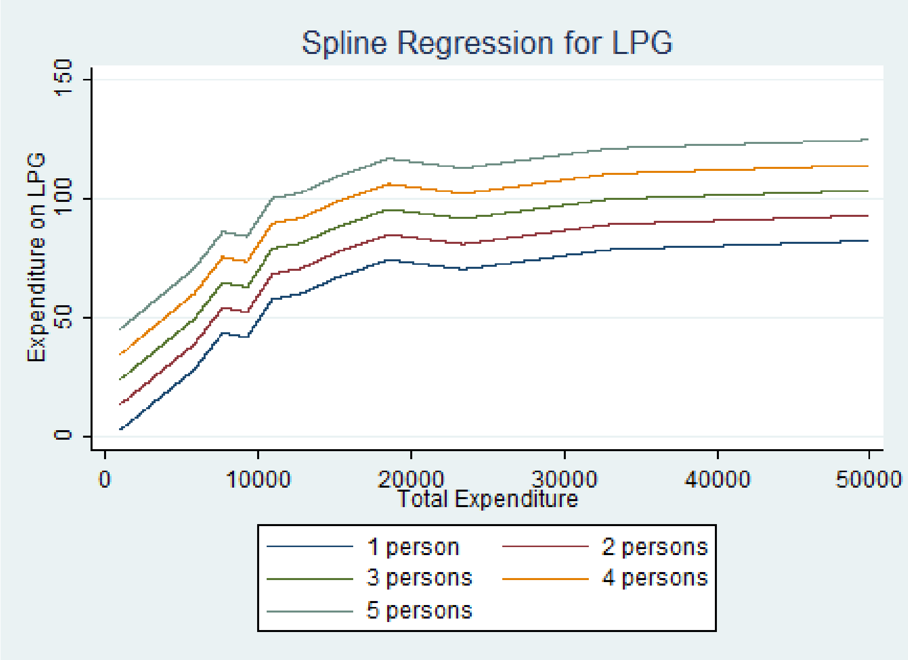


SOURCE: SES 2011 DATA FROM 2011 WAS USED TO AVOID CONFUSION RESULTING FROM THE END OF BENZENE 91 IN 2013.

Spline regressions have similarities to Engel curves, showing the effects of marginal income on marginal expenditure on energy. The vertical axis is the dependent variable and shows expenditure on energy. The horizontal axis is household income (expenditure is used as a more stable proxy). The elasticity of income can be calculated from the slope of the regression lines. The spread of the five lines show the effects of family size.

In the figure above, the spline regression for electricity shows that the demand for electricity continues to grow rapidly as income increases. Family size has only a moderate effect on household electricity use, suggesting subsidies at the household level are probably appropriate for electricity. It is not obvious where family's switch from one use of electricity to another suggesting a steady progression as income increases, is the best fit.

Figure 38 Spline Regression for LPG

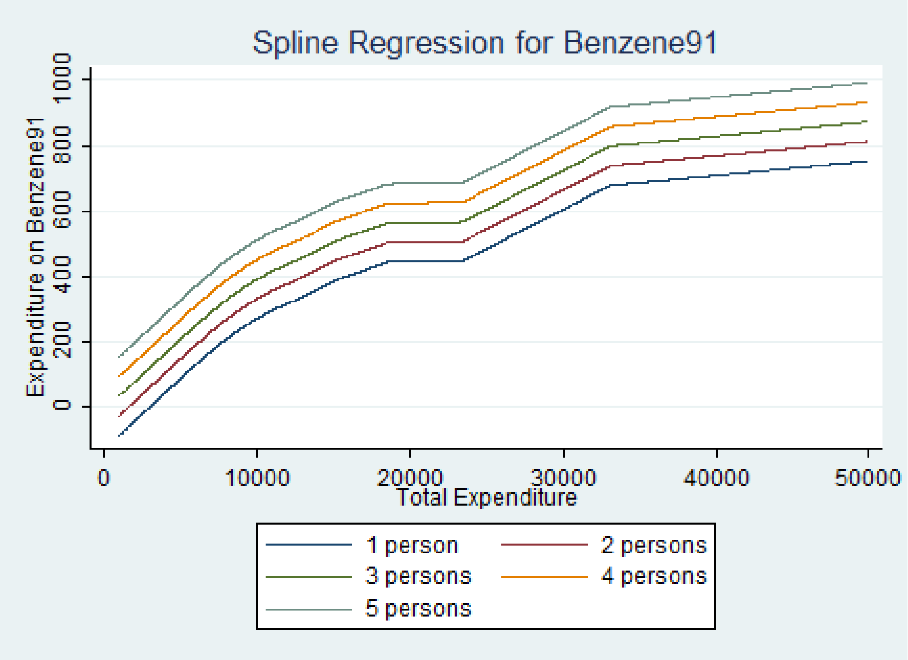


SOURCE: SES 2011

In the figure above, expenditure on LPG increases as family income increases until household expenditure reaches a level of 20,000 baht month (600 USD). Thereafter additional income does not lead to additional LPG use. The parallel lines show expenditure for a household of 1,2,3,4 or 5 persons respectively.

The spread of the five lines show family size accounts for a sizable share of the difference in spending on LPG. The absolute scale is also important though. LPG is not likely to be a large part of household expenditure for most households.

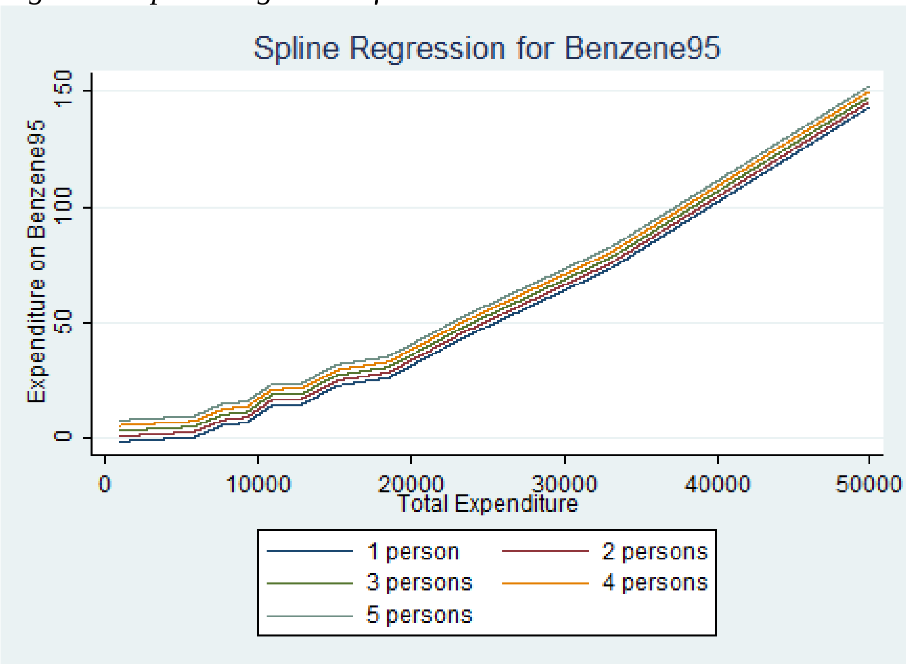
Figure 39 Spline Regression for Benzene91



SOURCE: SES 2011

The use of benzene 91 shows a very interesting pattern increasing rapidly over low levels of income, suggesting that income is limiting usage, then leveling off after a household income of 16000 baht suggesting households can buy all they need to about 25,000 baht, where demand jumps up again thereafter. What we may be seeing here is two separate uses for benzene 91, the first for motorcycles , and then after 25,000 baht households using benzene 91 for cars. In Thailand, about 20% of households own cars, and so it is suggestive that it is for the last two deciles that energy use jumps up. With a single value for household members, it seems that the size of the household is important in determining usage. That is likely to be true for motorcycles but not so much for cars. Benzene 91 at low levels could be a good candidate for a per capita subsidy.

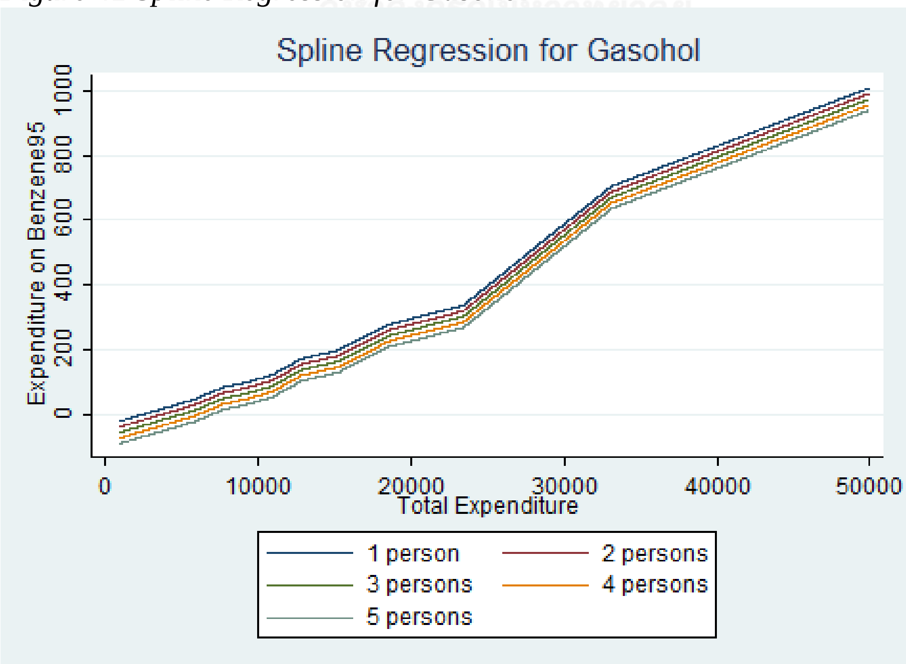
Figure 40 Spline Regression for Benzene95



SOURCE: SES 2011

Benzene 95 shows a very different pattern from benzene 91. Benzene 95 is used in passenger cars, and is not used in motorcycles. It is likely that most houses have only one car, so usage does not depend on the size of the family. Most interesting however, is how benzene 95 is clearly a luxury product, with energy usage increasing rapidly as income rises and families take more holidays and trips. This is a product that clearly should NOT be subsidized.

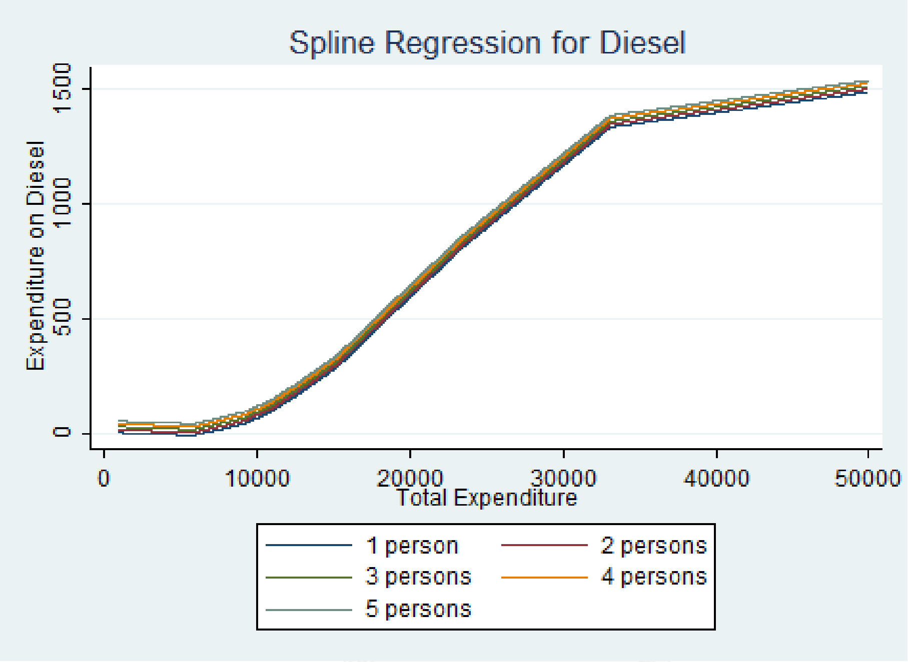
Figure 41 Spline Regression for Gasohol



SOURCE: SES 2011

The picture for Gasohol shows much the same pattern as the picture for Benzene 91 and for very much the same reasons. Poorer families use gasohol for motorcycles. Usage rises rapidly with income suggesting that income is limiting, then levels off. Then usage jumps again, presumably showing a switch to use of the fuel for passenger cars.

Figure 42 Spline Regression for Diesel



SOURCE: SES 2011

Our last Spline regression for diesel shows very low usage for poor families as should be expected, as it is used primarily in pickup trucks which the poor could not afford. Then usage increases very rapidly with income. This is likely not caused by being a luxury good, but because diesel is most often used in production and transport of goods. Usage is by household as each house is most likely to have one truck. Income is limiting in diesel to a high level of family income, in our picture around 35,000 baht of household income. However it would probably be better to have a higher number of knots at the high end to show more detail at this higher income level. Subsidies for production are a possibility, although diesel users are middle class or wealthy. Sometimes production subsidies are justified by claiming that the poor use goods that are transported by truck. However, this justification is much less likely to have weight for poor families than for others as they tend to live a more subsistence lifestyle and to buy locally.

3.5 Engel Curves

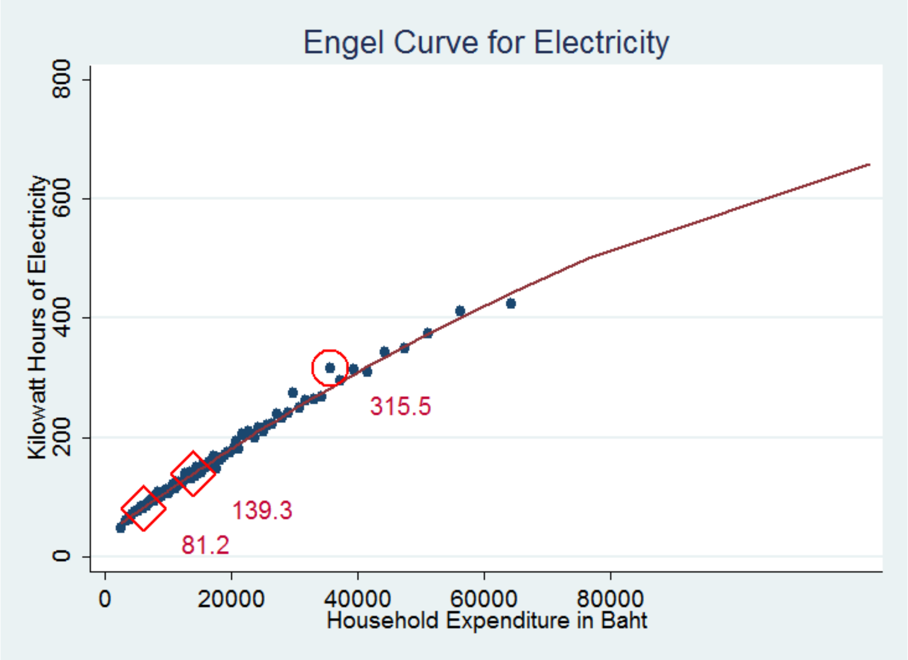
Engel curves are a popular tool often used by microeconomists to relate income to the quantity purchased of a good. They are a taught as theoretical construct which helps differentiate between luxury goods, normal goods and inferior goods.

Energy would generally be construed to be a necessity. Most subsidies are applied to necessities as the poor need to purchase them, but in limited quantities.
Our purpose in using Engel curves is more specific, to set a quantity at which we should subsidize energy. If energy is a necessity, then at least some people must use it for light, for refrigeration, for transport. We can use typical usage levels for different income groups as targets.

3.5.1 Towards Optimal Subsidy Levels

It is often difficult to determine the level of energy to subsidize. This paper takes the approach of adopting both a philosophy and a tool. The philosophy postulates that there will always be poor people, at least relative to the rest of society, and it is our objective to always help the bottom 10% to achieve the standards and living conditions of the rest of society. Thus we never solve the problem of poverty, but we always reach down to the ones on the bottom to pull them up, so that as a society we can all grow together.
In practice, this means we want to assist the bottom 10% of households in terms of income to achieve energy use similar to those immediately above them in terms of income.
The tool we use for this purpose is based on the familiar Engel curve, which plots income versus energy use. In our version, the sample is divided into one hundred income group centiles, with each centile plotted on a graph as below.
Engel curves are used to study the effect of changes in income on the purchase of goods. In the charts and regressions displayed below, household expenditure continues to be used as a proxy for income. Each dot represents 1% of the population. The horizontal axis is in baht/monthly expenditure. The vertical axis is measured in quantity of energy used by each centile. Expenditure data on energy have been converted to quantities by using data from (EPPO, 2017) and (MEA, 2017). Details available on request.
The graphical Engel curves were created by dividing the sample households into income centiles and calculating energy use for each centile. The slope of the Engel curve describes whether the good is a luxury (> 1), necessity (0 < elasticity < 1) or inferior good (< 0) . Each dot represents a separate income centile, while the circle at the top represents the 90th centile and the two diamonds represent 10% and 50% expenditure deciles respectively. Data was included only for those households who used that energy source, excluding zero values.

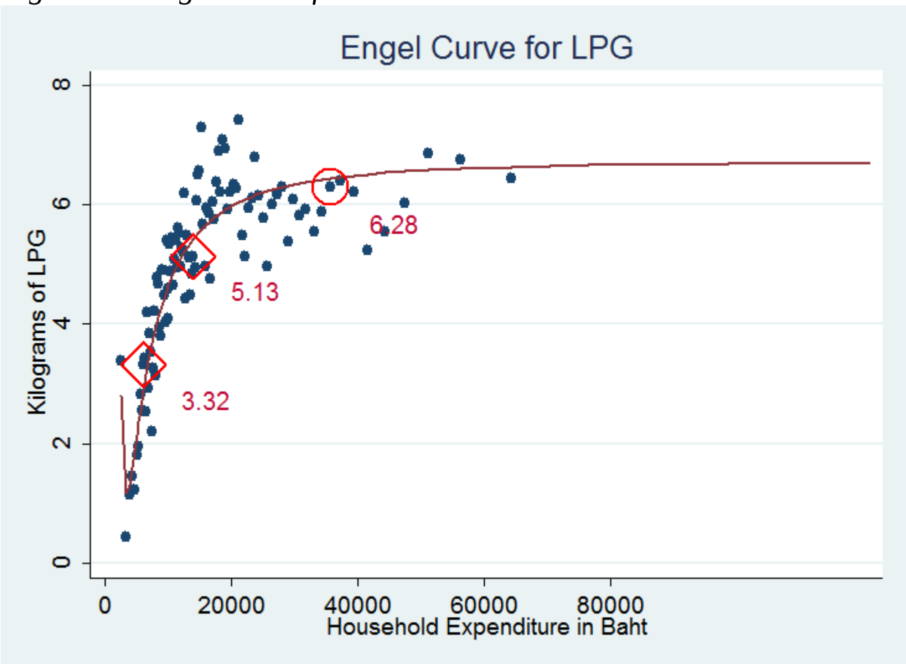
Figure 43 Engel Curve for Electricity



NOTE: OPEN DIAMONDS ◊ DEMARCATE THE 10% AND 50% INCOME LEVEL, WHILE AN OPEN CIRCLE ○ DEMARCATES THE 90% INCOME LEVEL. ONE APPROACH TO SUBSIDIES IS TO HELP THOSE BELOW THE 10% INCOME LEVEL TO USE ENERGY AT A TARGET LEVEL BETWEEN THE OPEN DIAMONDS. FOR INSTANCE, IN THE ELECTRICITY ENGEL CURVE ABOVE, WE WOULD TRY TO ENSURE THAT EVEN THE POOREST HOUSEHOLDS COULD USE BETWEEN 81.2 TO 139.3 KWH OF ELECTRICITY.
SOURCE: SES 2013

One practical approach to setting subsidies is to set a goal that the poorest 10% of households should ideally be able to consume at similar levels to others in the 10%-50% income level of society. In our Engel curves, we can use the range between the diamonds as the target level and the dots lower than the lower diamond (lower than 10% expenditure level) as the targeted group.   The implication for electricity is that we would like everyone in the country to be able to use between 81.2 and 139.2 kilowatts of electricity in each month. In fact current policy is accomplishing that admirably due to the subsidy program in place. The lowest dot on the Engel curve (median for lowest 1% in terms of expenditure) for electricity is still about 67 KwH. However, the implication is that 81.2 KwH would be a good cutoff point for free electricity.

Figure 44 Engel Curve for LPG



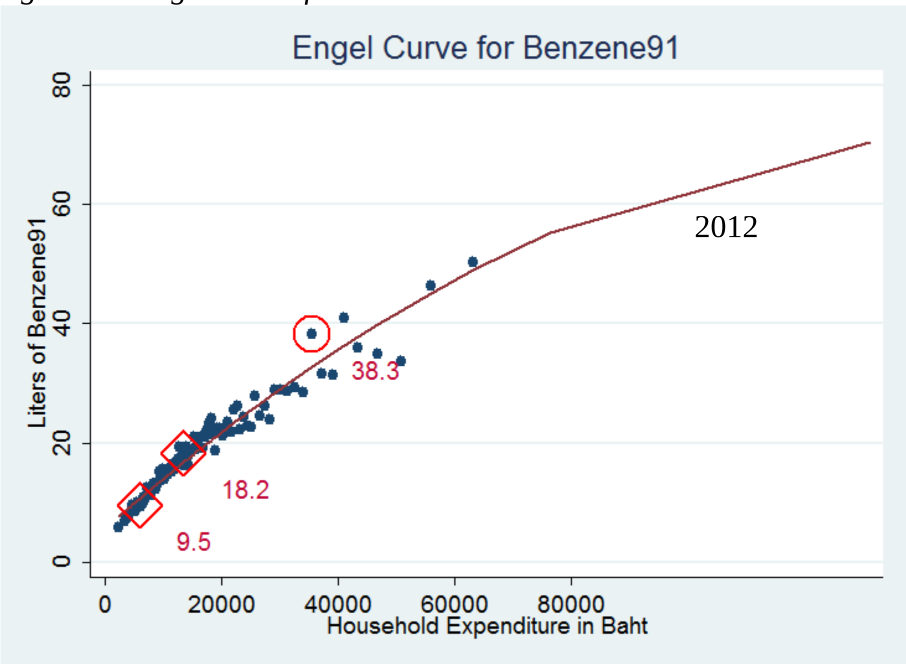
SOURCE: SES 2013

The Engel curve for LPG suggests that poor households use much less LPG. Ideally we would like to allow all households to be able to use between 3.32 and 5.13 kilograms of LPG a month, and the poorest households do not achieve this. Once income reaches about 20,000 baht a month, usage is constant in income. More income presumably does not involve more cooking. One way that has worked in Indonesia is to sell subsidized gas only in 3 KG tanks. Wealthier households are willing to pay a higher price to avoid the hassle of changing tanks very often, while poorer households use less and are more price sensitive so they buy the small tanks.[3] The majority of poor households do not use cooking gas at, exposing the family to smoke and carcinogens. The policy implication is that there is a need to encourage LPG use to improve quality of life and health. Supplying free stoves may be better than energy subsidies. In Indonesia free stoves and an original 3 kg tank of gas were given away free to the poor since a switch resulted in a reduction in lung disease from smoke particles. The original stove and tank full of gas were given to the poor directly to promote switching and to avoid rather expensive startup costs. (IISD, 2013), p.10.

---

3 The data used here are for families who reported using LPG as their primary cooking fuel. Families do not buy a 15 KG tank of gas every month, so for this graph only, if they report using LPG as their primary cooking fuel, but did not buy any gas, a value of zero was used in the average.

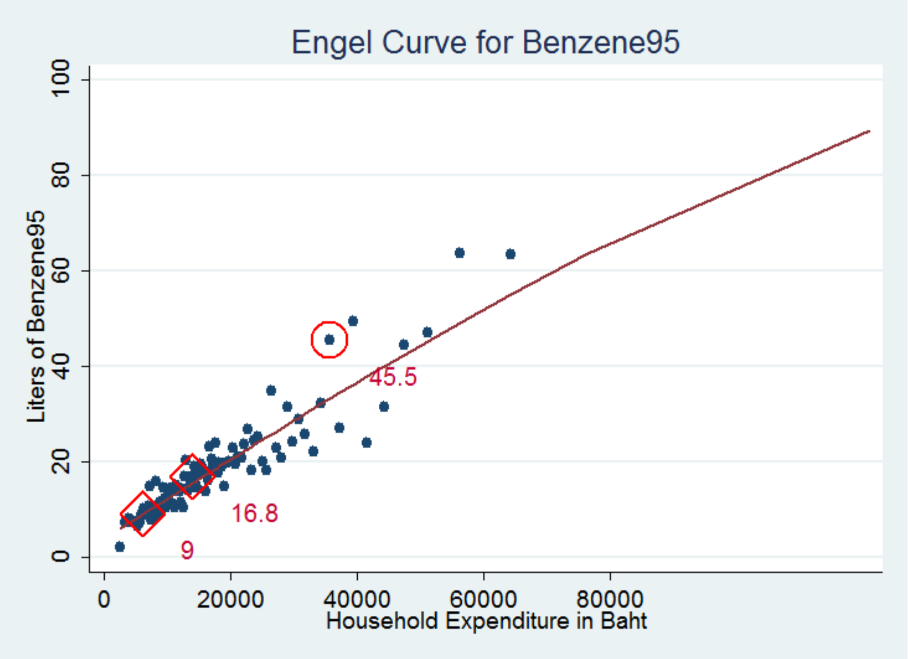
Figure 45 Engel Curve for Beneze91



SOURCE: SES 2012 USES DATA FROM 2012 SINCE BENZENE 91 WAS NO LONGER FOR SALE IN 2013.

The Engel curve for Benzene 91 shows a cutoff for the poorest decile of 9.5 liters a month. The poorest half of households have a range of 9.5 to 18.2 liters per month. If we were to subsidize Benzene 91 it should probably be no more than 10 liters a month per household. The Engel curve shows a typical shape for a necessity, but does not flatten out as much as for LPG as cars can use much more benzene than motorcycles. From our earlier work, we saw that benzene 91 depends on the number of persons in the household, so that may a concern in the design of a potential subsidy. Data is for 2012 as Benzene 91 was no longer for sale in 2013.

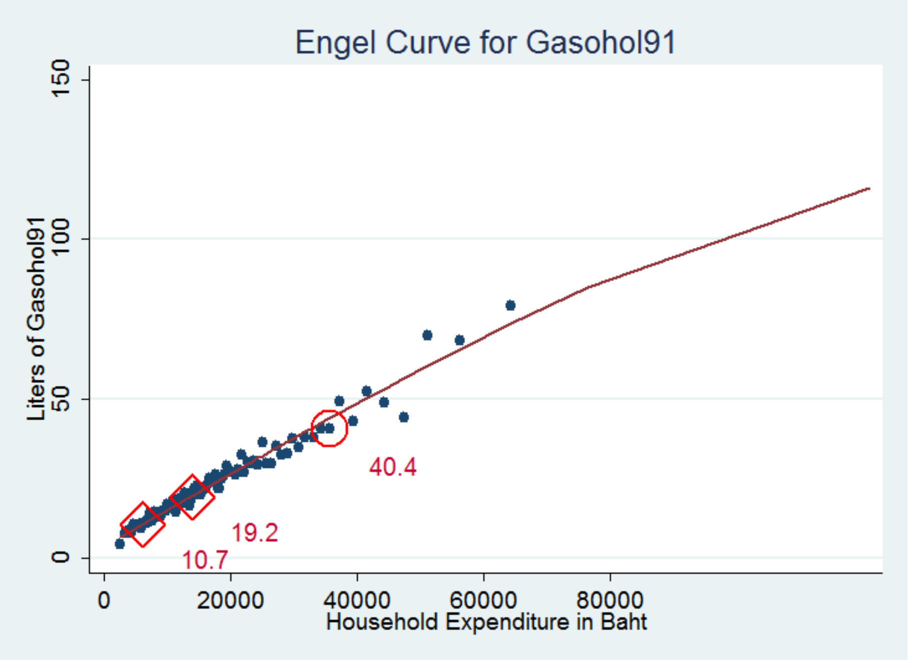
Figure 46 Engel Curve for Benzene 95



Source: SES 2013

Benzene 95 shows a steadily increasing demand across the measured income range. At least in Thailand, we are not near satisfying demand at current price levels. The cutoff for the bottom decile is 9 liters which is very close to what we see for Benzene 91 supporting a fuel subsidy of 10 liters per month. However, almost no poor households purchased Benzene 95 so we would be very unlikely to subsidize if our objective is redistributive.

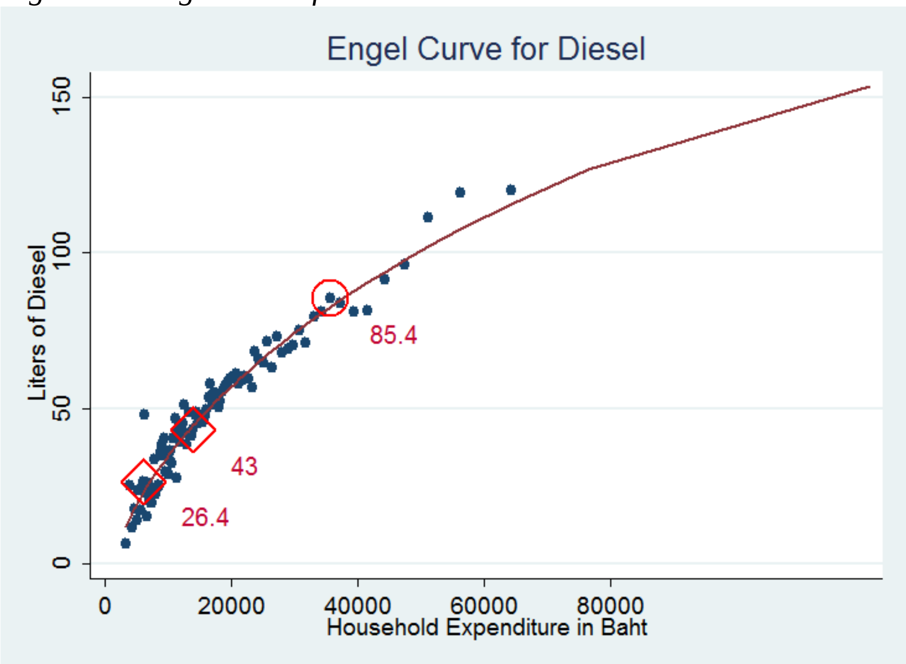
Figure 47 Engel Curve for Gasohol 91



Source: SES 2013

Engel curves are included for both gasohol 91 in 2013 and for benzene 91 a year earlier in 2012. Given the popular idea that gasohol would damage motorcycle engines, it was interesting to see if the switch affected people's behavior. In fact, about 5% of users did switch to benzene 95 in the first months after the government forbid the sale of benzene 91. But the most striking thing about the gasohol benzene Engel curve is that essentially nothing changed implying that the transition went smoothly. The cutoff is slightly higher at 10.7 liters, but the range of the bottom half the population does not change much at 10.7 liters to 19.2 liters. Again this lends support to the argument that the government could subsidy transport transportation fuel, but no more than 10 liters per household per month.

Figure 48 Engel Curve for Diesel



Source: SES 2013

Diesel is a more specialized fuels used by a small share of households. It should be reemphasized that these Engel curves are only for those households who use the fuel. Households that do not use any are not averaged in. So the Engel curve for diesel looks at only households who buy diesel.

Poorer households buy less diesel. Diesel, however, shows an elastic income elasticity. As income increases the use of diesel increases rapidly One of the first production tools that households may purchase as they reach a middle income is to buy a pickup truck for farm or business. This supports the argument presented elsewhere, that the government might want to subsidize the purchase of pickup trucks "First Pickup" policy. However, subsidizing diesel would be expensive and wealthy households use much more than poor households. Those poor households that do have a truck, are already using 26.4 liters. Subsidizing diesel is mainly subsidizing the transport sector but is unlikely to be needed for production.

3.6 Conclusions

3.6.1 What Should we Subsidize

LPG and electricity are important in terms of quality of life – even the poor should be able to eat refrigerated food, watch television, have lights at night, and use fuel that will not destroy their lungs by breathing air that is polluted with smoke particles. Benzene 91, (now Gasohol 91) may be necessary for production as well as consumption. Workers need to go to their place of work, school children may need a ride to school, markets may be located far from households.
In Thailand, and especially for the poor, this means using a motorcycle. Outside of Bangkok, more than 80% of households have a motorcycle. A motorcycle may be needed for each working person.
Using 10% and 50% expenditure levels as a target, then target levels of subsidies should ensure LPG use between 3 and 5 kilograms per household per month, electricity of more than 80 KwH a month and the use of between 10 and 20 liters of benzene 91 or gasohol 91 per month.

Chapter Four (Paper 2) Evaluating Energy Subsidies Over Time

4.1 Introduction

Subsidies are generally designed to assist the poor or other needy group with basic necessities of living, acting as a form of social safety net. Typically subsidies are provided for goods that are considered to be necessities. One does not provide a subsidy for chocolate cake, but a subsidy for rice is possible. One does not provide a subsidy for an amusement park ride, but a bus ride to take the person to work is possible. Most often subsidies of this sort are provided for health, for transportation, for food and for energy.

Although the motivation for subsidies is usually clear, it is difficult to judge their effectiveness. Governments often undertake programs to help the poor, only to discover years later that the subsidy has benefited another group altogether.

One approach to the question of who should receive subsidies comes from the literature on poverty lines. This paper will investigate this direction in some detail, However, the paper will also make use of a number of other tools to determine who should receive subsidies.

Many countries, Thailand included, are experimenting with a broad platform of energy policies to meet objectives such as promoting renewable energy, ensuring energy security or protecting the poor from high prices. What are lacking are adequate measures to evaluate these policies. Do the benefits outweigh the costs? Are the intended beneficiaries of certain policies the ones who benefit? An energy poverty line and a median index would add to the tools available to policy makers.

The focus of this paper is to track the effects of subsidies over time – with particular focus on three issues. Section 4.2 looks at the issue of energy poverty lines and considers the criteria that should be used to determine whether a subsidy is needed for a particular group. Several different approaches are considered, with the priority being subsidies that benefit the poorest groups the most. Section 4.3 uses a non-parametric approach to time series to track the use of expenditure items over time. The technique can be used to track expenditure for any product, but in this paper, the spotlight will be on energy policies, and who they reach. For this part of the analysis, groups are divided by income (expenditure) deciles and by month. Section 4.4 uses annual data to calculate the share of final expenditure on a subsidy that goes to each income group. The analysis is also extended to uptake, to look at the share of each income group that receives at least some benefit from a subsidy program. We would prefer that the poorest groups receive most of the budget allocated for the subsidy. We would prefer that a high percentage in the poorest group benefit from the subsidy in some way. Section 4.5 includes a brief conclusion section.

4.2 Energy Poverty Lines

4.2.1 Introduction

This section discusses the formulation and design of an energy poverty line and a median energy index, to be used to measure the effectiveness and distributional effects of government energy policies.

To be useful as a policy tool an index must be 1) stable, 2) simple, and 3) transportable between countries or regions. Strengths and weaknesses of various formulations are discussed, and generally the author comes down in favor of a pair of indices (poverty and median) relative to a base year. One candidate is a poverty line based on the average energy use of a 10th decile in terms of expenditure/income. Households below this level would try to be assisted to the extent that they could keep up with households a little better off.

Per capita energy expenditure data come from the Thai social economic survey, and are used together with retail energy prices. Several contending definitions of the energy poverty line are evaluated in terms of practical ease of use, stability and transportability. These three ways are 1) Minimum energy bundles per capita, 2) Share of income and expenditure spent on energy (in my paper disaggregated), and 3) current energy use of the lowest decile.

1) Minimum bundles per capita are identified in three ways – a) Engel curves for major energy types are estimated, with a target range bracketed by energy use of the 10th and 50th percentile determining the minimum quantity, b) regressing household energy use by family members for different deciles points to a minimum energy requirement per household, c) income elasticities are calculated for each income class, to study the effect of income and effects by group.

2) Share of income and expenditure spent on each basic need (cooking, transport, light, work, heat) are calculated separately to allow an additive approach to allow the measure to be used in different countries. A 10 percent level was used as a cut off.

3) Energy patterns of the bottom 10 percent of households in a base period, were considered, but rejected because of instability due to the wide range of fuel types and high in-kind payments of the poorest groups.

In addition,

4) Access to modern energy implies the ability to afford to maintain certain goods such as LPG stoves, refrigerators and motorcycles. Survey data was used to related ownership of these goods to energy use using a simple OLS estimation.

Conclusions from the paper are as follows. Indices based on quantities are preferable as they are more stable. A quantity based energy poverty line is feasible because energy prices are standardized and there are just a few types of fuel. This is in juxtaposition to the general poverty line.[4] A median and a poverty index used together are more useful than just an energy poverty line. These tools are highly effective for evaluating the distributional effects of energy policy. Energy use is

---

[4] For the general population the wide variety of goods makes quantity indices practical. The exception are indices based on sufficient calories for the poorest countries.

highly skewed, so that subsidizing the prices of many types of fuel are not cost effective, and should be replaced with targeted programs where possible.

Over the past two decades high energy prices and a growing awareness of the rapidly diminishing supply of crude oil and eventually other energy sources have brought energy policy from simply an important topic to one of the most vital and politically sensitive topics in most nations. The fertile combination of political impetus, developments in microeconomics over the past 30 years, and strong interest by the populace due to high energy prices, have led to an amazingly broad and comprehensive range of policies, mostly directed at 1) promoting alternative energy, 2) policies linked to the environment and global warming, 3) energy security, 4) energy conservation policies, 5) supplying basic needs to the poorest, 6) energy investment policy, 7) energy integration, and 8) keeping prices low for at risk segments of the population.

This paper is relevant for most of these but focuses especially on the last of these, as it is especially relevant to poor and middle income countries. Higher energy prices and scarce energy will have a disproportionate effect on the poor. As energy prices are driven higher, those who are better off use less energy, but those who are poor are forced away from modern energy entirely. An energy poverty line can help focus on this issue. In addition, protecting energy prices due to idealism or political pressures threatens to bankrupt government. These polices need to be rationalized.

The specific purpose of this section is to search for an energy poverty line for Thailand. Over the past 15 years, the Thai government has promulgated perhaps 20 different energy polices. There is a debate in the press and among academics about how effective these policies are, who is affected, and whether the benefit is worth the cost. An energy poverty line provides one objective way to evaluate the effectiveness of policies. Does policy A reduce those below the poverty line, and if so, by how much? Is policy A more effective than policy B, and so forth.

In our context, an energy poverty line or energy index does not have to evaluate only the poorest of the poor, it could be set at the 30% level or the 50% of the population, to measure access to energy more broadly, and to study the distributional effects of energy policy. These options are explored broadly in the paper.

One circumstance that improves the ability of an energy poverty line to serve as a policy tool, is the recent change to the social economic survey so that it is now collected continuously. Although the survey is published once a year, survey data is collected continuously in all regions in every month, so that some monthly level evaluation is possible. Before 2006, data was collected continuously during the survey year, but then not collected in alternate years. Continuous data is useful for evaluating energy policy, since many energy policies have been short term experiments, abruptly initiated or terminated due to politics, popular pressure, or high global energy prices.

An energy poverty line is a fairly new concept in the literature, Therefore, conceptually there are many issues to be worked out. There are currently several candidate concepts of how to establish a poverty line, and new ideas are welcome. Attempting a poverty line with Thai data would be very useful to this debate and gives value to this current paper. Energy poverty lines have been attempted in South Africa, Bangladesh and the UK, among others.

4.2.2 Objectives

This study of energy poverty lines has the overall goal of designing an energy poverty line for Thailand.
There are three distinct objectives:
Objective I employs a variety of techniques to search for a minimum bundle of energy specific to Thailand.
Objective II discusses alternative designs for an energy poverty line and a median poverty index.
Objective III discusses evaluation of government energy policies.
It is useful to have some formal criteria to evaluate these candidate poverty index measures, which leads to the following three propositions:
Proposition 1) Any energy poverty line must be stable to outside factors.
To be used as a meter stick, the poverty line must be stable, robust to changes in factors other than energy.
For instance, using a maximum of 10 percent of total expenditure for energy products, a popular definition of energy poverty, would not be robust to overall changes in income.
Proposition 2) An energy poverty line must be simple and easy to use.
Most tools that are complicated are not useful in the real world as they confuse politicians and policy makers and occasionally befuddle even the economists they were designed for.
This could lead to a measure that is not as precise as theoretically possible, but that would receive broader adoption. If such a decision arises, the broader adoption would be chosen, within reason.
Proposition 3) Any energy poverty line should be transferable between countries or regions.
A third problem is that minimum quantities of energy are quite different country to country. This makes a tool designed in one country difficult to use in another. There needs to be flexibility in design. The policy should be flexible so that countries with cold weather could include heat, and countries with large land areas could include transportation.

4.2.3 Minimum Energy Bundles

4.2.3.1 Characteristics of Households – Missing Data

There are many households that have missing data for some types of energy. For instance about one half of households list LPG as their primary cooking fuel, while only a quarter of households list any expenditure on LPG. This is not too surprising given that the survey asks for expenditure in the previous month, and the typical household will only buy a tank of LPG every two to three months. A 15 kg tank of LPG costs approximately 300 baht, with the average monthly expenditure on LPG at about 100-150 baht, made up of some months with no expenditure, and others with 300 baht. This problem was resolved by using a different variable that asks whether the household uses LPG as its main cooking fuel, and then averaging expenditure across those who did and did not purchase that month.

The situation is not quite the same with benzene as usage could vary widely based on the transportation options (own motorcycles, cars?) and travel needs (live far from work or town?) of the family. 77% of all households reported owning a motorcycle and 38% reported owning a car or pickup truck, while 79% reporting buying gasoline. The implication is that missing data should not be too much of an issue here, although it is possible.

Diesel is used by a relatively small share of households, mostly wealthy ones. 50% of diesel is used by richest 10% of households. As such it pays a very small role in our poverty index.

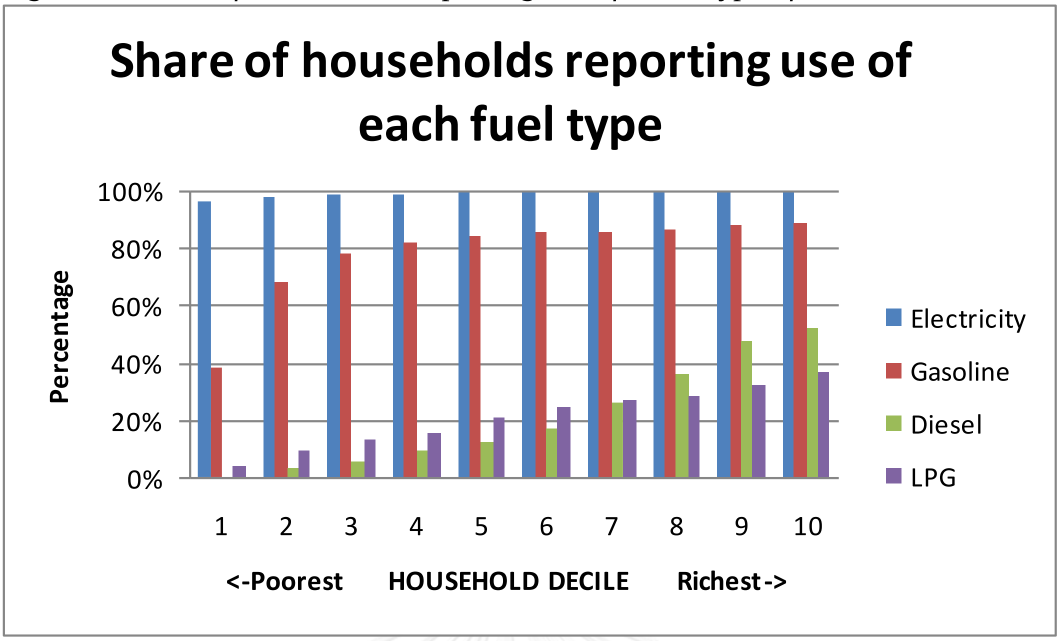
Electricity is also less of a problem as almost all households have electricity (99%) and are using a positive amount of electricity each month. However pricing for electricity is problematic, as electricity use of less than 90 KwH is free of charge due to government policies. Therefore pricing pressures are distorted for poorer households. According to interviews with villagers, if a house uses 91 KwH of electricity they would pay for all 91 KwH plus a 8 baht connection fee and a 15 baht counter fee, but if they used 90 they would pay nothing. These particular villagers had both a refrigerator and an iron, and therefore used far more than 90 KwH, having a typical monthly usage of 200 KwH. Electricity pricing is also complicated at the higher end as increasing block pricing (price per KwH increases with higher usage) is used to subsidize the poorer households and penalizing higher energy users. Where possible, this seems like a rational approach to energy problems broached by this paper.

The usage of electricity for the poorest decile households only rose about 1.7% comparing the year and a half before the policy and the year and a half after, so it does not seem that the policy had a dramatic effect on the way electricity was used.

4.2.4 Share of Households Using Each Fuel by Decile

Although we can calculate the amount of energy people use from each group, we would also like to know the share of households using each type of fuel type.

Figure 49 Share of Households Reporting use of Each Type of Fuel

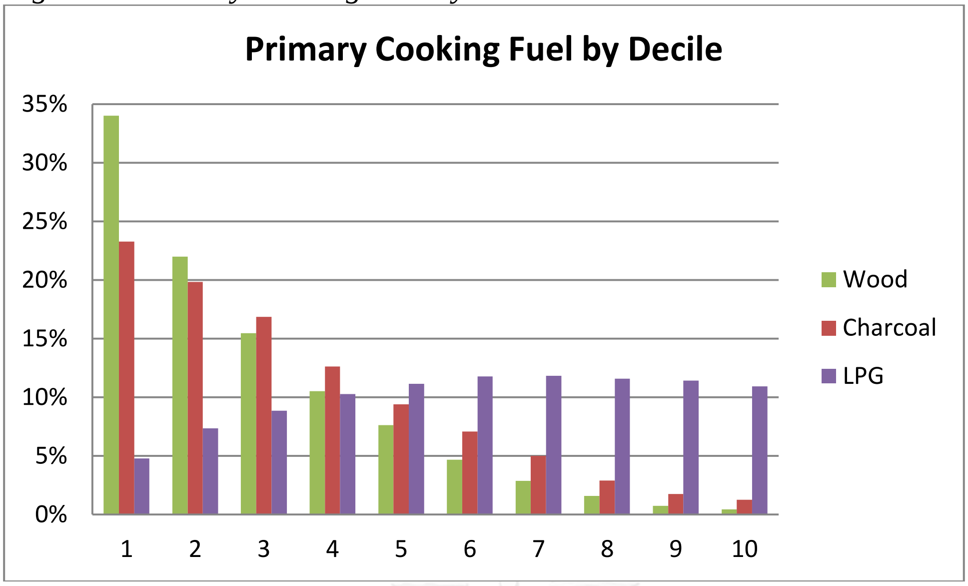


SOURCE: SES 2009 – HH DECILES

Electricity was the most commonly used energy source, with almost universal usage, but other types of fuel showed varying levels of access. For the 1st decile, less than 40% of households used any gasoline, suggesting that motorcycles were less available to this group. However, gasoline use increased rapidly with income with 65% of the 2nd decile used gasoline and 80% or more used gasoline thereafter. Diesel was used by less than 20% of households until about the 7th decile. Diesel is mainly used in pickup trucks or larger trucks which are not affordable to the poorest groups. Automobiles are also likely not available to the bottom half of households. Poorer groups could be affected by higher diesel prices indirectly though higher food or local transport prices, but the direct effect should be minimal. LPG is used by a small percentage of houses. Less than 20% of households in the poorest half of the population bought LPG during the survey month. This is likely because of use of wood and charcoal by poorer households, and because LPG does not have to be purchased every month.

The following chart shows the percentage of each household decile that reported that wood, charcoal or LPG was their most frequent cooking fuel.

Figure 50 Primary Cooking Fuel by Decile



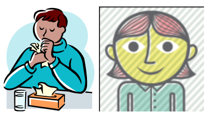
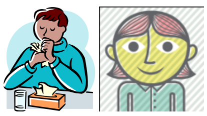
SOURCE: SES 2009 – HH DECILES

Households substitute LPG in place of wood and charcoal as income increases. Higher income households often did not cook at all, preferring to eat out, or used electric stoves (not shown).

4.2.5 Should Energy Bundles be Measured on a Household or Per Capita Basis?

Another problem emerges as to whether the proper unit to measure minimum energy needs is a household or an individual. A light bulb provides light for 3 people as well as it does for 1, a refrigerator can serve for a family or an individual, both arguing in favor of minimum energy needs at the household level. On the other hand, light may be needed in separate rooms by separate people, and clothing must be ironed for each person separately, arguing for per capita minimum requirements. It was decided to settle this issue empirically by seeing whether energy use per household or per capita better reflected expenditure patterns.

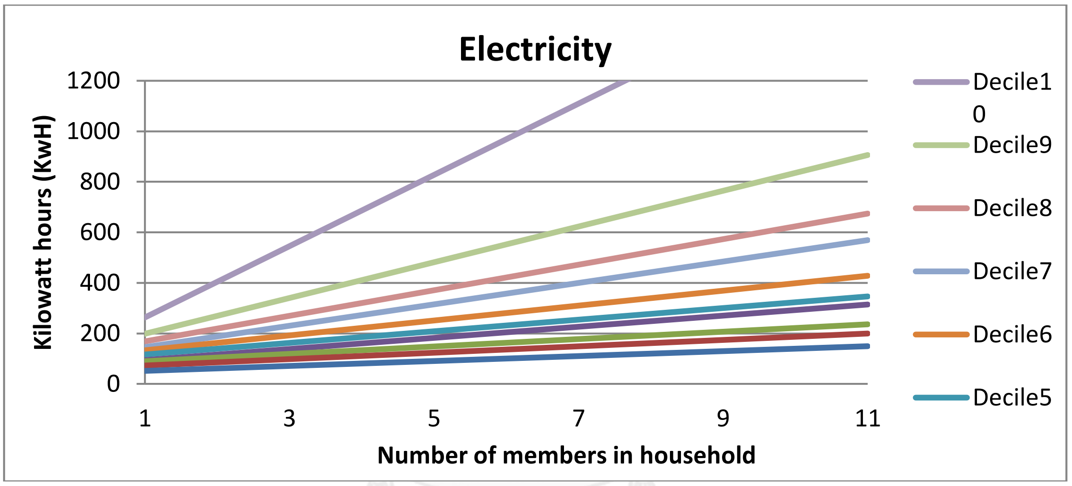
Figure 51 Comparing Energy by Household with Energy per Person

Family Members	House Uses	Each Person Uses
		
		

SOURCE: AUTHOR

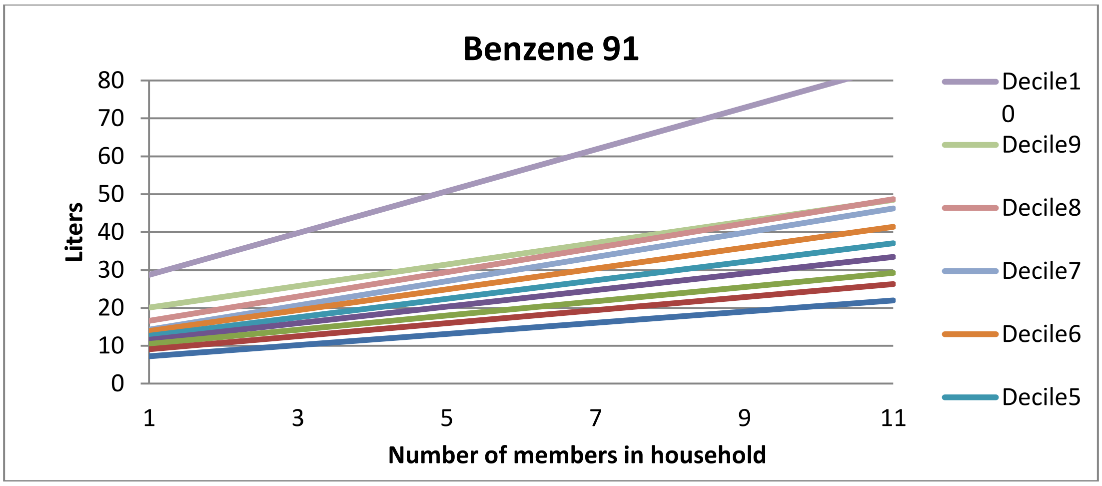
In this section we are trying to differentiate between the two cases shown above. In each row, the household uses seven units of electricity. But if the first row is true,

minimum bundles should be by household, and if the second is true, minimum bundles should be by person.
The following graphs are the result of regressing the quantity of fuel used by households in each decile by the number of family members. Most of the coefficients are significant at greater than the 1% level, but there are a few exceptions. Nevertheless data was included for all deciles for completeness, and because the lines as a group show a general trend.
The graphs start at 1 person since a household must have at least one person, and then slope upwards with other family members. There are two factors to look for:
If the **slopes are steep**, it implies that energy use is by person, since each person adds to energy use; if it is flat, energy use is by household.
If all the **lines converge** to a single point or to a narrow band on the left side, it implies that there is a clear standard amount of energy that each household must use.



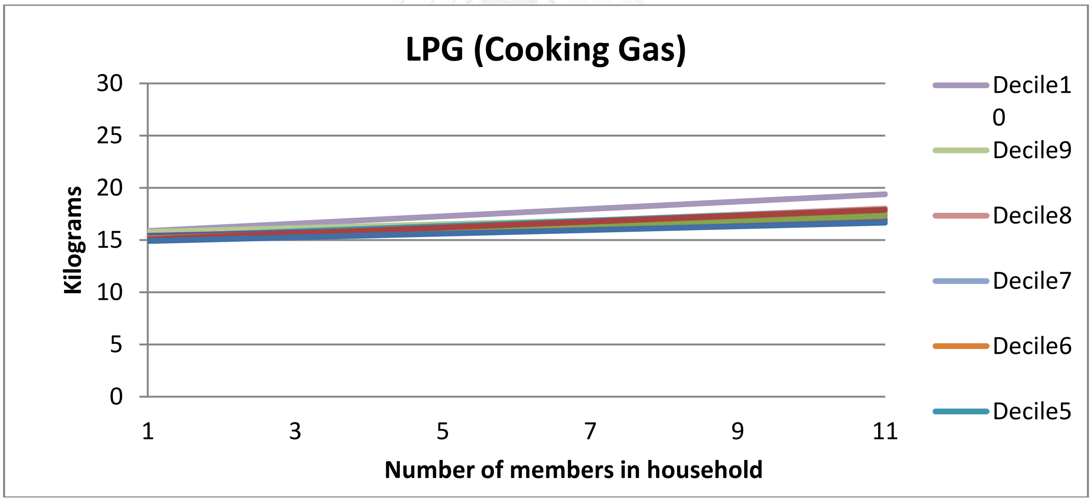
DATA SOURCE: SES 2009

The lines in this graph represent different deciles, with the poorest decile at the bottom and the richest at the top. The lines suggest that most electricity use is at the household level. The lines converge nicely on the left hand side, suggesting that there is some minimum level of energy required, and that it does not differ significantly for the poorer deciles. The relative flatness of the poorer deciles suggest that a certain amount of energy is needed to run any household, and that energy is shared amongst the family members. The lines become steeper as household expenditure increases. As family income increases, per person use becomes more important, culminating, perhaps, with each person sleeping in an air conditioned room, and having his or her own TV. The minimum bundle of electricity, bounded by the 2^{nd} to 5^{th} deciles, is approximately 70 to 125 kilowatt hours for a household with only one person.



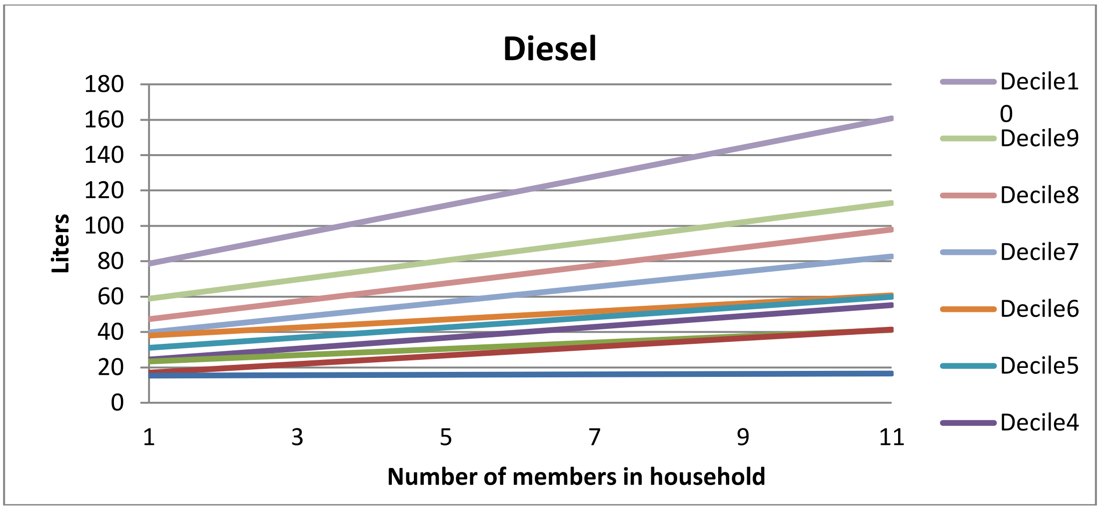
DATA SOURCE: SES 2009

The slopes in the graph for Benzene 91 are steeper than for electricity suggesting per capita use may be more appropriate, and are positive for every decile or level of expenditure, suggesting that the use of benzene increases as the number of people in the household increases. The minimum bundle seems to be 9 to 12.5 liters for the first person in the target 20% to 50% expenditure range. Benzene 91 requirements increase by about 2-3 liters per person as household size increases.



DATA SOURCE: SES 2009

From this diagram it appears as though demand must surely be at the household level. Every household of any expenditure level uses about 15 kilograms of LPG per month. However, LPG faces issues of missing data unlike the other energy series here. LPG is used in tanks of 15 kg explaining the shape of these lines. However, the typical family does not use a tank of LPG every month, and it may be two to three months between purchasing a tank. If a tank is not purchased in a given month, the expenditure on LPG is zero, and the data point is missing. If the tank is purchased the quantity used will be 15 kg. This diagram is not very effective. Another paper using Engel curves suggests that a per capita quota would be more effective for LPG.



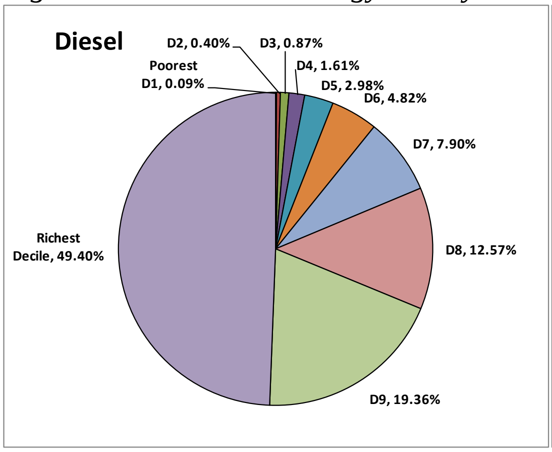
DATA SOURCE: SES 2009

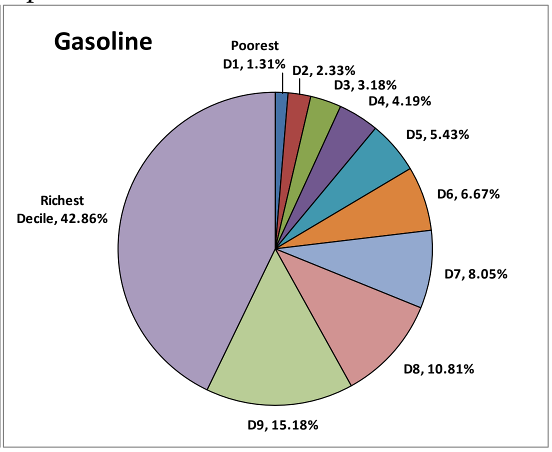
This graph is similar to the electricity graph in that the slopes of the decile lines change as income or expenditure increases. Diesel is used by only a small minority of the poor households, and data for the lowest decile line was statistically insignificant, unlike the other charts shown here. Interpretation of the lines is that diesel is generally used for work vehicles, and that a family is likely to have only one such work vehicle (borne out in a later section of this paper) , and at poorer expenditure levels the use of that vehicle will not be very dependent on the number or members of the household. For the few poor households who used diesel, minimum usage was higher than benzene 91 at 20 to 30 liters a month.

4.2.6 Energy Use by Decile

The energy pie charts show the amount of energy used by each household decile, calculated by adding up the energy use of all households in each decile and multiplying by survey weights. Use of all fuel types is dominated by the several richest deciles, but for diesel, gasohol and benzene 95, the results show half, or more, of all fuel being used by the richest 10 percent of households.

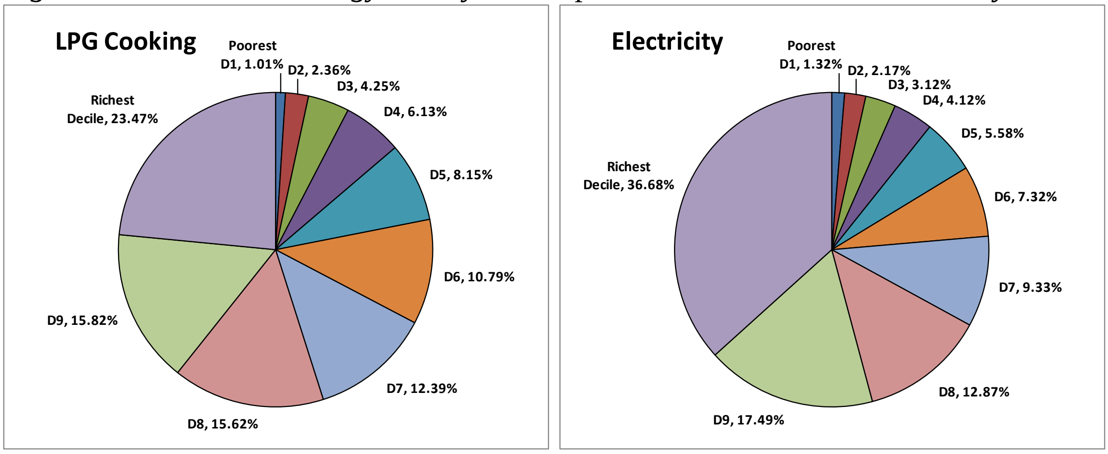
Figure 52 Residential Energy Use by Per Capita Decile – Diesel and Gasoline





SOURCE SES 2009, DECILE BY HOUSEHOLD

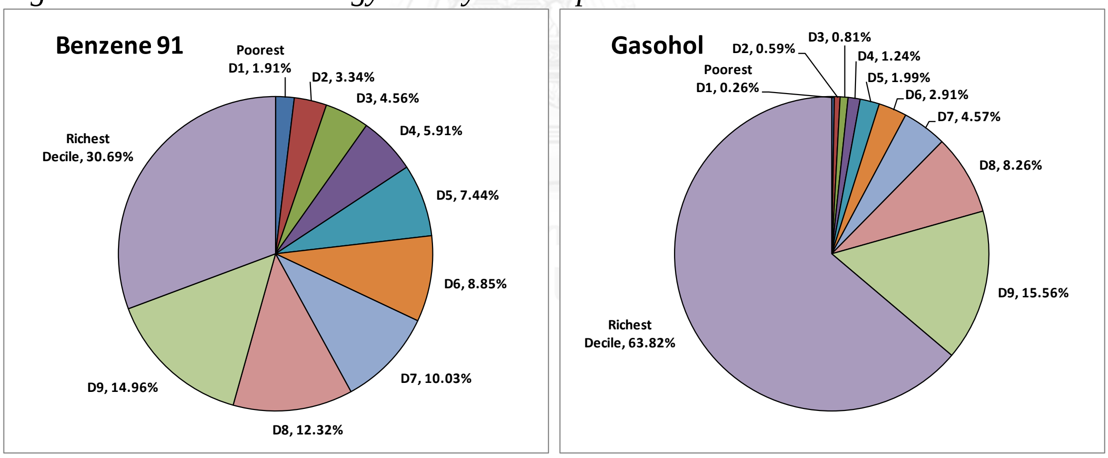
Figure 53 Residential Energy Use by Per Capita Decile – LPG and Electricity



SOURCE SES 2009, DECILE BY HOUSEHOLD

The implications are clear – Subsidizing the price of any type of energy directly will result in more than half the cost of the program going to subsidize the wealthiest 1, 2 or maybe 3 deciles. Targeted programs for the poorest groups are much more cost effective.
When first introduced, the Poor use Benzene 91, while Gasohol was only used by the wealthiest because of rumors that Gasohol would harm motorcycle engines

Figure 54 Residential Energy Use by Per Capita Decile – Benzene91 and Gasohol



SOURCE SES 2009, DECILE BY HOUSEHOLD

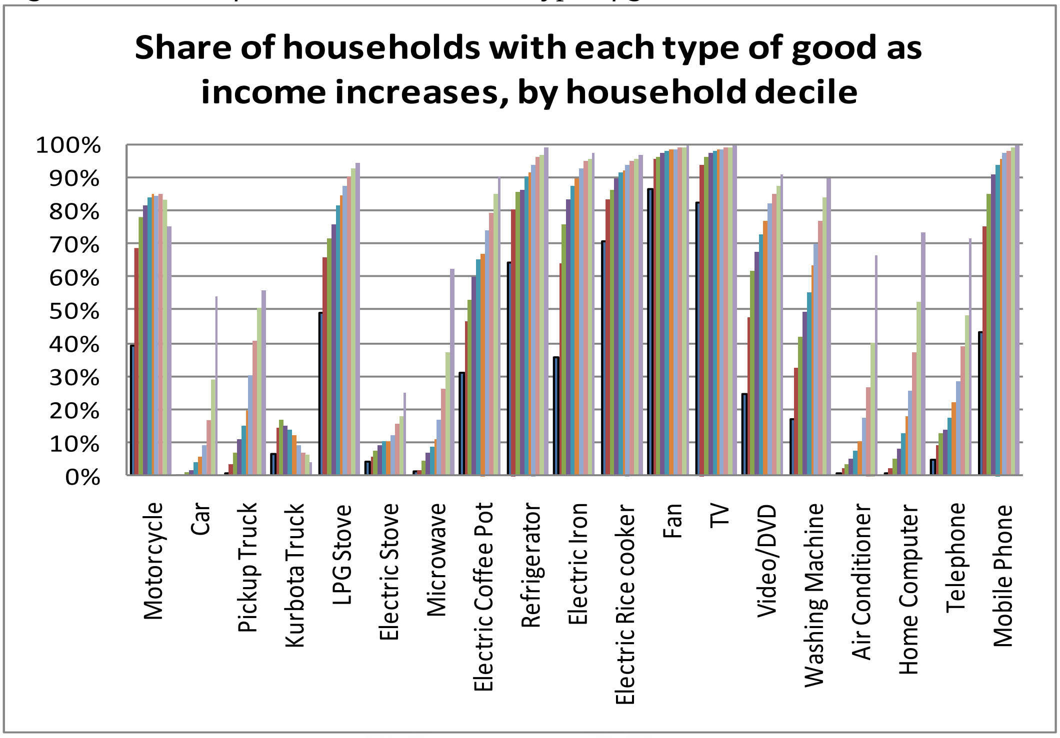
Since many recent policies have promoted Gasohol specifically, we also separate out Benzene 91 from Gasohol. Irrespective of the benefits that gasohol may present in terms of energy security and preparing for the future, it is clear that in 2009, subsidizing gasohol, was of value to only the very richest groups. Although it is not shown in the chart, in 2009, 21% of gasohol was used by the richest 1% of persons!

4.2.7 Energy use by end use & cost of maintaining energy using items.

Another entirely different approach to a minimum energy bundle would be to attempt to ensure access to certain types of physical goods for poor households. Energy

requirements would then follow from the energy required to maintain that particular good.
The following table shows the share of households of each decile that has access to an energy using good. In each group, the poorest decile is to the left in a darker font, and the richest to the right. The target should be to ensure access to goods that are currently widely used by the 2nd to 5th decile of households.

Figure 55 Share of households with each type of good, over household income decile



SOURCE: SES 2009 – HH DECILES – LEFT IS POOREST 1ST DECILE FOR EACH GOOD

As income increases, so does the use of virtually all electrical and other energy using appliances. In Thailand, access to motorcycles, refrigerators, televisions, LPG stoves, and mobile phones are available to almost all expenditure classes. Cars, pickup trucks, microwaves, air conditioning, and home computers are only available and commonly used by the wealthiest households. Although motorcycles, refrigerators, TVs and LPG stoves are available to almost all deciles, the exception is the poorest decile (shown as the left column for each good) in which the access to that good is substantially reduced relative to other deciles for virtually all goods. This is the group that we would designate as energy poor - without adequate access to modern energy. The problem is not a lack of electrification – even in the poorest decile, more than 97% of households have electricity.
Next we use simple regressions to look at the typical energy requirements to own different types of electrical appliances or transportation. The following regressions are a result of regressing electricity expense by household on home ownership of various electrical appliances, as described below. Results are mostly significantly different from zero and are as follows:

Table 18 Electricity Expenses per month in Baht for each appliance

ELECTRICITY EXPENSE PER MONTH IN BAHT	(1) Poorest 10%	(2) Poorest 50%
Household Monthly Expenditure (control)	0.0193***	0.0210***
	(0.00138)	(0.000499)
Electric Kettle	4.963	17.81***
	(3.258)	(2.191)
Refrigerator	30.99***	37.89***
	(3.088)	(2.745)
Electric Rice Cooker	23.73***	15.50***
	(2.432)	(2.418)
Fan	7.204***	16.88***
	(1.931)	(1.534)
Television	11.29***	17.61***
	(3.477)	(3.492)
Mobile Phone	6.096**	1.921
	(2.831)	(1.932)
Fluorescent Light Bulbs	2.199**	7.216***
	(0.945)	(0.675)
Constant	-11.70**	-58.32***
	(4.991)	(4.073)
Observations	3,836	19,773
R-squared	**0.218**	**0.246**

Robust standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1

In the table above, coefficients are estimates of baht per month spent on each electrical appliance. So that 30.99 baht per month were spent by the poorest group on refrigeration, and 37.89 baht per month were spent by the poorest half of the population on refrigeration, assuming, of course, that they had a refrigerator. Coefficients are per unit, so that in some cases, the family might have more than one fan (typically 2) or light bulbs (typically 4-5), with coefficients above representing the per fan or per light bulb cost in baht. For the other goods listed, the household would typically have just one. Some goods that are significant in terms of electricity use such as air conditioning, were not included as they were not an important item for the poorest 50% of households. Share of total expenditure on electricity is about 2% for the poorest 10% and poorest 50% groups.

The data comes from the following regression:

$$Electricity = \alpha + \beta_1 a07 + \beta_2 hh26 + \beta_3 hh27 + \beta_4 hh29 + \beta_5 hh30 + \beta_6 hh32 + \beta_7 hh40 + \beta_8 hh42$$

where, following nomenclature of the social economic survey, variables are defined as:

- electricity expenditure on electricity (eg181 and eg182) poorest 10% =131 baht, poorest 50% = 238 baht
- a07 total household expenditure
- hh26 # of electric kettles in household
- hh27 # of refrigerators in household
- hh29 # of electric rice cookers in household
- hh30 # of electric fans in household
- hh32 # of televisions in household
- hh40 # of mobile phones in household
- hh42 # of fluorescent light bulbs in household

Similar regressions are possible for gasoline and LPG as follows, where dependent variables are monthly expenditure on gasoline and LPG respectively, and other variables are # of the goods specified.

Table 19 Gasoline Expenses per month in Baht for each vehicle

GASOLINE EXPENSE PER MONTH IN BAHT	(1) Poorest 10%	(2) Poorest 50%
Household Monthly Expenditure (control)	0.0564*** (0.00548)	0.0354*** (0.00104)
Motorcycle	61.52*** (11.35)	109.5*** (4.244)
Automobile	258.1*** (55.83)	326.4*** (15.78)
Constant	-23.64 (24.56)	27.50*** (9.245)
Observations	1,529	14,180
R-squared	**0.100**	**0.163**

Robust standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1

Note from the household monthly expenditure variable above that expenditure of gasoline for those who purchased it was a greater share of income than electricity. For the poorest 10%, gasoline was 5.6% of their monthly expenditure while for the poorest half monthly expenditure was 3.5% of monthly expenditure. The fuel cost of maintaining a motorcycle varied from 62 baht to 110 baht for the poorest half, probably due to greater use.

Table 20 LPG Expenses per month in Baht for each appliance

LPG EXPENSE PER MONTH IN BAHT	(1) Poorest 10%	(2) Poorest 50%
Household Monthly Expenditure (control)	0.00470**	0.000935***
	(0.00197)	(0.000264)
Gas Stove	20.59	13.32***
	(38.79)	(4.790)
Constant	234.1***	267.9***
	(41.04)	(5.488)
Observations	165	2,711
R-squared	**0.036**	**0.008**

Robust standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1

Monthly expense on LPG per household was around 1% of monthly expenditure. The constant term is quite high in this equation at around 250 baht. Only those with gas stoves would buy LPG (Gas for vehicles is not included in this category) so the low value of the coefficient likely refers to the addition use if the family owned an additional gas stove. In other words, the family might spend 250 baht on LPG if they had one stove and 270 if they had two stoves. Note that only about n=165 or about 4% of the poorest group actually had a gas stove, and the coefficient on gas stoves is not significant.

4.2.8 Energy share in expenditure

A final approach to minimum energy bundles is to look at the share of total expenditure used on energy. Overall, no decile group was found to use more than 10% of their income on energy. This could easily change if energy prices increased in the future.

The following table shows the share of expenditure used by all persons on each type of energy. The subcategories do not have to add up to the total since not everyone uses all sources of energy.

Table 21 Share of expenditure spent on energy products, if purchased

Fuel Type	Share of Total Expenditure
All Energy	9.16%
Electricity	3.03%
LPG	1.98%
Gasoline	5.07%
Diesel	7.33% if used

DATA SOURCE: SES 2009

The share of expenditure used on a fuel can change substantially by decile for some types of energy and not for others. Below is a table for total energy, electricity, LPG and gasoline sorted by per capita deciles.

Table 22 Share of Per Capita Budget Spent on Energy by Type and Decile

PC Deciles	Energy	Electricity	LPG	Gasoline
D.1 - Poorest	8.0	3.2	4.2	6.6
Decile 2	9.0	3.3	3.7	6.2
Decile 3	9.1	3.2	3.2	5.9
Decile 4	9.3	3.2	2.9	5.7
Decile 5	9.4	3.2	2.6	5.4
Decile 6	9.6	3.3	2.3	5.1
Decile 7	9.7	3.2	2.0	4.8
Decile 8	9.4	3.1	1.7	4.5
Decile 9	9.1	2.9	1.3	4.4
D.10-Wealthy	8.6	2.4	0.9	4.4

SOURCE: SES 2009

The poorest decile does not spend much on energy since they own many fewer energy using items. LPG and gasoline decrease quickly as a share of expenditure, while electricity stays roughly constant as a share of expenditure. LPG refers to people who actually bought a tank of LPG in the survey month, so share of expenditure overall should be less.

Overall, this approach would be most useful if we could track expenditure share over time. At present, the data now available to us is cross sectional, so there is a limit to what can be done in analyzing energy poverty, but with better data this would be an effective way to track changes in energy poverty over time, especially if the price of energy increases.

4.2.9 History of Poverty Lines and Energy Poverty Lines

The first section of this paper dealt with trying to determine a minimum bundle of energy. This section discusses the possible measures we can use measure poverty, their strengths and weaknesses, and some historical background for poverty lines.

An energy poverty line depends heavily on the poverty line literature of the 1960s and 1970s. In the past, energy poverty lines were probably created occasionally as bi-products for a general poverty index, however, when energy was relatively cheap, but not accessible because of connectivity or availability issues, the debate naturally shifted to electrification, better stoves, in short, access to modern energy in any form. The debate at that time focuses on life-style changes to improve the health and welfare of the world's poorest. In the world's poorest regions, that debate continues today.

The poverty line literature is well-established and I will only discuss it a bit here. As national survey data became widely available it became easier to formally measure poverty in the form of a poverty line.

The poverty line index focused on the true necessities in order to survive healthfully and often this meant access to sufficient food calories to live. As such, it was a very critical line.

Starvation was a widespread problem in a number of regions, unlike isolated areas and times today. Poverty lines in this context were most appropriate for the poorest countries. Nonetheless, poverty lines were developed for middle income and wealthy

income countries as well to study the level of income inequality and the need for and effect of different economic policies.
Poverty lines, in this context were a bit artificial, and faced two severe challenges.
1) Evidence of a clear break, or change in behavior, or inability to forgo consumption was rare and hard to find. If there is truly a poverty line, consumption or food should drop with income to a certain level, and then be fixed.
2) Poverty lines were unstable over both time and space, so that income in urban areas had to be higher than in rural areas to support the same minimum lifestyle. Over time, as overall income levels rise, so did expectations, and the consumption considered to be a minimum level kept increasing. Nonetheless, energy poverty lines, indices, the Gini index and Lorenz curve all held sway in the development literature because they served a useful purpose, to evaluate government polices and development regimes in terms of this poorest group. Does a certain policy push people above a certain line in the sand, or not? What percent of people are below this line? Besides the idea had a popular cache. Both of the challenges mentioned above made poverty lines contentious and hard to use. It was difficult to say for certain where the poverty line was (except for the poorest group where caloric needs set a bar.)
Also it was hard to measure against a moving line in the sand. For instance, if we define the poorest 10 percent as below the poverty line, than by definition 10% would be below that line at any given time. If we defined it as a certain income level at a fixed period of time, the index would lose its relevance as income levels changed in the country.

A further problem, but not an insurmountable one, was inflation which could vary among products. Thus, with a wide array of products the poverty line had to be defined in terms of expenditure making it somewhat unstable.

4.2.10 Energy Poverty Lines

Over the past decade, energy prices have tripled making energy more difficult to afford. Meanwhile, in most of the world, famine is a thing of the past, and after decades of worthy effort, electrification and physical access to most modern energy are the status quo.
The problems for poor, and middle income countries have changed from providing access to modern energy to keeping access to modern energy affordable to its population as prices inevitably rise. An energy poverty line is a way to measure the effects of government policy, including those designed to meet this objective.
An energy poverty line shares some of the problems of traditional poverty lines in middle income countries. In fact, being below a certain line is not life-threatening, it just represents a certain quality of life. For instance, the poorest deciles in Thailand tend to burn wood or charcoal for cooking, but higher deciles use LPG. LPG is cleaner, more convenient, healthier, but using wood is not immediately life-threatening. A minimum bundle is easier to define. Wealthier households do not need to cook more than poor people, so the required level of LPG for any person to cook is easier to define. (in other words, alternative uses of LPG for transport, producing food to sell, etc., is limited in extent.)

Second, with only a few energy types, prices for each type of energy can be found, and the index can be based on quantities making it more stable over time. This also avoids issue associated with inflation.
There are three common candidate approaches to an Energy poverty line, each with advantages and disadvantages. They include 1) defining a minimum bundle of energy that a poor person needs 2) the share of expenditure spent on energy by the poorest 10 percent of the population, and 3) defining the current energy bundle of the poorest 10 percent at a fixed period of time. I also discuss a fourth approach based on the energy required to supply certain modern goods such as motorcycles, LPG stoves and refrigerators.

Table 23 Strengths and Weaknesses of Different Types of Energy Index and Energy Poverty Lines

Quantity Based Measures			
Criteria	**Stability**	**Ease of Use**	**Transferability**
Define energy as a certain minimum requirement of each type of energy per capita	Criteria is stable and is clearly defined	Easy to use, transparent and comparable across countries	Not easily transferable - Varies widely by country and region
Define energy requirements for certain goods (e.g. clean stove) and have ownership be objective	Criteria is stable and is clearly defined May require updates if technology changes.	Easy to use, but may require updates as technology changes.	Transferability, if define energy needs for each type of good, a country can choose from a menu of objectives.
Define energy in terms of basic needs, e.g. transportation, light, cooked food	Changes with income levels and across countries.	Not easy to use as basic needs may be met in many ways – restaurants cook food, transport can be private or public,	Not easily transferable, as a quantity index. However, basic needs defined as value index work well already
Value Based Measures			
Criteria	**Stability**	**Ease of Use**	**Transferability**
Share of expenditure on energy of poorest decile	Changes with prices of non-energy goods. Energy price changes will also lead to high volatility.	Very easy to calculate	Easy to transfer across countries and regions.
Define energy in terms of expenditure on basic needs, e.g. transportation, light, cooked food	Criteria is stable and is clearly defined	Fairly easy to use	Easy to transfer across countries and regions.

SOURCE: SUMMARY BY AUTHOR

Most people in developing countries are poor – emphasis is on access to a modern lifestyle for those who live in the country, insomuch as is possible! The share of energy by the poorest groups in a country is miniscule compared to the largest energy users. The emphasis cannot be just to reduce energy use for everyone. The issue needs to one of distribution of energy, with wealthier countries reducing more than poorer countries, and the wealthy in poor countries reducing enough to allow room for poorer groups to maintain a modern lifestyle. What does access to modern lifestyle entail? The poor need to have a way to have cooking lighting heat and cooling and transportation in a way that is healthy. It is not required that people would actually die without access to these things, but for a small burden on the world's energy supply the quality of these people could be much enhanced. Granted easing hardship must be within our financial means, but almost all societies can afford to do a little, and many can afford to do a lot.

Very little energy is required to sustain live. We can huddle in the cold and eat raw vegetables in the dark and still survive, and maybe that is the way our ancestors lived. But one of the goals of society is economic development, to ease this sort of hardship. With modern technology such as stoves, electricity, engines, an easier lifestyle is within reach of much of the world's population and is cheaper than ever before. Although only a little energy is required for the minimum requirements for this modern lifestyle, it is still some, and that is our goal to explore what is needed.

Energy use is tied to certain goods.

Wood -> Charcoal -> LPG -> Electric Stoves

None - > Kerosene and Candles -> Electricity

Walk -> Bicycle -> Motorcycle -> Car

Table 24 Hypothesized Effects of Subsidies on Different Income Groups

Fuel Type	Group Benefited Most	Effect on Energy Inequality
Electricity	Poor	⇩
LPG	Poor	⇩
Benzene91	All	⇔
Benzene95	Wealthy	⇧
Gasohol	Wealthy	⇧
Diesel	Middle and Wealthy	⇧

SOURCE: SUMMARY BY AUTHOR

4.3 Expenditure Time Series

4.3.1 Introduction

The idea behind the expenditure series is incredibly simple, although the econometric coding is not nearly as easy. The idea is that with a continuously collected survey, such as the Thai SES, it is possible to divide the sample along a couple of continuums and take averages, and that these averages will in themselves be sufficiently strong statistically due to large data properties, that they can be chained together to form a time series. With a total annual sample size of about 44,000 households, dividing the sample by 12 months and by ten income deciles gives an average sample size of 44000/120 = 367. By itself, this sample size is adequate, but barely adequate to use in statistics. However, once we use some sort of smoothing technique – in our case we use 6 month equally weighted smoothing, that makes the sample size increase to 6*367 = 2,200 which gives stable results over time for each of the income deciles. This stability can be observed in the time series shown below.

4.3.2 Methodology

4.3.2.1 Data

The data used in this project are mostly from the National Statistics Organization's SES data, with details as follows:

Table 25 Data used in Expenditure Time Series Project

Year	Source	Variables	Observations
2006	SES	570	44,918
2007	SES	566	43,055
2008	SES	385	44,969
2009	SES	595	43,844
2010	SES	392	44,273
2011	SES	545	42,192
2012	SES	411	43,762
2013	SES	594	42,738
1995-2013	NESDB	23 subcategories	77 provinces
2007	Electoral Commission	41 parties	76 provinces
2011	Electoral Commission	40 parties	77 Provinces

SOURCE: VARIOUS AS DENOTED IN TABLE

Altogether a bit over 20 million data points

4.3.2.2 Pseudo-panel and Sample Size Issues

The SES is a survey, not true panel data. There is, however, stability in the means of most data. From the central limit theorem, as sample size increases, the sample mean approaches the population mean for most variables, with an average value of population mean u, and the variation depending on sample as given by the t-statistic below.

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N}}}$$

In practice, as sample size grows we find that a series of samples give very similar sample means. From the t-statistic about with a sample size N of 400, sqrt(N) = 20, and unless the standard error is very large, the mean remains nearly the same from sample to sample as observed empirically. As the SES survey consists of about

44,000 observations, it is generally possible to divide each survey into approximately 100 pieces and still get stable means. The "pieces" could be provinces, professions, months, regions, or a combination of these or other variables.
In the expenditure database I employ three versions of each variable j*k times:
A) The average expenditure on a good averaged over all possible households even if they don't buy. (10 time series)
B) The average expenditure on a good averaged over only those households that do buy (10 time series)
C) The share of all possible households in a group that do buy the product. (10 time series)
All 3 are needed to answer different sorts of questions. There are k time series of each type, where j could be monthly values strung together to make the time series and k could be income decile.

$$Version\ A) \quad \forall\, j\, \forall\, k \qquad all_{jk} = \sum_{i=1}^{q_{jk}} x_{ijk}/q_{jk}$$

$$Version\ B) \ \forall\, j\, \forall\, k \qquad users_{jk} = \sum_{i=1}^{l_{jk}} x_{ijk}/l_{jk}$$

$$Share = \ Version\ C) \ \forall\, j\, \forall\, k \qquad share_{jk} = {q_{jk}}/{l_{ik}} * 100$$

q_{jk} = possible incidences of the i variable if j=J and k=K (the sample size)
l_{jk} = actual incidences of the i variable if j=J and k=K (non-zero sample size)
Some possible divisions of the data

- by month (12) and expenditure decile (10)
- by province (77)
- by month (12) and region (5)
- by month (12) and profession (7)
- by education level (8) and income (10)

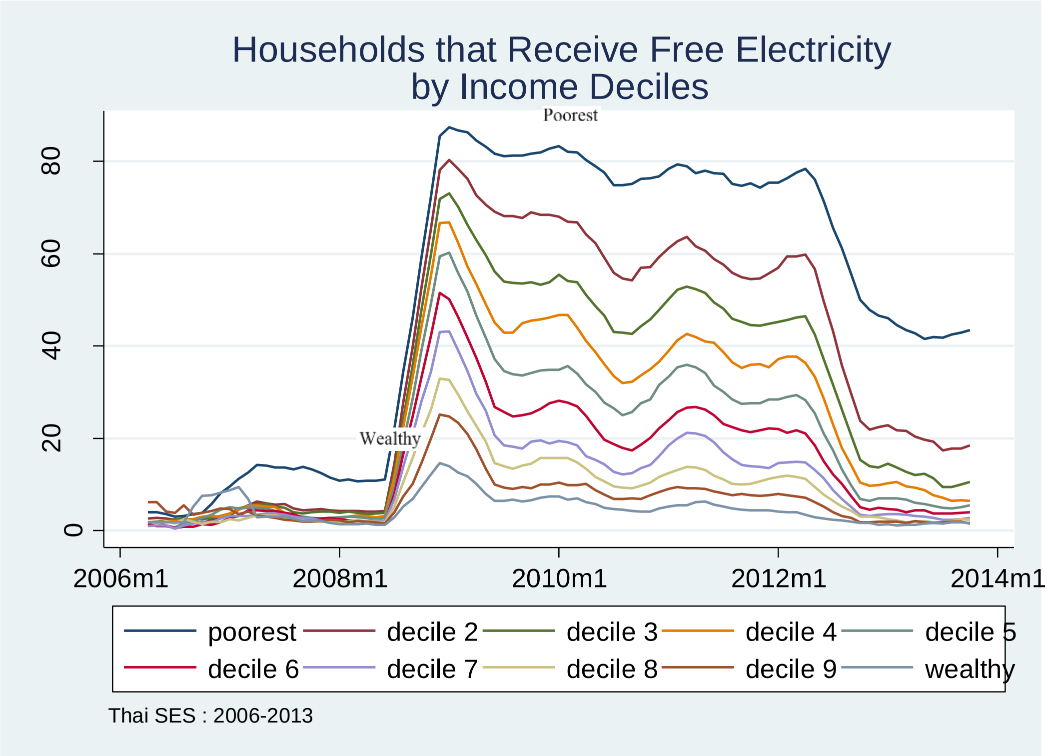
… don't know others yet.
Although this technique is very flexible, I have mostly only experimented with the first two of these options.
Although we have used the mean in experiments so far, it may be better to use the median in some cases. It is a simple matter to change a line in the code to do so.

4.3.3 Time Series Representations

4.3.3.1 Free 90 KwH electricity program (2008-present)

The Free electricity program (lifeline levels) was initiated in 2008. The purpose was to help households that used exceptionally low levels of electricity under the assumption that they were the poorest. Compared to other aid programs this one was surprisingly effective. Using 330,000 households from 8 years of the SES expenditure survey, we can see the pattern of benefits from the program.

Figure 56 Timeline of Share of Households Using Free Electricity



SOURCE: SES 2006-2013 NOTE: IN THIS CHART, EACH INCOME DECILE HAS ITS OWN LINE SHOWING THE SHARE OF THAT DECILE THAT BENEFITS FROM THE PROGRAM. THE HORIZONTAL AXIS IS TIME IN MONTHS. IN 2006, THE PROGRAM DIDN'T EXIST, SO NO ONE BENEFITED. IN MID 2008 THE POLICY WAS INTRODUCED , WITH FREE ELECTRICITY FOR THOSE WHO USED BELOW 90 KWH. IN 2013, THE CUTOFF WAS REDUCED TO 50 KWH.

A small trial program was floated in 2007, and the full program began in the middle of 2008. This graph shows what share of each decile received free electricity, with the line on the top being the poorest decile. Eighty percent of the poorest group took advantage of the program. About 60% of the second poorest group took advantage of the program, 50% of the 3rd poorest group, etc. The problem comes when we get down to the 6th poorest group, or households that are slightly wealthier than the average. In this group, about 20% of households took advantage of the program. Programs that target the poor must strike a careful balance between helping those they wish to help and helping those who do not need help. This could be thought of as a false positive and true positive problem. We want the true positives, but not the false positives. Aid programs always attract free riders who don't need subsidies.

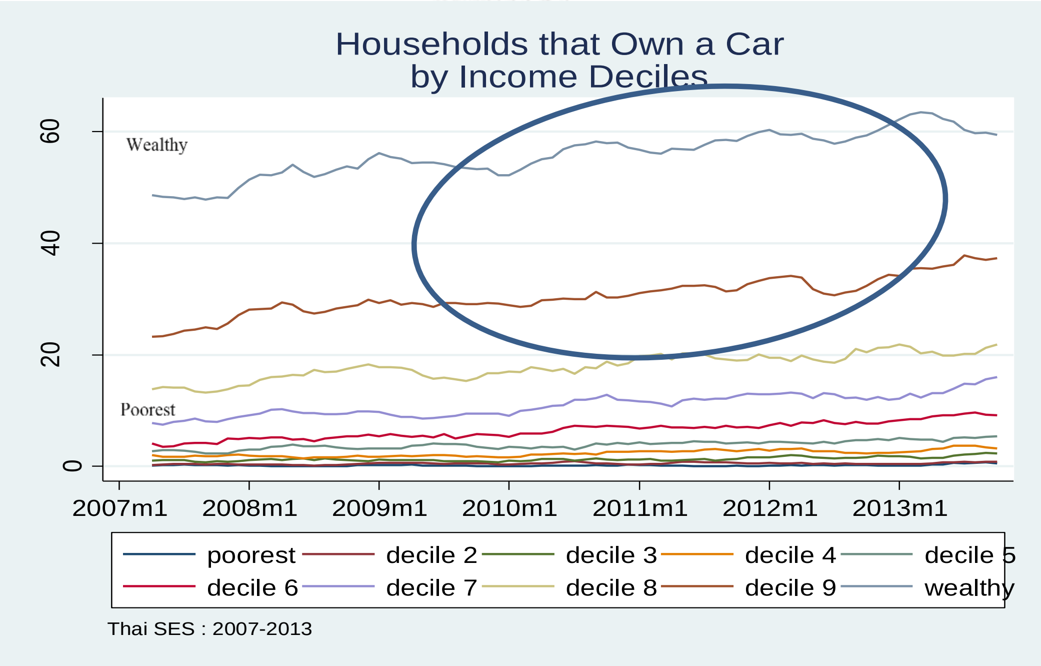
The government was upset about these "false positives" and in 2013, the lifeline program was cut so that only households using less than 50 KwH of electricity received it free. Immediately this eliminated the false positives in decile 6. However, the uptake rate of the poorest group (decile 1) dropped dramatically from 80% to 50%. In other words, to get rid of the false positives we had to get rid of the true positives as well. Was this a good decision? I would argue No – that this program has better targeted poor people than any other energy program, and it should have been left as it was, or possible could use a higher cutoff KwH of about 80 KwH as suggested by the Engel curves in section 7 above.

4.3.3.2 First Car Program

In 2011 Bangkok experienced severe flooding and many of foreign automobile factories were damaged due to what they attributed to negligence on the part of the Thai government. The argument was that the government protected Bangkok to the detriment of provinces north of Bangkok where many car producers were located. These foreign car producers threatened to move to Indonesia which has a much bigger population.

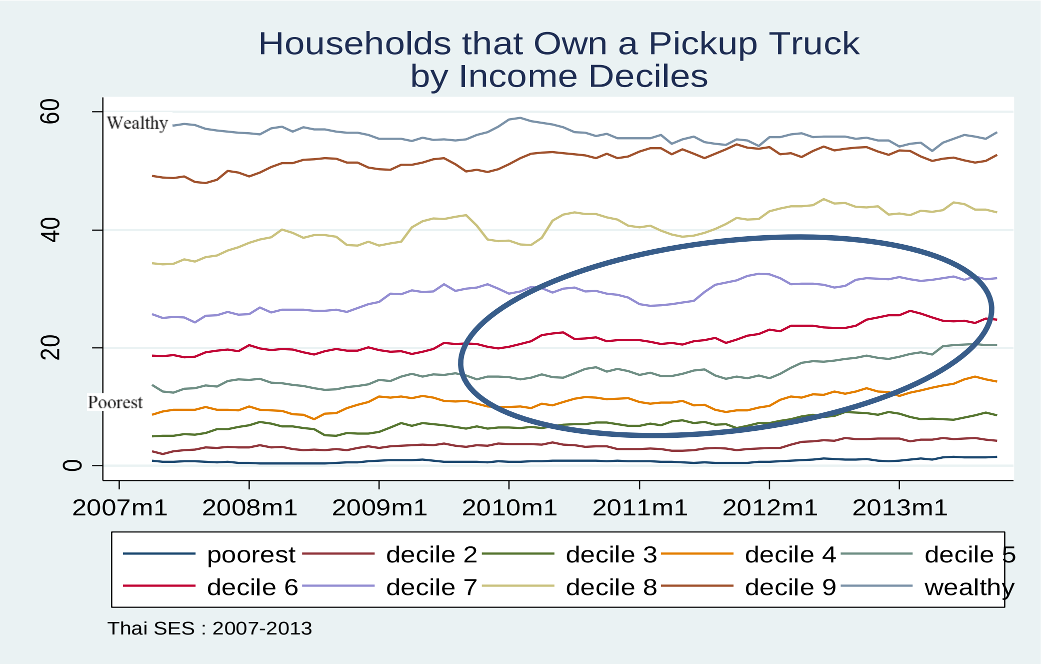
The First Car Policy, which eliminates a 25% tax on new cars, had recently been initiated and was enlisted to help keep these car companies from leaving Thailand. In the popular media the program was portrayed as a way to help struggling new families obtain their first car. However there were always suspicions that benefits went to wealthy families who already had vehicles. Looking at the accompanying time series data helps us investigate the truth of these allegations.

Figure 57 Timeline of Share of Households by Decile that Own a Car



TIME SERIES RESULTS SHOW THAT MOST OF THE BENEFIT OF THE FIRST CAR POLICY INDEED WENT TO THE TOP TWO DECILES IN TERMS OF INCOME AS THE SLOPE OF CAR OWNERSHIP INCREASED FOR THESE DECILES AFTER THE POLICY TOOK PLACE.

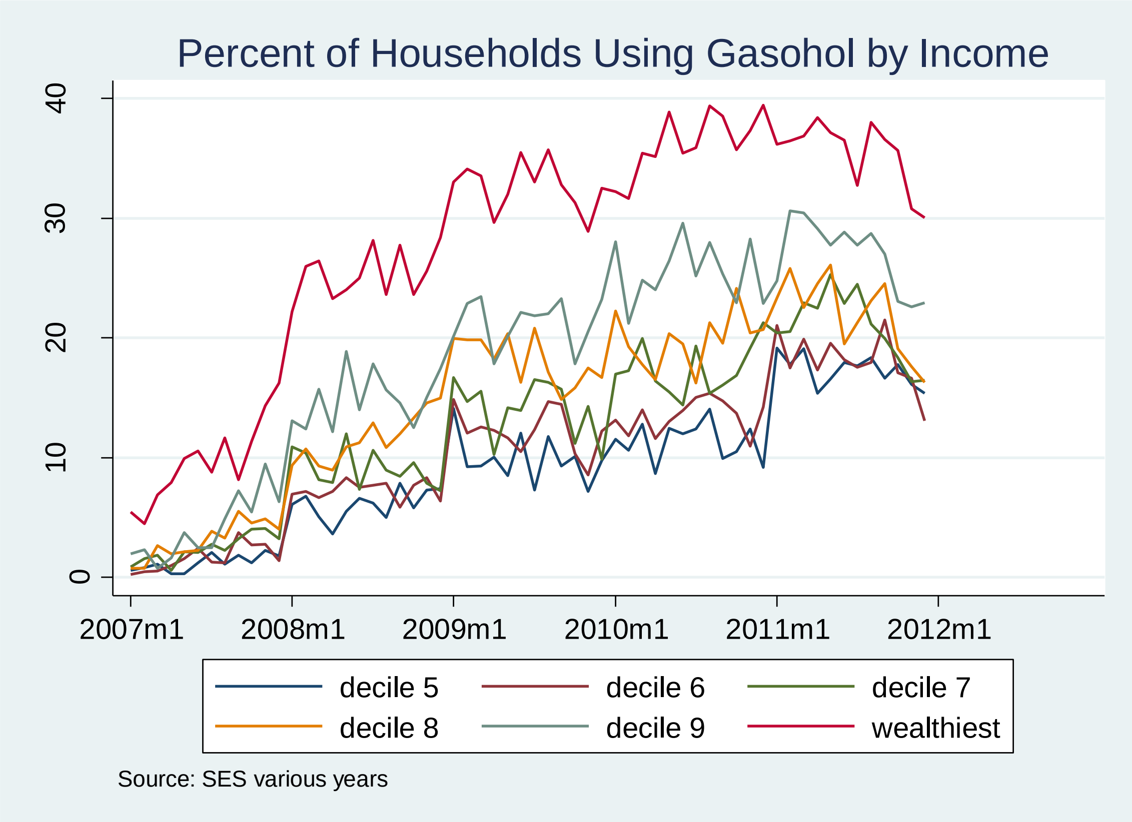
Figure 58 Timeline of Share of Households by Decile that Own a Pickup Truck



NOTE: IN FIGURES 10 AND 11 THE WEALTHIEST LINE IS THE HIGHEST ONE. THE LINES SHOW THE SHARE OF HOUSEHOLDS IN EACH DECILE THAT OWN A CAR OR PICKUP TRUCK. IN 2011 WHEN THE FIRST CAR POLICY WAS STARTED, WE CAN SEE AN UPWARD CHANGE IN THE SLOPE OF SOME OF THESE LINES, REFLECTING INCREASED OWNERSHIP OF CARS AND PICKUP TRUCKS. POLICY IMPLICATIONS ARE THAT THE FIRST CAR POLICY WAS MORE USEFUL TO THE KINGDOM IN PROMOTING PICKUP TRUCKS THAN IT WAS IN PROMOTING PASSENGER CARS, AS IT LIKELY ENABLED POORER FAMILIES TO OWN WHAT IS USUALLY A PRODUCTIVE VEHICLE.

However, when we look at Pickup trucks, the first car policy seemed to be much more beneficial to lower deciles, without really affecting the wealthiest groups. Perhaps we should have had a FIRST PICKUP TRUCK policy?

Figure 59 Timeline of Share of Households by Decile that Use Gasohol



SOURCE: SES 2011

4.3.3.3 Gasohol Analysis

Subsidizing Gasohol was primarily a subsidy for the wealthiest decile and producers until the policy changed in 2013 to require gas stations to replace benzene 91 with gasohol. There was a widespread belief that gasohol would damage motorcycle engines and so most persons did not switch until required to do so, despite a small price advantage. In fact, many of the users were well educated persons who were willing to have their own views and beliefs, especially about helping the environment. In 2013 there should be a massive up spike in gasohol usage but it is not shown on this graph. In fact because not all gasoline is branded – e.g. motorcycles buying gasoline from dark bottles by the side of the road, it is not clear that everyone knew that benzene91 was removed from the market. About 10% of households report buying benzene 91 in 2013 after it was no longer for sale. This chart is an example of what the technique looks like if no smoothing is employed.

4.3.4 Conclusion

In summary, expenditure time series are a useful way to study the effect of a subsidy over time. The fortunate inclusion of a date in the Thai socio-economic survey allows for a continuous time series of expenditure data, but averaging over all users in a particular month. Data is collected in every province and region of Thailand on a continuous basis allowing decent statistics to be obtained for each month. It is possible to generate an aggregate index of the time series variables to test how inclusive or how pro-poor a policy can be. Such a development should be the next step in this research agenda. The figure below gives an idea of what such an aggregate index might like, with the index using a second y-scale on the right hand axis.

4.4 Expenditure Budget Pies

4.4.1 Introduction

Energy Expenditure Pies are a tool to look at how our subsidy budget is being spent. The objective of this part is to answer the question – which subsidies are worth supporting? There are two tools provided in this section. The first, and most important, is an energy budget pie that shows out of the budget that goes to the final user, how much goes to each income group. Policies that are proposed as a way to help the poor, but are primarily a subsidy for the rich, should be considered for termination. Policies that reach mainly the poor should be continued. The energy expenditure pie charts proposed here look at the final amount received by customers – they cannot consider administration costs, but nonetheless they provide useful and hard to obtain information. The second tool are bar charts that show what percent of an income group benefits from a subsidy in terms of percentage of a decile or in this case a quartile. This uptake issue is very important, but it has already been looked at closely in several other places in this paper, so it is included here just as check on our previous work.

4.4.2 Methodology

The SES contains data about how much a household spends on each energy source. It also uses weights so that by multiplying each household by the number of other households it represents, we can get back to the 20 million or so households that exist in Thailand. For instance suppose that in our representative sample a household in the 6th income decile spent 140 baht on LPG. According to the survey weights this particular household represents 324 other similar households. Therefore if we multiply 324*140=45360 baht, we can say that this household and households like it added 45360 to the 6th income decile. If we continue with all of the 44,000 households in the survey we can obtain an estimate of how much the final amount spent on LPG for a month was, and also we can show what percent of that total went to households of each income decile.

The number we obtain will estimate the total amount end users spent on LPG. This number will not be the total amount of subsidy. Suppose the subsidy is 20% of the final price. Then the amount of the subsidy will be 20 percent of the figure we calculated earlier. However we do not need to worry about the actual amount of the subsidy because generally we really only want to know the percent of the subsidy that goes to each group. That percentage will be just the same as the percent spent on LPG by each group. Consider the subsidy on diesel oil. The subsidy will be a very small part of the total amount spent on diesel, but the share of the subsidy for each group will exactly match the share each group spends on diesel.

The share of each group that uses a subsidy is calculated in a similar manner to that used in previous sections of the paper. The share of each group with non-zero expenditure is calculated, and reported in this case in quartiles.

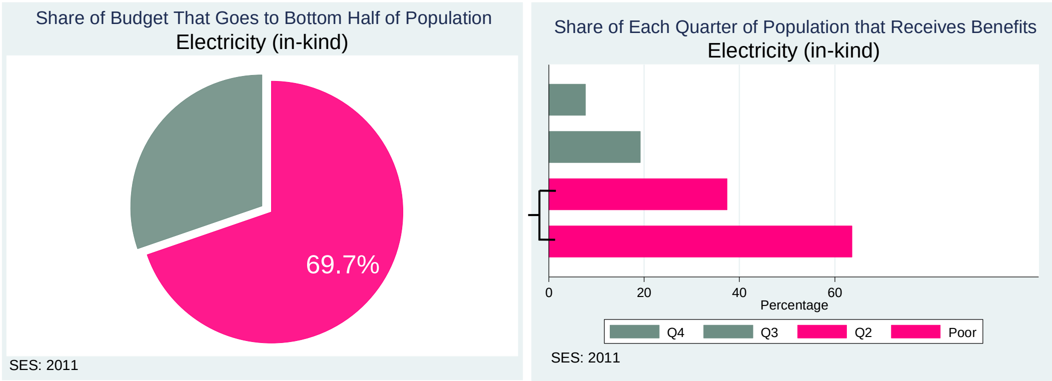
The pie charts used in this section only differentiate between the poor half and wealthy half of the population. Although a more detailed pie may be useful for in-depth policy makers – the current approach is much more appealing in its ease of understanding by politicians and other power brokers, when considering supporting or damning a policy.

4.4.3 Budget Share of Thai Energy Policies

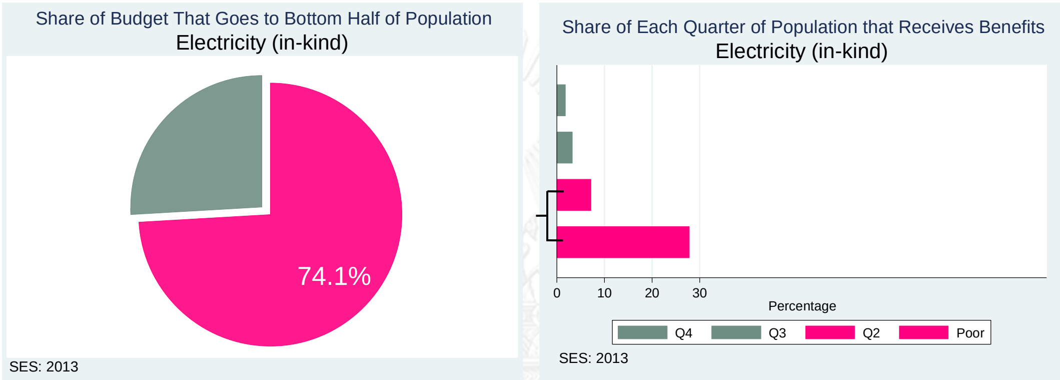
The following pie and bar charts track the beneficiaries of Thai energy policy. The goal is to see what share of the final benefits accrue to each income quartile. The analysis relies on SES survey data and does not include the administrative costs of the program, but rather tracks the benefits to the final recipients.

Figure 60 Electricity (in-kind) Budget Share and Use Share per Quartile, by Year

2011



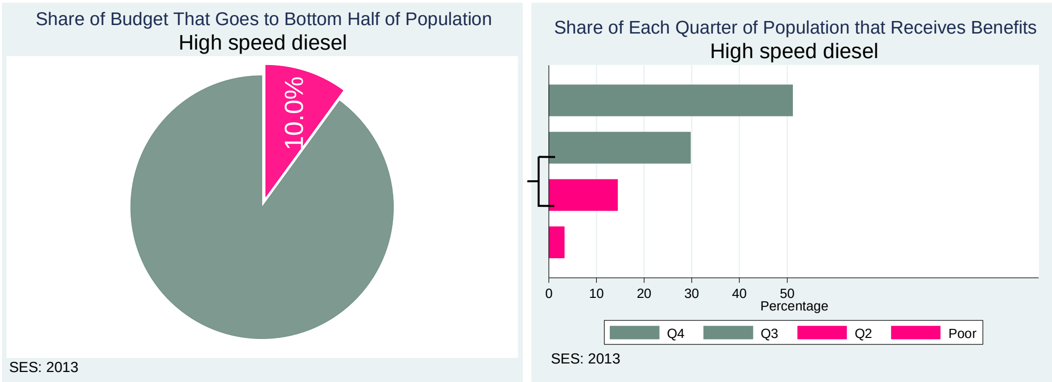
2013



NOTE: IN THE PIE CHARTS ABOVE, THAI HOUSEHOLDS ARE DIVIDED INTO TWO HALVES, THE WEALTHIER HALF AND THE POORER HALF. THE PERCENTAGE SHOWS THE PERCENT OF TOTAL PROGRAM BENEFITS (MONEY SPENT) GOING TO THE POORER HALF. FOR INSTANCE, IN 2011, 69.7% OF THE BENEFITS OF THE FREE ELECTRICITY PROGRAM WENT TO THE BOTTOM HALF OF HOUSEHOLDS IN TERMS OF EXPENDITURE. THE BAR CHARTS SHOW THE PERCENTAGE OF EACH QUARTILE WHO RECEIVED ANY BENEFIT AT ALL FROM THE PROGRAM WITH THE POOREST QUARTILE (25%) ON THE BOTTOM. THE BOTTOM TWO QUARTILES MAKE UP THE LOWER HALF OF THE POPULATION IN TERMS OF EXPENDITURE AND ARE COMBINED TOGETHER IN THE BUDGET SHARE PIE CHART ON THE LEFT.

As discussed above, starting in 2008, if a household used less than 90 KwH of electricity, they would not have to pay for it, except a nominal 40 baht usage fee. In 2013, the free electricity program was changed so that only electricity usage under 50 KwH would be given away free. This change meant that the government budget for the program was better targeted – i.e. 74.1 percent of the budget went to people in the poorest half of the population. But it also meant that it no longer reached many of its intended recipients. From the bar charts, we can see that uptake for the poorest quartile of the population was 65% at 90 Kwh and only 28% at 50 KwH.

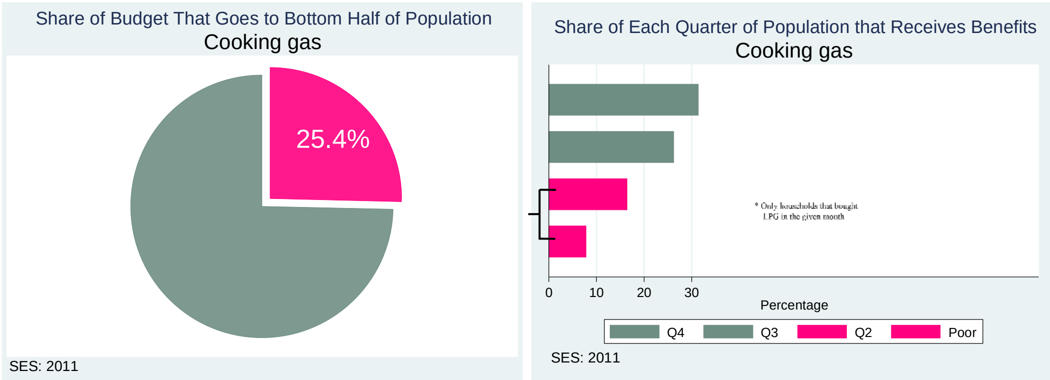
Figure 61 Diesel Budget Share and Use Share per Quartile, by Year 2013



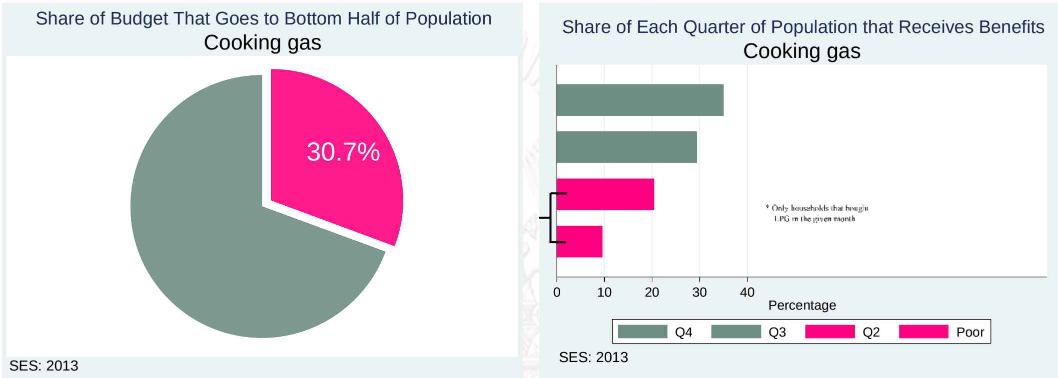
The diesel program was never intended as a program for the poor. Generally it is the wealthier families who own and use pickup trucks or farm equipment. Fifty percent of the wealthiest quartile benefited from this program. Only 10 percent of the subsidy went to the poorest half of the population. There is some pass-on effect as goods that poor people buy may be cheaper if diesel is cheaper, but poor people are much more likely to buy local goods or grow or collect crops themselves, so that cheaper transport costs will mostly go to wealthier groups.

Figure 62 LPG Budget Share and Use Share per Quartile, by Year

2011



2013



The subsidized cooking gas program was designed to reach the poor, but was not very effective at doing so. Depending on the year, a quarter to 30% of the subsidized gas went to the bottom half of the population. This is because many poor households were using charcoal or firewood, especially in the Northeast. There is a strong mismatch between those who use free electricity and those who use LPG Gas.

Table 26 Primary Cooking Fuel of Household Relative to Electricity Use

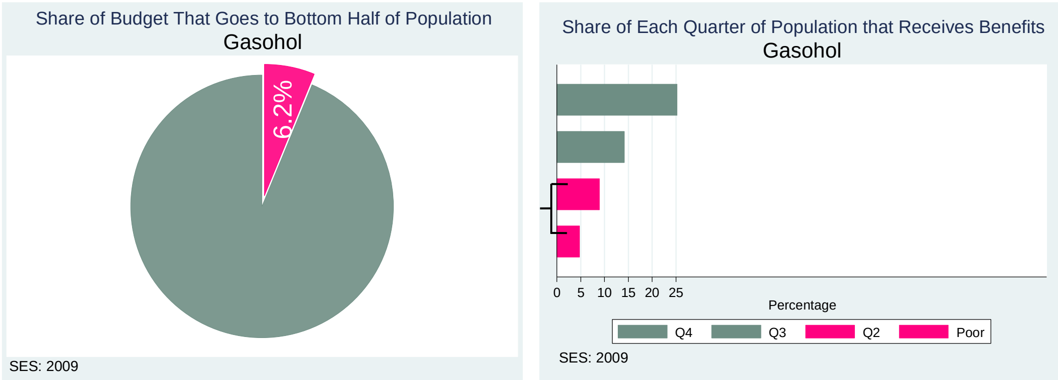
Electricity Use	Gas	Wood or Charcoal	Other
Use less than 50 KwH electricity	29%	63%	8%
Use less than 90 KwH electricity	39%	54%	7%
Use more than than 90 KwH electricity	75%	14%	11%

SOURCE: SES 2011

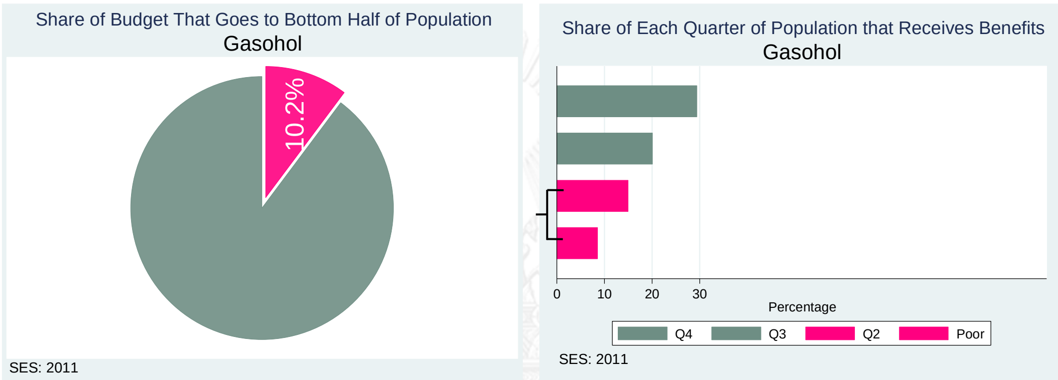
It was thought that electricity use would be a useful way to identify poor people eligible for other subsidies, as subsidies eligibility could be tied to a dwelling with proven low energy use. Unfortunately, for LPG subsidies, the target demographic is very different.

Figure 63 Gasohol Budget Share and Use Share per Quartile, by Year

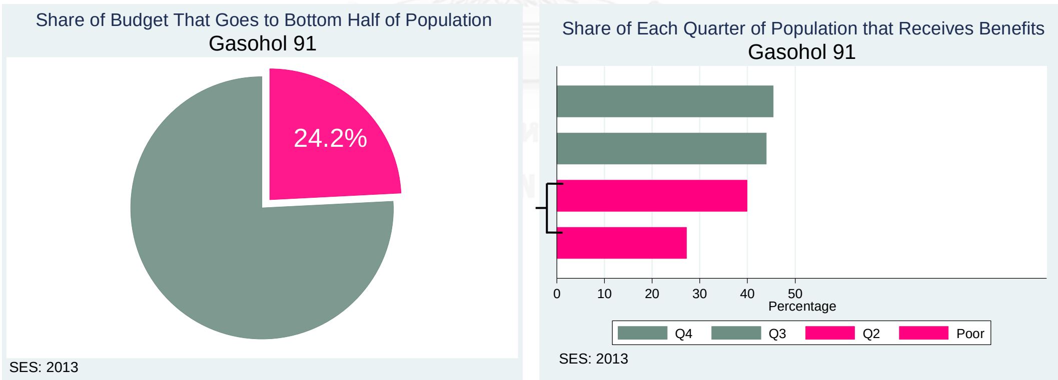
2009



2011



2013



Reducing the price of Gasohol (by reducing the tax) was a policy primarily addressed at encouraging the use of renewable energy. Adoption was very slow in the early years, and almost all of the usage was by the wealthiest quarter of the population, in fact the wealthiest 10% although it is not shown here. There was a widespread perception that gasohol could not be used in motorcycles and might hurt cars as well. The government took a strong arm approach to this problem and removed benzene 91 (unleaded gas) from the market in 2013, leaving people with no option except gasohol. At which point, usage of gasohol jumped. However, poor people drive much less than wealthy people and any policy aimed at cars and trucks will benefit primarily the wealthy. Even after eliminating benzene 91, only about 25% of the budget went to the poorest half of households.

Note: Data used for the pie charts depends on actually or in-kind expenditure from the SES. In-kind consumption is assumed to be the same as the subsidy for electricity. A very few wealthy people do receive large amounts of free electricity presumably from their companies, they are statistically minute in the SES data.
For other energy products, if the product is subsidized and a household buys it, it is assumed they get the benefit from the subsidy. The amount the household spends on the product times the share of the subsidy in the price of the product goes into the pie graph on the left, while whether the household uses any of the product, goes into the bar chart on the right.

4.5 Conclusions

This paper has addressed several different approaches to understanding subsidies over time. Section 4.2 looked at Energy poverty lines and considered several approaches to using them to identify who needs help from the government and when. It is possible to watch energy poverty lines over time and to use them as a justification for the government to step in and subsidize a hard hit group. Several approaches were considered and the appeal of a quantity based approach was significant, although in other parts of the paper, we stay with our goal of helping the poorest ten percent of the population, no matter how badly off they are or are not.
Section 4.3 looks at time lines of expenditure use that allow us to track who is receiving benefits from subsidies or other policies, what their income levels are, how uptake is changing over time, and finally suggests an appropriate level of subsidy as well.
Section 4.4 looks at the final amount received in subsidies by each income class, and suggests a criteria by which we can decide if a subsidy or policy is meeting its objectives. By looking at a series of pie charts over time, it is possible to determine if changes in policy or population are leading the policy in unexpected directions. These pie charts provide information that is not usually available to policy makers to help them justify or damn government programs and policies.

Chapter Five (Paper 3) Targeting and Distribution of Energy Subsidies

5.1 Introduction

Subsidies are generally designed to assist the poor or other needy group with basic necessities of living, acting as a form of social safety net. However, ensuring that subsidies reach the desired target without being used up by more wealthy groups is notoriously difficult. The primary problem is that wealthy persons consume more of just about everything, so unless the good is really unique to the poor, subsidies tend to be diverted. Success with targeting has been successful when the middle class would not use the subsidized good, such as with Kerosene for lighting (Indonesia and India) or LPG for cooking (India), but attempts to limit rights to poor groups have proven to be so difficult that the majority of commodity subsidies often get used by the wrong groups.

With the advent of big data tools, the opportunity to truly target subsidies has become much greater. It is now possible to have a much better idea of how poor families use goods, what channels they use, what other goods they buy as well, which makes the job of the policy maker considerably easier. In middle income countries, one thing we often cannot do is identify particular families that are poor. Knowing a family's financial status is difficult because there is no data collected about income in countries in which most individuals are paid outside the banking system. Nevertheless we can identify a typical poor person, and target our subsidy so it appeals to that person, much as a modern supermarket can use sales data to target certain types of customers.

5.1.1 Outline of the Chapter

Section 5.2 discusses different options for providing and subsidizing subsidies, including i) Helping the Poor (Targeted Subsidies) ii) Social Society (Give the Same to Everyone) ii) Social Safety Net (Self Selecting Subsidies) iv) Cash Transfers (Monetarizing Subsidies) Strengths and weaknesses of each are discussed. The focus of the rest of the paper will be on three approaches that practitioners can use to better target recipients and better distribute energy subsidies. Section 5.3 Discusses a Loop Model for identifying characteristics of subsidy users that can be used to target desired recipients. The model is used to predict LPG usage, but it is actually a general "big data" correlation model that makes use of forward stepwise aggregation of R-square to build the best predictive model possible. The model proceeds by an iterative process in which a type of loop is run, the results are analyzed for power and reasonableness, variables are included in the model, and another type of loop is run. The model could be used to identify any group in society. Section 5.4 discusses the ability to use Data Visualization to identify patterns in target groups that are very hard to notice using traditional econometric techniques. Some brief discussion of the legal issue of "who owns data" is included. Section 5.5 very briefly suggests several other ways of identifying groups of target recipients, such as classification trees and nearest neighbor techniques while section 5.6 provides a very brief conclusion section.

5.2 Four Ways to Distribute Subsidies

Table 27 Four ways to give subsidies

Nickname	Policy	Example
Help the Poor	Give the subsidy to just one targeted group, usually the poor	Free school lunch
Social Society	Give everyone the subsidy in equal amounts	Public libraries
Social Safety Net (Libertarian)	Let those willing to put up with the conditions get the subsidy	Soup kitchens
Cash Transfers (Authoritarian)	Give cash instead of subsidies to targeted group	Welfare payments

Table 28 Common objections to the four forms of subsidy

Nickname	Against	Counter-Argument
Help the Poor (Targeting)	Targeting doesn't work well in middle income countries, and in wealthy countries poor people self-identify as poor and don't try to better themselves	Even if there are cheaters, the poor need the subsidy so badly we should just tolerate
Social Society	No need to waste public money subsidizing wealthy households who don't need the money anyway.	Wealthy pay the most tax, so should get benefit too
Social Safety Net	Long waits in hospitals, filling out forms to get benefits, social ostrification - all are unnecessarily inefficient and reduce welfare.	Makes programs economical, people should have incentive not to get subsidy
Cash Transfers	Benefits to those in a club - Requires high level of information which excludes marginalized poor such as migrants, those working outside system	Economically efficient since people buy what they choose. Why subsidize migrants?

5.2.1 Inefficiencies from Subsidies

Inefficiencies from subsidies arise from a variety of causes. They may result from poor targeting whereby much of the subsidy goes to individuals outside of the target group. Much of the cost of subsidy programs goes to the middle and upper class as they consume more of the fuel. The inefficiencies could also be a result of subsidizing a good which gives ephemeral relief but cannot assist the target group in any long term or sustainable way. Fuel subsidies may reduce your cost for this month but improving the road or the transportation network could reduce your costs far into the future. Finally inefficiencies may result from an inefficient increased use of the commodity that result in negative externalities such as pollution or an excessive demand on foreign exchange.

5.2.2 Government Budget

Subsidies can have a very substantial negative impact on fiscal budgets. Especially with the run-up of energy prices over the last decade, the cost of subsidies has increased dramatically. Reducing the drain on government coffers has become a focus of government policy in many countries (IMF, 2013). In a recent publication, the IMF lists 8 countries in Southeast and South Asia that spend more than 8 percent of their government revenue on energy subsidies. In South East Asia, the size of subsidies has been a political issue in Indonesia, India, Thailand, and Malaysia among others.

Subsidies on consumption goods are most common in middle income countries that still have a large number of poor inhabitants, but have enough resources to try to help the poor achieve a better quality of life. Typically rapid development results in unequal growth and high income inequality, which is the rationale for reducing prices to help the poor. Consumption subsidies are often placed on basic commodities such as food and energy, and basic services such as transportation, health and education. As the cost of providing subsidies has soared, governments have responded by trying to back out of their positions by raising prices of subsidized goods to be in line with world prices. They have not had much success. Politically, subsidies are very popular with a large segment of the population, and a democratically elected government (Dansie, Lanteigne, & Overland, 2010)will find it hard to revoke them, since almost all voters will see the immediate negative effect on their own lives of higher prices, while not seeing the more hidden effect of higher tax burdens.

As second best options, governments have been experimenting with i) better targeting subsidies, ii) changing the form of subsidies so that they self-select their target clientele, iii) giving the same fixed amount of subsidy to each person in the country, or iv) substituting for the subsidies with some sort of lump sum transfer. All of these are ways to limit the liability of the government while still providing services to the most destitute persons.

5.2.3 Problem of excessive subsidies

"An energy subsidy is defined as any government action that lowers the cost of energy production, raises the revenues of energy producers or lowers the price paid by energy consumers." (OECD; IEA; OPEC; World Bank Joint Report, 2010)

Southeast and South Asian countries have generally adopted high levels of energy subsidies. This is particularly burdensome in countries that are net energy importers. The table below gives subsidy estimates from a report released in April (IMF, 2013)for major South and Southeast countries.

Table 29 Pre-tax Subsidies for Petroleum Products, Electricity, Natural Gas, and Coal (selected countries) (as a percent of government revenues)

Country	Petroleum Products	Electricity	Natural gas	Coal	Total
Bangladesh	7.56	22.12	13.45	0	43.13
Bhutan	1.39	n.a.	n.a.	n.a.	1.39
Brunei Darussalam	3.77	1.57	0	0	5.34
Cambodia	0	n.a.	n.a.	n.a.	0
China	0	0.68	n.a.	n.a.	0.68
India	6.75	1.72	0.9	0	9.37
Indonesia	14.51	3.69	0	0	18.2
Laos P.D.R.	0	n.a.	n.a.	n.a.	0
Malaysia	5.67	1.49	1.41	0	8.57
Myanmar	9.35	n.a.	n.a.	n.a.	9.35
Pakistan	1.02	10.23	19.89	0	31.14
Philippines	0	0	0	0	0
Singapore	0	n.a.	n.a.	n.a.	0
Sri Lanka	7.99	3.26	0	0	11.25
Thailand	0.66	7.24	0.61	1.08	9.59

Source: IMF 2013 (data generally from 2011)

Energy subsidies are primarily a middle income country problem. Most OECD countries do not subsidize, but rather, tax energy. Most poor countries cannot afford to subsidize energy. Middle income countries have sufficient resources to subsidize, and often face uneven economic development which gives the reason to subsidize energy. There is also a clear positive relationship between being an energy exporter and subsidizing energy. The political rational is that the resources belong to the country and should be available cheaply to the citizens.

5.2.4 Problem that Subsidies try to Solve

Although subsidies are expensive, few would argue that they should be phased out entirely. The cost of providing subsidies only to the poorest is very low since they use so little of total energy.

Figure 64 Engel Curve for LPG Use

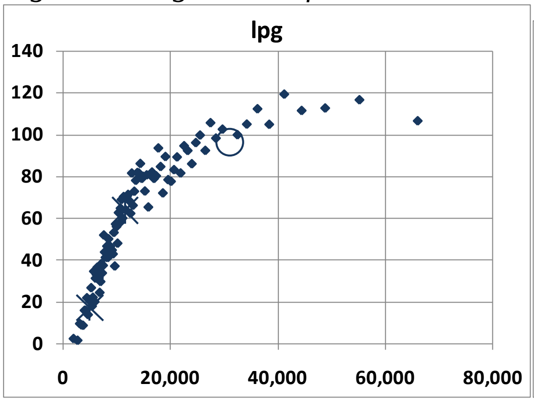
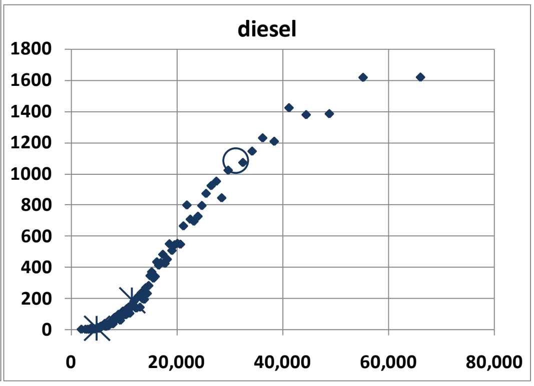


Figure 65 Engel Curve for Diesel



DATA SOURCE: SES 2009

Although data for compressed natural gas is not available in the Social Economic Survey, subsidies on other similar products are represented. The diagrams above use data from the Social Economic Survey (n=43,844) to show Engel curves for purchases of cooking gas and diesel fuel with income on the horizontal axis and expenditure on the good on the vertical axis. The two Xs on each graph show the location of a household at the 10th percentile in terms of income and a household at the 50th percentile, and the circle shows a household at the 90th percentile in terms of income. A key goal of a subsidy program should be to bring people in the bottom 10% of the population, below the bottom X, closer to the median in terms of energy use, or to the upper X. The median household has an income of about 12,000 baht per month and spends about 65 baht on LPG and about 180 baht on diesel fuel. There are about 20 million households in Thailand. To subsidize 10% of households to use 65 baht of LPG per month would cost 130 million baht a month or about 50 million USD a year, even if the government paid the entire cost of the LPG. However, much of the cost of subsidizing LPG comes from use by wealthier households, vehicles, and industry. According to (IISD, 2013), LPG for households in Thailand is cheaper than in any other major country in Southeast Asia.

5.2.5 Problem of Politics

The Thai government has been trying to remove subsidies on LPG and NGV for at least ten years, while the subsidies have existed for about twenty years. Whenever the idea is broached, there are objections by vested interest groups, often through protests in the city. There are frequent elections in a country that has seldom seen a government complete its 4 year term, and this leads political parties to retain current subsidies as well as add new programs for the poor by populous rural population.

In the latest episode of this ongoing challenge, the Thai government announced in late 2011 that they would allow the price of NGV and LPG to rise to market prices while sheltering some groups through special programs, in particular public transport through the energy credit card.

This experience has also been mirrored in Indonesia where the government has been trying to back away from subsidies on fuel for a decade. Although the current increase in price, bringing gas from half of the world price to two thirds the world

price may hold, if experience from the past five years is repeated, the government may need to back down again.
Finally state and national elections in India have led to a plethora of subsidy schemes that are hard to repeal once they are in place.

5.2.6 Problem of better targeting

If it is not possible to remove subsidies entirely, is it possible to reduce their scale so that they present less of a burden to the already stretched national budget? Subsidies in middle income countries are generally justified as a way to protect the poor from the effects of uneven development and income inequality (and less frequently but occasionally as a stimulus to national production). If the purpose is to help with the cost of the living to the poor resulting in high or volatile energy prices, how can it be arranged that the subsidy is more directly targeted to the poor? Generally it has been found that commodity and energy subsidies accrue mostly to wealthier households who can afford to use more energy.

1) Targeting

If the intention remains to only give the subsidy to the poor, some way of targeting the poor is needed. How to prove you are poor in a world without income tax, or in fact any formal receipt of payments or income for most of the country?
One (unsuccessful) option that has been tried, is to approach the village elders and ask for a list of names of who would qualify, Other ways include observing house characteristics, subsidizing by profession (for instance the energy credit card), or by family size.
Some countries have adopted an aggregate checklist as a metric of poverty. The BPL (Below Poverty Line) group in India was identified using the metrics in Table 5 below.
The Thai government recently agreed in principle to limit LPG subsidies based on those who already qualify for free electricity (i.e. people who use less than 90 KwH a month.)According to (Lewis, 2013a) this is a fairly ineffective method as most households that use less than 90 KwH do not use LPG, while nearly all of the poorest households do not use LPG. Furthermore the policy is biased against larger households. Nevertheless, the attempt is worthwhile in its attempt at targeting.

Table 30 Probability that a household uses LPG given certain characteristics

Below Poverty Line	19.6%
Wealthiest Half of Households	74.1%
Rural Area	52.4%
Average Household Size (3)	58.9%
Use Electricity Less than 90 KwH	37.6%

SOURCE: LEWIS (2013A)

Table 31 Example of Scoring Pattern, Socio Economic Survey 2002-2003, India

Sr No	Characteristic	Scores: 0 to 4 (Code)				
		0 (A1)	1 (B2)	2 (C3)	3 (D4)	4 (E5)
1	Size group of operational holding of land	Nil	Less than 1 ha. of un-irrigated land (or less than 0.5 ha. of irrigated land)	1 ha.-2 ha. of un-irrigated land (or 0.5- ha. of irrigated land)	2 ha. -5 ha. of un-irrigated land (or 1.0-2.5 ha. irrigated land)	More than 5 ha. of un-irrigated land (or 2.5-ha. of irrigated land)
2	Type of house	Houseless	Kutcha	Semi-pucca	pucca	Urban type
3	Average availability of normal wear clothing (per person in pieces)	Less than 2	2 or more, but less than 4	4 or more, but less than 6	6 or more, but less than 10	10 or more
4	Food Security	Less than one square meal per day for major part of the year	Normally, one square meal per day, but less than one square meal occasionally	One square meal per day throughout the year	Two square meals per day, with occasional shortage	Enough food throughout the year
5	Sanitation	Open defecation	Group latrine with irregular water supply	Group latrine with regular water supply	clean group latrine with regular water supply and regular sweeper	Private latrine
6	Ownership of consumer * durables : Do you own (tick (ï))	Nil	Any One	Two items only	Any three or all items	All items and/or Ownership of any one **
7	Literacy status of the highest literate adult	Illiterate	Up to primary (Class V)	Completed secondary (Passed Class X)	Graduate / Professional Diploma	Post Graduate/Professional Graduate
8	Status of the Household Labour Force	Bonded Labour	Female & Child Labour	Only adult females & no child labour	Adults males only	Others
9	Means of livelihood	Casual labour	Subsistence Cultivation	Artisan	Salary	Others
10	Status of children (5-14 years) [any child]	Not going to School@ and working	Going to School@ and working			Going to School@ and NOT working
11	Type of indebtedness	For daily consumption purposes from informal sources	For production purpose from informal sources	For other purpose from informal sources	Borrowing only from Institutional Agencies	No indebtedness and possess assets
12	Reason for migration from household	Casual work	Seasonal Employment	Other forms of livelihood	Non-migrant	Other purposes
13	Preference of Assistance	Wages Employment	Self Employment	Training and Skill Upgradation	Housing	Loan / Subsidy more than Rs. one lakhs or no assistance

Source: (Commissionerate of Rural Development Govt of Gujarat, 2002)

Notes: * B/W TV, Electric Fan, Pressure Cooker, Radio, ** Telephone, Fridge, Colour TV, Computer, 2/3 Wheeler, LMV, Tractor etc, @ including Non formal Education, Note : The Total Score for a household will vary between 0 and 52. A score of 17 or less is BPL (Below Poverty Line)(Commissionerate of Rural Development Govt of Gujarat, 2002)

Smart cards are an ideal way to target subsidies once the recipients are known, but they do not help with identifying the poor.

2) Self-Selection

One of the most humane ways to address the issue of subsidies is by allowing people to self-select service based on convenience. In Bangkok, the whole public bus fleet is subsidized, but there are many levels of service. There are infrequent and crowded free buses, there are non-air-conditioned busses that cost about 7 baht or 21 US cents. There are air-conditioned public buses with service from about USD 50 cents to USD 1 dollar. Finally, moving to the private sector are taxis which are very cheap compared to other countries, costing from 1 to 3 US dollars for a typical trip. The fuel for Taxis is subsidized at about ½ of its delivered cost. Fares are also fixed. In this way, passengers choose the level of service and comfort that they prefer without any conditions on who uses what mode of transportation. Self-selection as a mechanism for controlling subsidies is very popular in Thailand, and can be found in health care, and schooling, as well as transportation.

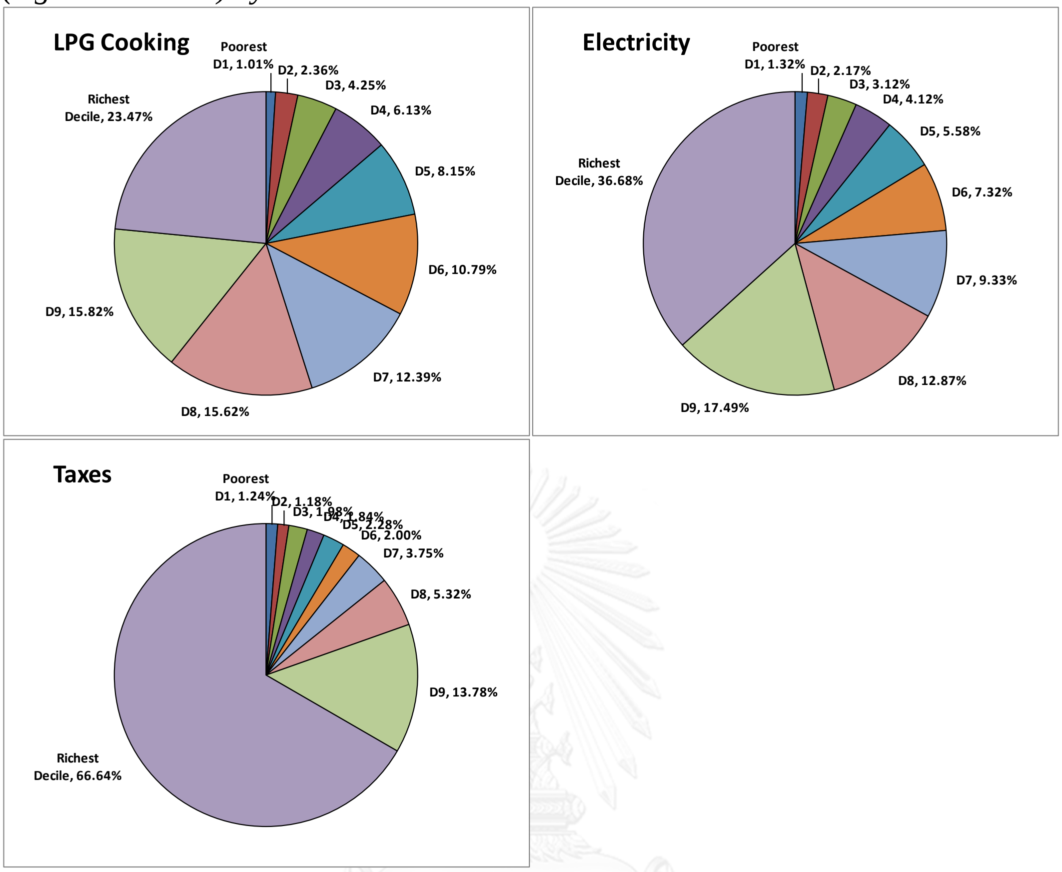
Another example from around the region would be subsidies on LPG from Indonesia that depend on the size of the container, with 3 kg tanks being subsidized the most. Ten kg tanks have a higher price, and 30 kg tanks have the highest price. The idea is that the inconvenience of changing tanks frequently will keep wealthier households from using the smaller tanks.

Tunisia is also known for successfully implementing this self-selection type of subsidy.

3) Give the Same to Everyone – One Size Fits All.

Subsidies for the poor have an element of unfairness about them, in that those who pay for them, are not eligible to receive the benefits. Wouldn't it be fairer to give equal opportunity to everyone in the society, and have people pay as they are able? On the other hand, when subsidies are available indiscriminately, because of higher purchasing power, most of them are absorbed by the middle class and upper classes, subverting the original intention of the subsidy to provide a social safety net, and miring the government coffers in debt.

Figure 66 LPG Used for Cooking, Electricity Used by Households, and Direct Taxes (e.g. Income Tax) by Income Decile



SOURCE: AUTHOR, FROM SES 2009 HOUSEHOLD LEVEL DATA. THE TAXES REFERRED TO HERE ARE DIRECT TAXES AND EXCLUDE VALUE ADDED TAXES AND CORPORATE TAXES (TO SOME DEGREE PASSED ON TO THE CONSUMER) WHICH WOULD MAKE TAX INCIDENCE SOMEWHAT MORE EVEN.

From the above figure it is clear that although the richest decile uses the most of each type of energy, thus receiving the most subsidy, they also pay the most taxes. To exclude them from receiving the subsidy seems unfair.
One alternative is to give the subsidy to everyone in exactly the same quantity. Each person in the society would have the right to so many kilos of LPG or so many kilowatts of electricity per month. This gives everyone the same opportunity, but provides all of the need from the poor, and only a part of the need from the wealthy.

4) Cash Transfers

The 4th way to manage subsidies is to transfer them into a single cash payment. Iran is the best known example of this. Trying to cope with among the highest energy subsidies in the world, in 2010 Iran decided to do away with subsidized prices and instead give a cash transfer to each of its citizens. "On December 19, 2010, the government [of Iran] ended the decades-long subsidy program for bread and energy products like gasoline, and replaced it with direct payments of about $45 per month per individual…Two features of the program have political salience, prices of bread and all energy products were increased to market levels in one shot, and at the same time money was transferred back to people's bank accounts."(Salehi-Isfahani, 2011)

This subsidy was the same for each person, so that it might make up 50% of the income of a family in the bottom decile, or 5% of the income of a wealthier one in the top decile. In the aftermath of the program, Iran has experienced extremely high rates of inflation tied both to the increase in price of commodities and to increased purchasing power of the poor. The cash transfers have made the poor better off, but the middle class worse off. (Shabani, 2013)

Countries have often tried to replace an inefficient and poorly targeted energy subsidy (which tends to increase consumption) with a non-distortionary cash payment that can be used as the recipient sees fit.

Other countries such as Brazil (Bolsa Familia) and Mexico have had success with this approach.

5.3 Big Data Loop Model

5.3.1 Efforts to Identify Poor Households

5.3.1.1 Overview

For half a century, development economists have been trying to identify poor households so that programs for the poor can be better targeted. The problem seems simple, but in a society like Thailand in which only 20-30% of workers receive a formal paycheck, or in India where 97% of transactions are in cash, it is very difficult to identify the poor. It is relatively easy to prove you are rich by showing your bank balance, but much harder to prove that you are poor. That is not to say that it is impossible, just very time consuming and expensive to do so.

A paper published in the American Economic Review (Alatas, Banerjee, Hanna, Olken, & Tobias, 2012) tried to test the best way to identify the poor by looking at physical characteristics of households, and by consensus opinion of village elders and a mixed method, but the conclusion seemed to say that the job is difficult.

5.3.1.2 Identifying the Poor Using the SES

The question addressed in this chapter is – is it possible to identify the poor by looking at the way they behave, and then use that to better target subsidies? Of course it is possible to see which households are poor simply by looking at their expenditure in the SES. However, households chosen to be included in the SES are confidential and make up only a small sample of the population.

Suppose, however, that we do not try to identify particular households, but only the "characteristics" of poor households. This can give us a wealth of information about better ways to reach this group without specifically identifying them. Likewise, we can find some characteristic of poor households that is observable and verifiable.

The first thing to know about poor people in Thailand is they are old. Age is something that is measurable and verifiable, and if a cash transfer program were to be developed in Thailand it is likely should be built around the pitiful existing government pension program of 700 baht per person a month, about 1/5 of the official poverty line. (Thailand has a near zero unemployment rate, but a young retirement age of only 60 years.) Old people do not travel much, so the implication would be Not to subsidize transportation fuel.

Now another developing country may have a different pattern of poverty and energy use so that a different method of targeting subsidies would be appropriate. But the point is that it is very easy to find by using big data techniques to build a portrait of specific groups.

The following section describes an experiment using big data to identify a particular group of families, in particular those who make use of the LPG subsidy. The variable under investigation are households that use LPG as the primary source of cooking fuel. This status is self-reported, but the household would have no reason to prevaricate.[5] This study will serve as an illustration of a technique that could be used to investigate the subsidy of any product, or any household type of interest.

Often when we are engaged in econometrics we want to know the overall effect of an independent variable and a dependent variable. However, the object of this research is to identify specific individuals in the sample who are affected by the variable. This is similar to what happens in marketing in which the marketer attempts to identify individuals who would be interested in a product. In this case our objective is to identify individuals who benefit from a policy so that we could rationalize the subsidy process.

We may want to redesign the subsidy to effectively reach the target group without reaching other groups. This could be done by searching for commonalities that would allow us to identify those belonging to the target group but exclude those who do not.

Or, we may want to redesign the survey so that recipients will self-select to use or not use the subsidy. This selection may be due to convenience versus cost, social perception (sell the commodity in a low social class shop)

Or, we may want to redesign the subsidy so that it is more sustainable e.g. roads / fixed cost.

Or, we may want to redesign the subsidy to offer something equally desirable but cheaper to a group who benefits from the subsidy.

5.3.1.3 Predicting Households That Use LPG

The objective of this section is to introduce a “Loop” model that uses a series of loops through types of regressions searching for the best predictive model of a binomial objective.

The overall strategy will be to use a training dataset to build a predictive model based on regressions of the binomial on all available continuous variables, categorical variables, and deciles of values seeking always to increase correlation or for increased convenience, R2, or correlation squared.

This type of model may partially reproduce the results of large correlation models, but has the advantage of allowing an understanding of the underlying causal relationships that more “black box” methods such as a stepwise forward regression do not allow. It is a fortunate compromise for policy makers who need to know both “why” as well as “who” in designing subsidies.

The “training” data used comes from the 2009 social economic survey (SES) produced by the Thai National Statistics Office. The National Statistics Office (NSO) conducts an annual Household Socio-Economic Expenditure Survey (SES) augmented with demographic and household characteristics. The survey is a

[5] The survey is confidential, it is completely separate from attempt to identify subsidy users, and the subsidy is universal at any rate – it is not targeted.

carefully designed stratified representative survey using sample weights to readjust household sampling probabilities. Every other year (odd years) questions about income and employment of household members are also included in the survey. The survey is conducted yearly, but a different set of households is chosen in each year. Sample size is 43,844 households (about 0.2% of all Thai households) of which 59% use LPG as a primary fuel for cooking. Test data was a similar sample of households from 2011. The objective is to build a model that can predict whether or not a household will use LPG with 80% accuracy in test data. This model should be mostly independent of the original training data (little or no over fitting) so that it can be used with fresh instances of the same variables without much loss of precision, which is why the 80% test is used on a fresh data set.
The naïve predictive model is 59% accurate. If we predict all households use LPG, we will be correct 59% of the time, with 100% accuracy about households that use LPG and 0% accuracy about households that don't use LPG. This illustrates a key feature of this sort of model – we need to predict those who don't use LPG as well as those who do use LPG.

5.3.2 Deductive Model

The first step is to build a preliminary deductive model. This will allow us to informally compare the two approaches. Details of the deductive model are hereby given.

5.3.2.1 Dependent variable

There are three candidates for the dependent variable for our model available in the SES. All relate to LPG, but in different ways. The candidate variables are as follows.

1) Bought LPG during the month of survey
2) Use LPG as (primary) cooking fuel
3) Own an LPG stove

As an expenditure survey, the first is the obvious choice, but unfortunately the 15 KG LPG tank widely used in Thailand lasts for several months and the survey asks households to report only expenses in the previous month. The third variable is also tempting in that it seems to be a form of revealed preference, that does not rely on self reporting. However, it appears that owning a stove does not necessarily imply use of the stove. Finally the second option was used in the model: Use LPG as (primary) cooking fuel.

5.3.2.2 Choosing Independent Variables

Some possible explanatory variables were considered for several reasons. It was hypothesized that income would likely have an effect on LPG use. Familiarity with the data suggested a simple above the average income and below the average income. (Actually the relationship between income and LPG use is not linear.) A second income variable Below Poverty Line was included to help address policy questions. The Thai poverty line was 1,586 baht per person per month in 2009. (NESDB 2009) Rural-Urban was proposed as lifestyles and eating styles are quite different between the two. Region was proposed as a variable but abandoned. The Thai administrative regions do not easily classify LPG users. Although province was later added as a categorical variable, later work with data visualization showed that the real regional classifier should be mountainous versus plains. Household size, of course matters to how much energy is used, and families with children may cook in different ways than families without. Similarly the original theory also suggested a variable for older persons for the same reason. Households that use Less Than 90 KwH was included as a policy variable, as there was a proposal at the time to tie LPG subsidies to only low electricity households.

- Below Poverty Line
- Income in Top 50% of Households
- Rural – Urban
- Region
- Household Size (additional household members)
- Children under 15
- Adults more than 60
- Electricity Less than 90 KwH

Eventually the model was reduced after some preliminary investigation to the following:
Final Model (with predicted signs)

$$CookLPG = \beta_0 + \beta_1 BelowPov + \beta_2 Wealthy + \beta_3 Rural + \beta_4 HHSize + \beta_5 ElectricLT90 + \mu$$

- **CookLPG** - if household uses LPG as primary cooking fuel
- **BelowPov** – if per capita income in household less than 1586 baht (expect negative)
- **Wealthy** – if household is in top half of income per capita (expect positive)
- **Rural** – if household is in rural area (expect negative since easier to obtain firewood)
- **HHSize** – [0 to 16, mean=2.18] number of additional people over one in household (expect positive since increasing returns to scale in cooking)
- **ElectricLT90** – electricity less than 90 KwH, gov't gives this much free (expect negative since tied to poverty, small HH)
- (All variables are binomial except HHSize, which is count)
-

5.3.3 Results of Probit Deductive Model Regression

Probit regression		Number of obs =		43844	
LR chi2(5) =		4607.62		Log likelihood = -10867643	
Prob > chi2 =		0.0000		Pseudo R2 = 0.1793	
cooklpg	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]
belowpov	-.784	.044	-17.64	0.000	-.871 -.697
wealthy	.566	.022	25.22	0.000	.522 .610
rural	-.134	.019	-7.13	0.000	-.171 -.097
hhsize	.112	.007	15.43	0.000	.098 .127
electricLT90	-.782	.022	-36.18	0.000	-.824 -.740
_cons	.248	.031	7.89	0.000	.187 .310

Signs are all as expected and significant and negative coefficients are large on belowpov and electricLT90. People are less likely to use LPG as a cooking fuel if they are poor, and if they use little electricity, and if live in rural area (since can use firewood instead) Probit results are not marginals as they would be with OLS, so they cannot be used directly to calculate marginal changes in probability. We can multiply the coefficients by the cdf over a range of one near the point of interest to find the actual slope / marginal at that part of the regression. The pdf of a single point is 0, but the sum of pdfs from t = -.5 to .5 gives a cdf of .38 which is what you want to multiply the probit coefficient to find the marginal at the mean level which is very close to the OLS coefficient. We can choose other ranges of one (e.g. t = .5 to 1.5) to find about .24 to find marginal at other places along the probit regression.

Is it a good model? We can use two methods to evaluate this -
Method 1: Pseudo- R^2 = .18 Roughly, predictions are better than NO model by 18%.
Method 2: Predictions from model? Naive – No model (In pop, LPG = 59%, so predict use LPG)

Predict that use LPG True = 100%
Predict Don't use LPG True = 0%
Weighted average correct = 59%

Model predictions

Predict that use LPG True = 76%
Predict Don't use LPG True = 62% <-BIG IMPROVEMENT
Weighted average correct = 69%

We can also find Marginal Results at the means (what we normally get from OLS)
We are measuring the average effect of switching a binomial value from 0 to 1.
Imagine we start at probability=.59

	dy/dx	Std. Err.
belowpov	-.246	.014***
wealthy	.178	.007***
rural	-.042	.006***
hhsize	.035	.002***
electricLT90	-.246	.006***

These values are at the mean, so values of coefficients will likely be very close to regression coefficients from a linear probability model we would have got if we had used a simple OLS model instead of probit.
However, it is possible to calculate Marginal Results at a particular point we are interested in. The following is for a below poverty line family that has 4 family members, lives in a rural area, and using very little electricity.

	dy/dx	Std. Err.
belowpov	-.189	.005***
wealthy	.137	.008***
rural	-.032	.005***
hhsize	.027	.002***
electricLT90	-.189	.009***

Essentially all coefficients will be reduced as the probability of using LPG at this income level is low so the marginal effects are reduced as well.
Average probabilities of using LPG
Our model predicts the probability of each household using LPG. We can find the average P_probit for each dependent variable.
Probability of using LPG (Avg P_probit)

belowpov	19.6%
wealthy	74.1%
rural	52.4%
hhsize (3 persons)	58.9%
electricLT90	37.6%

5.3.4 Sensitivity Analysis

Sensitivity analysis can be used to test how much our model depends on the particular specification of variables we chose. If the variables are close to independent we would expect that their values would not change much as other variables were added or deleted from the model. Below, some alternative model specifications are compared.
Model 1: probit cooklpg belowpov wealthy rural hhsize electricLT90 [pw=a52]
Model 2: probit cooklpg expendpc expendpcsq rural hhsize electricLT90 [pw=a52]
Model 3: probit cooklpg ib(last).decilepc rural hhsize electricLT90 [pw=a52]
Model 4: probit cooklpg ib(last).decilehh rural hhsize electricLT90 [pw=a52]
Model 5: regress cooklpg belowpov wealthy rural hhsize electricLT90 [pw=a52]

Results were not very sensitive to specification

Model	Model 1	Model 2	Model 3	Model 4	Model 5
Criteria 1: Pseudo-R2	.18	.14	.19	.18	-----
Criteria 2: Predict Use LPG	76%	79%	77%	78%	76%
Criteria 3 Predict Don't Use	62%	57%	61%	61%	60%

5.3.5 Inductive (Loop) Model

If the naïve model gives us predictive power of 59%, and the deductive model gives us predictive power of 69%, the question becomes "Can we do better?"
In the following section an Inductive model is developed that does just that. A deductive model, as used above, requires the researcher to include only variables thought to be important ex ante in determining the state of the Y variable. An inductive model includes all possible information and searches for the best fit. Classification trees, forward step-wise regressions and the loop model included here are all examples of inductive models.

Table 32 Comparing Problems of Small Data and Big Data

Small Data vs.	**Big Data**
Problem of the Past	***Problem of the Future***
Finding statistically reliable estimates with few data points	Finding efficient ways to extract information from large data sets
Advantage of the Past	***Advantage of the Future***
Easy to understand data	Lots of data adds new information
Deductive Reasoning Possible	Inductive Reasoning adds value

SOURCE: COMPILED BY AUTHOR

The reasoning behind using inductive models, when possible, is that there are limits to the human brain's ability to process and accurately include predictors. We may miss some variables that should be included because we didn't think of them, or include unimportant variables because of a mismatch between theory and reality.
The danger of using inductive models is over fitting. Over fitting is commonly encountered with small data sets when simply by chance one variable correlates with another, such as sunspots and the stock market. There is no relationship of course, but because of small sample sizes, a surprisingly large share of variables will appear to have large correlations. Over fitting is reduced by increasing sample size. That is why regression fishing is strongly discouraged in traditional econometrics. As the sample size increases, these random relationships become vanishingly small. Over fitting can also be strongly reduced by testing the model on data that was not used to create it.
A second issue around using inductive models is that it can be increasingly hard to understand about causality in the model. In a deductive model, causality was implicit

in the original specification of the model, since we only added variables that theory suggested. Inductive models are much less clear about causality relationships, in some cases not even showing what variables were included. Think of HAL the computer saying "There is a 7% chance the spaceship will survive re-entry." The inductive loop model outlined here, although requiring more work than a fully automatic model, will allow the researcher to understand most causal relationships involved.

5.3.6 Process

The first step is to run simple regressions of each variable included in the SES against the dependent variable. We are looking for variables with a high degree of correlation (R^2) with the dependent variable. High correlation will give us predictive power. We are not especially interested in statistical significance. Statistical Significance is generally not useful in Big Data applications. In fact we will find that most variables (about 80% in our case) will be statistically significant. Statistical significance depends on sample size, and as the size becomes large enough, tiny differences (in means, for instance) will be statistically significant, but still not have predictive power. In the formula below, as n becomes large, standard error (s) approaches zero, giving us a statistically significant result with possibly very low explanatory power.

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Although the survey data set is 43,844 households, not all of our regressions will have that same sample size. In fact, whenever the household does not buy a certain good, data for that good will be denoted as a missing value in the survey data set. Therefore the original data set does not have zero values, but it does have many missing values. Data is carefully and completely recorded, so that for expenditure items, a missing value refers to zero expenditure, not a lack of information about whether that variable was purchased.

Suppose only 245 families received soap and shampoo as a free good. Records with missing values will not be included in regressions, so the sample size of regressing free soap and shampoo on cooking with gas is 245. Overfitting and spurious relations are definitely possible with this small sample size.

Below are Monte Carlo experiments performed by the author to see how often an R^2 of .02 or above could be expected in randomly generated variables. A random generator was used to generate random Y and X variables which were regressed against each other. The experiment was repeated 5,000 times for each sample size to estimate frequency of spurious .02 R^2 results. Results follow:

Figure 67 Results of Monte Carlo Experiment to Test for Spurious regressions with different sample sizes.

Sample Size	Percent R^2 > 5%	Percent R^2 > 2%	Percent R^2 > 1%
50	12.1%	33.4%	47.5%
100	2.7%	16.9%	32.1%
300	~0	1.5%	9.0%
500	~0	0.18%	2.2%
1000	~0	~0	0.14%

The risk of spurious causality is much reduced as sample size passes 500 and 1000 records. Another approach is to use the binomial sample size rule of the thumb from statistics. A sample is "sufficiently large" if it meets the following criteria:

$$from\ theory:\ n_0 = \frac{(\text{critical value})^2 * (\text{prob.}) * (1 - \text{prob.})}{(\text{sig. level})^2} = ?$$

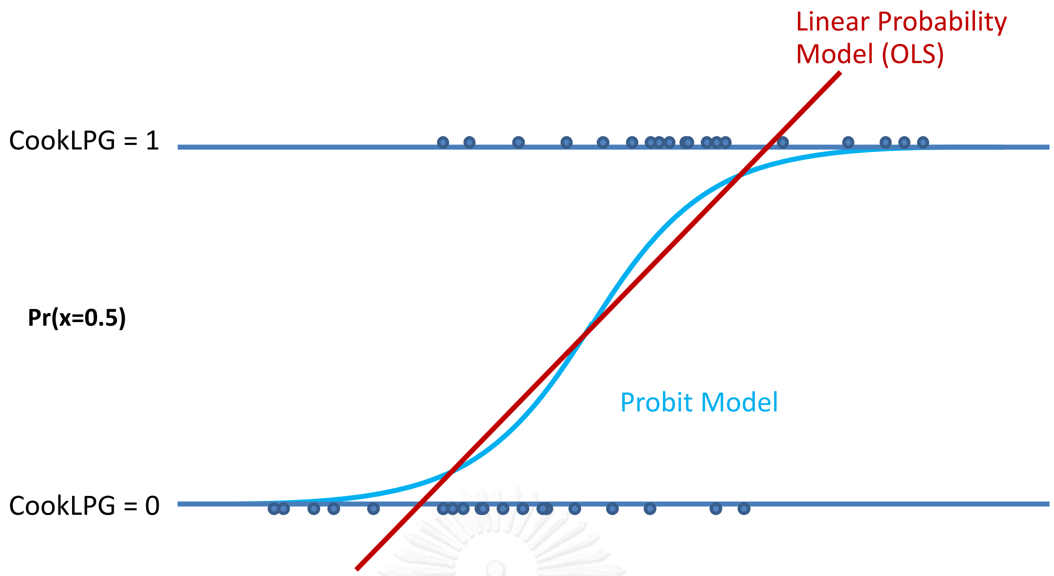
Or in our case,

$$for\ us:\ n_0 = \frac{(1.96)^2 * (.59) * (.41)}{(.05)^2} = 372$$

To reduce as much as possible over fitting in the model, every variable with a sample size less than 500 was eliminated automatically from consideration in the model.

Although the final objective is a binomial probit model, preliminary regressions are run using a linear probability model to reduce computational burden and computing time. Only final versions of the model will be evaluated using the probit model. Near the mean, values of each should be about the same, although standard errors may differ.

Figure 68 Comparison of Probit and Linear Probability Model Regression Results



In the figure above, the slope of the straight line (linear probability model) is near the slope of the center of the probit regression – (probit marginals evaluated at the mean.) In a probit regression the slope changes with the value of X which is why we don't get coefficients directly without specifying a point of interest. Therefore we can use the straight line as a rough approximation of the blue line until arriving at the final stages of the project to save time and computer power.

5.3.7 Choosing Variables for the Inductive Model

The approach used was as follows

- Cycle using all variables 1 by 1
- Cycle using income plus one by one.
- Cycle using categorical variables
- Cycle using decile versions of all variables
-

The first pass of everything uses the linear probability model. Variables that add significantly to R^2 will be kept. This model is built to work with any dataset.

Loop 1

Our first step selects variables with adequate sample size. If the sample size is too small by chance the dependent variable might be correlated with the independent variable. The first loop eliminates variables with a sample size of less than 500.

Loop 2

A loop that searches for an R^2 of 0.05. The second loop cycles through about 300 variables in the SES database looking for anything that gives an R^2 of 5%. Generally the t-stat is above 10, the lowest was 4.67 for this group. Operations are hidden until such a variable that gives a sufficiently high R^2 is found, after which that particular regression is printed. All successful regressions (38 of 279) were examined to try to understand the reason for their significance. Examination of results led to the following conclusions:

1) Examination suggests that many variables may be acting as proxies for income.
2) Some household and some demographic variables are significant.
3) Some regressions are specious such as

reg dgas
reg dgas hh10 (dgas=1 if hh10=4)

4) Some regressions are too obvious

reg dgas hh23 (do you own a gas stove)

Conclusion from Loop 2: Need to include INCOME in the regression and then to see if other variables still add predictive value.

Loop 3

A loop including an additional ***income*** term that searches for an R^2 of 0.1085

- Income by itself has an R^2 of 7.85%
- The third loop includes household income (a07) and looks for any variable that can add R^2 of 3% to the 7.85% from income.
- Two addition regressions are needed as checks – since the sample size varies in each regression, we want to see what the effect of income alone and of the variable alone would be for this particular sample.

Examination of results leads to the following conclusions:
1) Many variables can be dropped since they add little, after income is included.
2) It is clear that household financial assets are an important variable besides income, and should be added to the model.
3) It is clear that some household variable needs to be included as a proxy for a stable, permanent household

- Washing machine
- Hooked up to commercial water supply
- Own land and house
- Various other household variables are moderately significant

Conclusion: Add FINANCIAL ASSETS to probit model and also add some measures of HOUSEHOLD STABILITY

Loop 4

A loop that searches for an R^2 of 0.1 for categorical variables
Many variables in the SES are categorical so that each value signifies a different profession or a different region. In other words, the values within the variable are not ordered, and each number represents something different.
The fourth loop searches for such variables (first line) and then tests for an R^2 of 10% for all values of each variable against dgas.
Examination of results leads to the following conclusions:
1) The location of the household has a strong effect on whether they use LPG or not. Southern households do use it, NE households do not use it. Region does an okay job of capturing this, but province is significantly better.
2) The type of ownership of the house is important adding to the stability of the household as a significant **factor.**
3) Profession and education do not add any predictive value to the model
Conclusion: Need to include PROVINCE and HOUSEHOLD TYPE into the regression.

Intermediate results

By adding in the suggestions from the last three loops we can modify the model to better predict LPG usage. Of course the final test will come when we are done with deciles and interaction terms, and when we test using 2011 data. However, intermediate results give reason for hope. The model is now:

$$CookLPG = \beta_0 + \beta_1 BelowPov + \beta_2 Wealthy + \beta_3 Rural + \beta_4 HHSize + \beta_5 ElectricLT90$$
$$+ \beta_6 Wealth + \beta_7 WashMachine + \beta_8 WaterHookup + \beta_9 i.\Pr ovince + \beta_4 i.HouseOwner + \mu$$

This model was found to predict whether a household used LPG with **75.4%** (was 69%) accuracy, predicting correctly 81% of the cases that did use LPG and 70% of the cases that did not use LPG.

Loop 5

A loop that searches for an R^2 of 0.15 using deciles of each variable

Often there is not a linear relationship between the dependent variable and the independent variable. For instance a poor family does not use gas since it is expensive, while a rich family does not use gas as it is an inferior substitute for electricity. This gives an irregular shaped pattern which is hard to capture with OLS. This loop creates deciles for each variable and searches for a significance or R^2 of 15% using all deciles for each variable in turn against dgas.

Examination of results leads to the following conclusions:

1) The kinds of food a households buys is tied to the use of gas stoves. It is possible to predict gas stoves by looking at households who buy fish, grains, fruits and vegetables. These variables were added to the model.

Conclusion: Need to include FISH, GRAIN, VEGETABLES AND FRUIT into the regression.

The final results of using an inductive model to predict LPG use is that we can predict LPG use with 80.4% accuracy using the original training data. Accuracy diminishes slightly to 79.3% when we use Test Data – a new untouched dataset.

Predictions (compare to Naïve = 59% and Original Probit = 69%)
Predict Do use LPG True = 84% ***Using Training Data 2009***
Predict Don't use LPG True = 77%
Weighted average correct = 80.4%

INDEPENDENT TEST DATA

Predictions (compare to Naïve = 61% and Original Probit = 68%)
Predict Do use LPG True = 84% ***Using Test Data 2011***
Predict Don't use LPG True = 75%
Weighted average correct = 79.3%

5.3.8 Conclusions

Some Statistical Problems with Big Data

- **Statistical significance (t-stat) not very useful** since almost everything is significant with big data, need to use minimum R^2 instead.
- **Changes in Alternative-R^2 not useful** as test for discarding variables since will barely decrease, need to use whether variables add to predictive value instead.
- **Non-linear relationships are likely, but traditional solutions are ad hoc**. Non parametric solutions such as deciles may be a substitute.
- **Over-fitting is a danger**, so using separate data for designing and testing model is important.
- **Measurement errors in original data** due to poor survey design are a concern that cannot be solved by increasing sample size. Increased care is needed in collecting data.

Without any model the naïve prediction of which households will use LPG would have 59% accuracy.
With just a Probit model the prediction improves to 69% accuracy.

In the Intermediate Stage, the Model was able to predict which households use LPG for cooking with 75.4% accuracy using a combination of Probit Models and Big Data and Data Mining Techniques
Final Stage, the model was able to predict which households used LPG with 80.4% accuracy. When used with Test Data (fresh data that was not used in the creation of the model) a small amount of predictive power was lost, and the model could predict with 79.3% accuracy.
The Final Loop Model is presented below in an illustrated form. The box shows the original deductive form of the model.

$$CookLPG = \beta_0 + \beta_1 BelowPov + \beta_2 HighInc + \beta_3 Rural + \beta_4 HHSize + \beta_5 ElectricLT90$$

$$+ \beta_6 WashMachine + \beta_7 PipeWater + \beta_8 HouseType + \beta_9 FinWealth + \beta_{10} Province$$

$$+ \beta_{11} Fish + \beta_{12} Grains + \beta_{13} Fruits + \beta_{14} Vegetables + \mu$$

5.4 Data Visualization and Public Data Availability

5.4.1 Introduction

The world is awash with data but big data is an impenetrable jungle, its message visible only to a few. Data visualization may provide a flexible window into that jungle, especially when public access is limited.
The special concern of this section is with those organizations maintaining libraries with their own source of data, especially government departments and ministries. How can the great quantities of data available through these ministries be made accessible to the public?

5.4.2 Data Ownership

Who does data belong to? Government data was collected from the citizens, by government agencies, using tax payer funds. The data should belong to us all. However, in its disaggregated form it is seldom released to the general public for a variety of reasons.

1) Privacy concerns – some parties may benefit or lose or be uncomfortable due to privacy invasion.
2) Data is sold – some data that is collected is sold to those who can afford it to help offset the cost of collection.
3) Data is as-is – data is in an imperfect state, and the process of cleaning it is too difficult or time-consuming or expensive to complete.
4) Data acts as a source of power – those who control data maintain importance in relationships, even if no money changes hands.

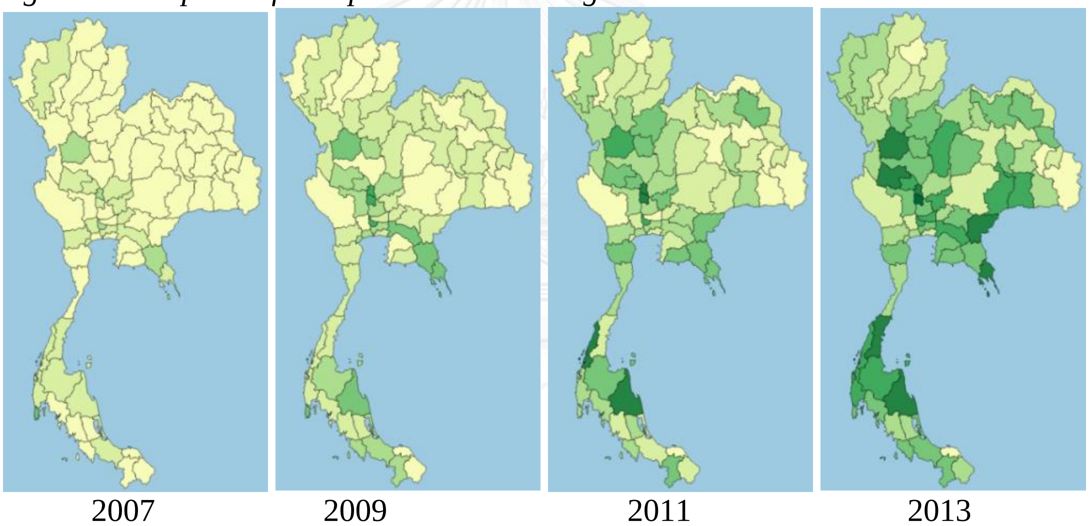
Nevertheless, the data belongs to the public, and needs to be accessible to the public in one form or another. One general solution has been to package the data into summary statistics. This could be five to ten tables on a website, or a book of summary statistics, or excel / csv files available on a website. However, all of these solutions are imperfect as a window into big data.
The author of the website must necessarily choose those summary statistics which will be most of interest to a general audience. It is impossible to provide everything a specialized researcher would need.
The summary statistics book was by necessity a hard copy of tables of statistics which required repeated data entry by the researcher. Most disaggregated series in the book were rarely used.
Excel or csv tables are in a much more convenient form since data for all series can be plotted and graphed once downloaded. However the data is already aggregated into summary statistics so that questions other than those anticipated by the compiler are difficult to research. For instance trade statistics are aggregated into codes which make the study of particular products opaque.

Fortunately, as data accumulates, public organizations are beginning to question the risk of releasing disaggregated data to the public. The Thai National Statistics Office now considers survey data collected at the provincial level to be sufficiently anonymous to allow the easing of requirements for release of disaggregated data to those with proven academic interests. Nevertheless data access is still carefully monitored, and many ministries continue to hesitate to release any disaggregate data. One alternative to releasing the raw data could be data visualization. Data visualization could work on the originally disaggregate data while leaving it in the possession of the owner. Data visualization could allow unreleased data to be available as versatile summary statistics, greatly increasing the uses they could be put to.

5.4.2.1 Example

The following maps were created by combining a JavaScript program written the author with, a json map of Thailand and a simple excel file containing 77 observations. The 77 observations were averages compacted from about 160,000 original observations. The data is about adoption of compact light bulbs over time.

Figure 69 Adoption of Compact Florescent Light Bulbs over Time



Source: Thai SES

This mapping program requires an input of only an excel or csv file with a single observation by province or state. It would be easy to create an app that could create a graph like this for every variable collected at the province level by various Ministries. We wrote the above JavaScript program to look at the 600 separate variables asked in the Thai Socio-Economic Survey.

5.4.2.2 Private Company Data

Who owns private company data? Both sides of the transaction. Any time we make a transaction with a company we create data. That data should be available to both parties. Owners of a phone service should have access to the aggregate data that is collected, sort of like a club membership. There is no justification for the phone company to retain as much data on us as it does, unless it helps us as well. Furthermore, just because data <u>*can*</u> be collected on a transaction, doesn't mean that it should be. There are presently three large sources of data available:

- Internet
- Phones
- Financial

The case for collecting personal information seems strongest for financial data. Financial transactions need to be verified, and records need to be kept.

The telephone service provider needs to know your location to provide cell service, but it doesn't need to keep this data. In this case, opportunity leads to practice, and ***possession is 95% of the law*** when it comes to data. Furthermore, the need for the phone company to keep records of who you call has been much less important in the era of unlimited calls and data by the minute. Maintaining record of who you call for your convenience is often used as a justification but could easily be kept on your local handset, as few people would use more than one phone. Maintaining information on billions of people to catch a few terrorists seems to be an extreme measure more attuned to a controlling state.

Fair Usage Laws – Basis for EU Privacy Law

- For all data collected there should be a stated purpose.
- Information collected by an individual cannot be disclosed to other organizations or individuals unless specifically authorized by law or by consent of the individual — NOT TRUE / Okay – Terms of use
- Records kept on an individual should be accurate and up to date — Impractical
- There should be mechanisms for individuals to review data about them, to ensure accuracy. This may include periodic reporting — Impractical
- Data should be deleted when it is no longer needed for the stated purpose — NOT TRUE
- Transmission of personal information to locations where "equivalent" personal data protection cannot be assured is prohibited — Maybe Untrue
- Some data is too sensitive to be collected, unless there are extreme circumstances (e.g., sexual orientation, religion) — Impractical, but Best Effort

SOURCE: EUROPEAN FAIR PRACTICE / PRIVACY LAWS SEE FOR INSTANCE, KUNER (2007)

Likewise, keeping internet searches could be done locally. If it is not, you should be entitled as one party to the transaction to the same access to data including summary statistics produced with your data. Just because data can be retained, doesn't mean that it should be. Again, possession seems to be 95% of the law.

5.4.2.3 A level playing field

One of the artifacts of big data is that the playing field is becoming increasingly un-level as access to data is not universally available. Facebook and Google and data aggregators sell data for very high prices to large international companies, who use the data to compete with smaller firms. Large firms have access to information about consumer behavior and ways in which consumers can be persuaded. Such data should also be available to the consumers.

Libraries have long been one of the forces that help to make the playing field level by allowing access to resources not elsewhere available. Big data needs to be available somehow through the public sector. How this would happen remains unclear as there is so much data that collecting everything that anyone might want would be impossible.

One area in which it would be possible to help level the playing field would be with the previously mentioned ministry libraries. At the Thai National Statistics Office disaggregate data is available for those willing to pay for it. Aggregated statistics are however easily available to all, and allowances are made for academic use. For many other ministries, the lack of an adequate interface makes data access difficult, unintentionally giving advantage to those with teams of employees.

5.4.2.4 Example

While working on a project for the Land Development Department I discovered they were collecting data on soil conditions at a scale of 1:25000 (100 meters is 4 cm). This was too detailed to easily put online. I asked who used this data. Apparently only Chareon Pokphand (A giant Thai agricultural conglomerate) used the data. They regularly sent teams to the department to study the maps and buy up the best farm land. If this data could be available online it could reach a broader audience.

5.4.3 Challenges of Working with Ministry Data

It may be that public is available, but it is difficult to release it except as summary statistics. Some of the problems that often arise are

1) Privacy concerns
2) Data is sold
3) Data is as-is
4) Data acts as a source of power
5) Capacity of Staff

5.4.3.1 Privacy concerns

All survey data is sensitive to some extent, but some is more sensitive than others. If data visualizations made use of summary statistics, data of individuals would become lost in the crowd. Traditionally if sales data referred to less than 3 companies, it would not be reported, as company A would then know the sales of company B. Frequency for most ministry data would be much greater than this, but caution would be needed for disaggregation to small increments.

Average Income for Districts of Chonburi Province

District	Income	Sampled	If N=20	If N=50
1	23,049	129	23,049	23,049
2	16,499	34	16,499	N/A
3	16,644	9	N/A	N/A
4	23,033	134	23,033	23,033
5	16,020	57	16,020	16,020
6	22,282	49	22,282	N/A
7	18,686	42	18,686	N/A

Perhaps a value of no fewer than 20 included in an average would be sufficient to insure anonymity, or if it was wished to be conservative, 50 or even 100 would be possible. This would still be much more disaggregation than is currently allowed.

5.4.3.2 Data is sold

Funding for Free services will always be in short supply. Collecting survey data is both difficult and expensive. One way in which survey collection is subsidized is through sale of data to those who can afford to pay. Usually this represents a tiny proportion of the income for a government ministry. Losing this revenue should be compared to the great increase in the ways that data can be used nowadays. Ministries are understandably reluctant to part with what costs them so much to obtain. The author once experienced this directly while undertaking a study of illegal drugs in Thailand. There were three organizations responsible for collecting and analyzing drug data. Each organization would not give me the raw data which they had spent so many resources collecting, but each was willing to share the raw data of the other two organizations which they had received for free!

It would be a great waste to keep creative energies from using data in novel ways as that is why it was collected in the first place. That being said, data visualization can take the pressure off of releasing data.

5.4.3.3 Data is as-is

Often data is not shared because some step along the way was not completed – there is no documentation, the data was not adequately vetted, the excel sheets are not systematically labeled. The emphasis should be on keeping the requirements for graph able as minimal as possible. If data can be put into an excel file by province than it can be graph able using averages or charts. Therefore, special features of data visualizations such as floating tips should be avoided or optional.

5.4.3.4 Data acts as a source of power

Different people within the organization may be using the data as a source of power. This power could result in giving them more political power, bargaining power, keep them included in decision making processes and could in some cases even result in financial gain.

Land Development Department Example

The Land Development Department has about 500 government employees, responsible for all areas of land degradation and land quality. This is a lot of people, and it is typical of many government offices. Since 500 people are responsible for a narrow branch of the ministry of agriculture, each individual is responsible for one very narrow specialty. If you want soil maps you talk to this person, for water maps talk to that person. Projects are done in teams of 15-20 with each person filling their required role. Lots of work gets done and done well, by working in teams. However, for the soil map person to talk with expertise about water maps would be inappropriate and likely lead to conflict.

Making data widely available takes power away from the person who holds it. Therefore there is likely to be resistance from each expert to share their data. This problem can be mitigated if the authority comes from the highest level of the department, and if the program is department wide, rather than focused on one single group.

5.4.3.5 Capacity of Staff

The final issue is capacity. It is likely that government officers are neither technically capable, nor interested in I.T. issues. They were hired because of their expertise in other areas such as law, science, criminology or sociology.

Training could be three pronged:

- Training the common denominator
- Training the experts
- Training the user

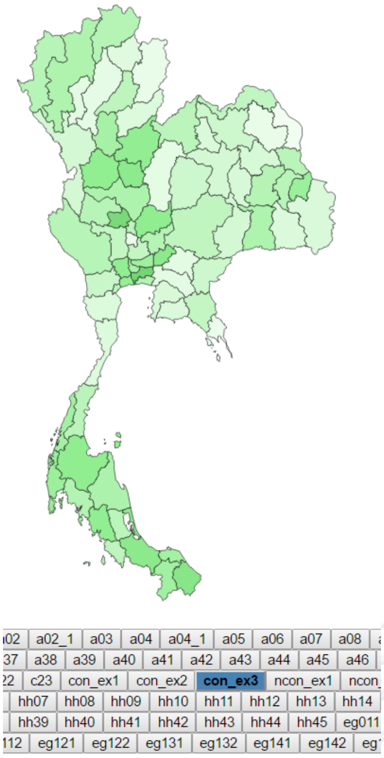
Training the common denominator could be focused on questions such as: How does final data need to look like in order for it to be useful afterwards? Is it possible to adjust the way data is kept to make it more available to outsiders? What kinds of data could each group contribute that would increase value?

Training the experts would be focused on upgrading the skills of computer administrators to handle tasks associated with data visualization. Issues would include training in platforms, security issues, levels of disaggregation that would be considered appropriate and handling user requests.

Training the user - Users of government data are also likely to be unsophisticated, or at least completely new to the ministry's system. What data is made publically available should be simple to use and explained in detail. Ideally there would be a common software used across the ministries.

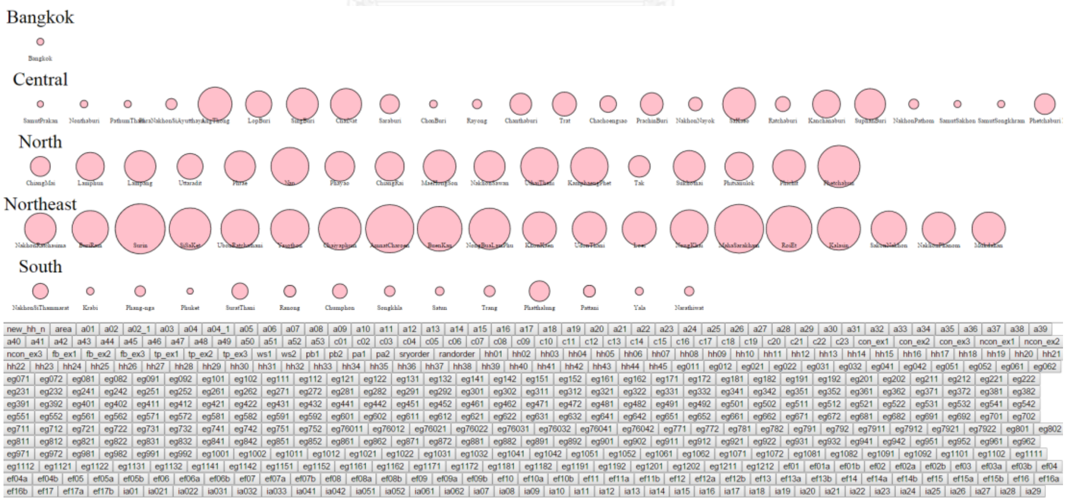
5.4.4 Introducing an Online Visualization Tool

One way to make data from small ministries and departments generally available would be to tie them to a simple visualization tool. The following screen shots are from a tool that the author designed in D3.js for just this purpose. The map or visualization is written in D3, but to use the visualization the buttons or map need to be linked to any csv or excel file. This allows a visual presentation that is easy to understand and use. The buttons on the bottom are automatically generated from the column headings, and clicking on the buttons allows the user to see a graphical presentation of the data for that variable. Data can either be linked to maps or to general purpose buttons as shown below.



SOURCE: DATA SCIENCE FOR DEVELOPING COUNTRIES (2017)

National Statistics Office Expenditure Data for 600 variables



SOURCE: DATA SCIENCE FOR DEVELOPING COUNTRIES (2017)

If anyone is interested in help with this sort of project, code is freely available. For more information consult the website for ***Data Science for Developing Countries*** hhttp://www.bigdata.econ.chula.ac.th or consult the author.

5.4.4.1 Data Availability Summary

Data visualization provides one way to provide access to public data even when the original data is not released to the public. Because the original information has become anonymous through aggregation, data visualizations can present information that would otherwise need to remain confidential. Data visualization can present data in many flexible ways that would not have been anticipated in summary tables. It allows access to data that would not be easy to publish. However, whenever possible , data should be released to the public. Privacy and ownership rules for data should be modified and rationalized.

5.4.4.2 Information Monopoly Act

Given that many technology products are now necessary to lead a normal human life, and given that these technology products are provided by private companies, which are by their nature necessarily natural monopolies, there is a need for the government to regulate the collection and distribution of data so that it is non-abusive, fair for the company and the public on which data is collected, and available in anonymous or aggregated form for a fee to third parties and the government for purposes of the public good.

In line with the above, the following principles should be guaranteed

1) Some return for those who collect data in line with cost of collection
2) Some ability for others to use data – prices subject to approval by the government
3) Some ability for governments to make use of data for public good projects
4) Ability for companies to preserve original data, that can be licensed.
5) Ability to reserve data that is specific to the company for competitive reasons.

- **Indian Statistical Institute, Bangalore 2016**

5.4.5 Data Visualization and Energy Subsidies

5.4.5.1 Liquefied Natural Gas (LPG)

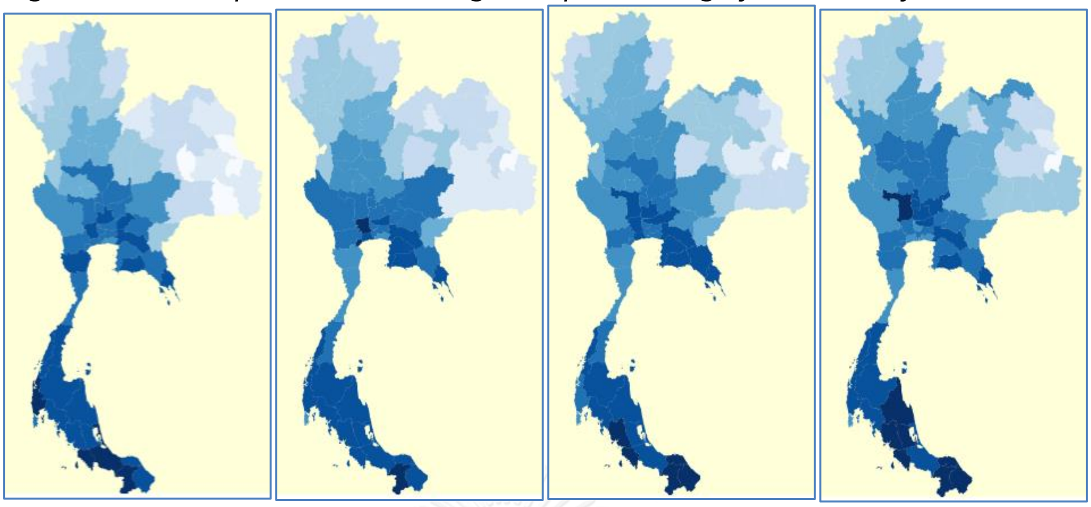
Although the subsidies for LPG have recently been abolished, Thailand has had a long history of subsidizing LPG, and it is worth considering if they are needed.

It was Thai policy since the 1990s to keep the price of LPG cheap. Thailand was an exporter of LPG gas, and the idea was that LPG gas was a national resource and belonged to the Thai people and should be used by all Thais, not sold to foreigners at a profit. Unfortunately, because of low prices, it was not long before LPG in Thailand was less than demand, and imports of LPG became a sizeable expense.

Gas is used both for production and for consumption, and often affects the poor. Street merchants use it for cooking and domestic households use it for cooking. Gas also has industrial uses – and from the beginning it was attempted to keep gas for industry at a higher price. As energy prices rose, gas was also used for transportation. Later governments tried to raise the price for gas for transport be higher than for residential use, with some protection for the somewhat poorer taxi drivers. A succession of governments attempted to close down the subsidy completely, but it was only after the price of LPG declined on its own after the time frame of this paper, that they were able to do so without the wraith of the voters.

Any policy subsidizing LPG will have strong regional biases as natural gas comes from the south, entering either from the Gulf of Thailand or from Burma. LPG is not cheaper in the South than other regions but there may be issues of availability as the gas does not need to be transferred far.

Figure 70 Share of Households Using LPG for Cooking by Province by Year

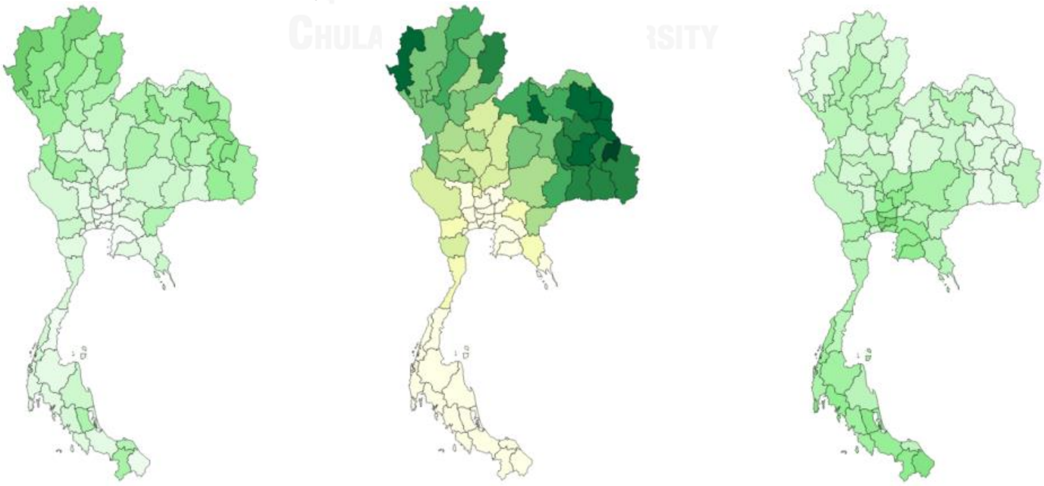


2007 2009 2011 2013

NOTE: THE DARKEST COLORS ON THESE MAPS IN THE DEEP SOUTH REPRESENT MORE THAN 90% OF HOUSEHOLDS ARE USING GAS FOR COOKING WHILE IN THE NORTH AND NORTHEAST REGIONS THE SHARE MIGHT BE AS LOW AS 20% OF HOUSEHOLDS. SOURCE: SES VARIOUS YEARS

Over time, the use of LPG has been spreading from the South northwards towards the North and the Northeast regions as this series of maps shows us.
In both 2011 and 2013, the highest use of LPG was in Naratiwat province at 98% and 97% of sampled households using gas for cooking, while the lowest usage was in Amnart Chareon province for both years with only 15% and 17% of houses using gas for cooking.
The implications for policy are that Parties that wish to win votes in the South and the Central region could promote policies that reduce the price of LPG.

Figure 71 Regional Analysis (continued)



Free Electricity Use Firewood or Charcoal Own Car

SES: 2013

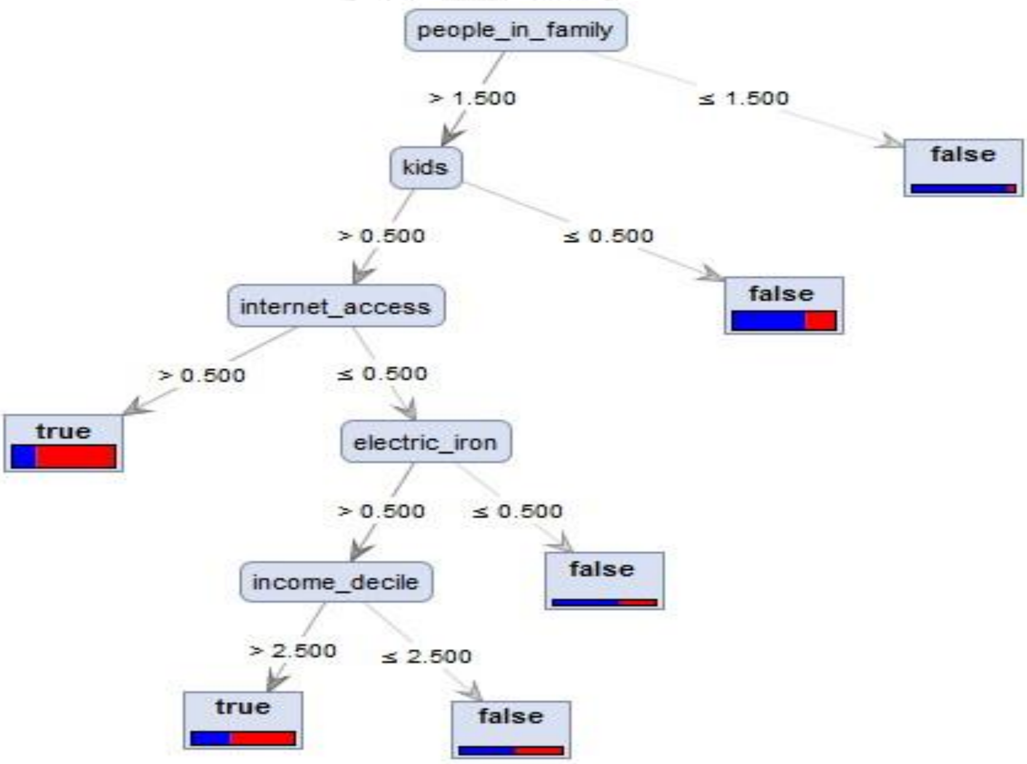
Several other maps can help us with regional interpretations. In part 6 regressions we found that free electricity most benefits those with smaller households. In addition, there is a strong regional component to free electricity, with those living in cooler mountainous regions receiving free electricity more often. This may not reflect a difference in lifestyle so much as energy needs, since it requires more electricity to keep a refrigerator cool in a hot climate than in a cool climate. Hotter households may also use fans more. The policy implication is that perhaps free electricity cutoffs should vary by province.

Firewood also is available in mountainous forested areas. By far, the northern and northeastern regions rely the most on firewood and charcoal. Almost all of the firewood was collected for free, and 2/3 of the charcoal was free, presumably with the household producing it themselves.

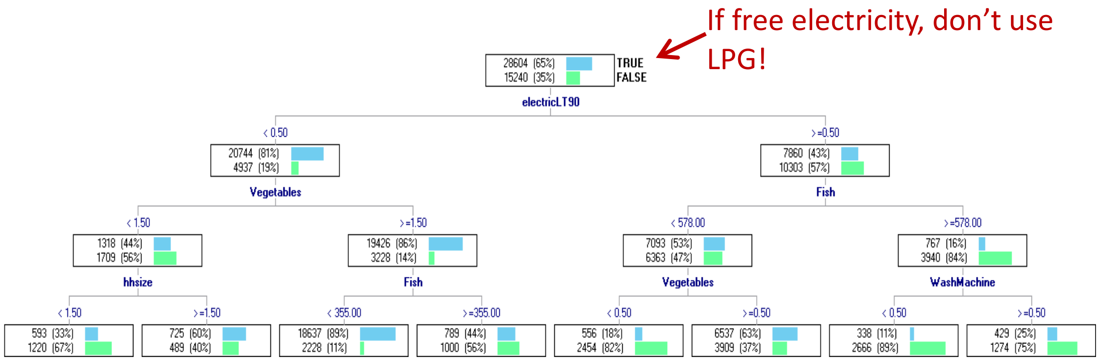
Although less visible in the final map, automobiles are much more common in Bangkok and the surrounding area, while motorcycles and pickup trucks are more common farther away from Bangkok. Policy implications are that programs that benefit car owners, such as support for benzene 95, or the first car policy benefit the Bangkok area and perhaps the south, while programs that promote diesel and benzene 91 benefit those outside of the Bangkok area.

5.5 Other Techniques to Identify Groups

5.5.1 Classification Trees



This is a decision tree generated by the author showing characteristics of those who utilize a school bus. Data is from the Thai Socio-Economic Survey. The objective is to predict those who could benefit from this as a public service. Decision trees go beyond description into predictive science. Access to the raw data would not be a requirement if the software were sufficiently advanced.



5.6 Conclusions

There are four parts to this paper, all of which have conclusions that can help with the targeting and distribution of subsidies. Section 5.2 is a theory section that discusses four different approaches to subsidies, all of which can useful in one situation of another. The four approaches are:

Table 33 Four ways to give subsidies

Nickname	Policy	Example
Help the Poor	Give the subsidy to just one targeted group, usually the poor	Free school lunch
Social Society	Give everyone the subsidy in equal amounts	Public libraries
Social Safety Net (Libertarian)	Let those willing to put up with the conditions get the subsidy	Soup kitchens
Cash Transfers (Authoritarian)	Give cash instead of subsidies to targeted group	Welfare payments

"Help the Poor" is the approach most broadly discussed in this dissertation. It is most useful when i) there are substantial income inequalities such as in Thailand so that some people are in special need, ii) when it is possible to identify and target a particular group, and iii) when subsidizing the entire population would be too expensive. Much of the rest of this dissertation is helping practitioners to make use of this technique more effectively.

"Social Society" is most appropriate when income distribution is not too unequal, when the subsidy is for a good that everyone could only use a little bit of, and when there are arguments about fairness such as subsidies should go also to those who pay the most taxes. It is important that this technique is used ONLY if the quantity is clearly fixed to be the same for each person, or it will turn into a subsidy for the upper classes.

“Social Safety Net” – Thailand has had significant success with self-selecting subsidies, which are subsidies that can only be obtained by waiting, accepting crowded conditions, returning to the shop often and are in other ways inconvenient. This type of subsidy is best when the differences in preferences between the poor and the wealthy for time, convenience, comfort and spending money naturally create this cheap and libertarian approach.

“Cash Transfers” are best for a more authoritarian society that keeps close track of its citizens. It requires more knowledge of individuals within its population than many middle income countries can provide. Best for small and easily identified target groups.

Section 5.3 discusses a Loop Model which is a compromise between a traditional Economic style deductive model approach and modern data mining techniques. The data mining part allows the inclusion of any informative variable, while the iterative and reflective part only allows variables that have some clearly identifiable justification. The objective is to predict a certain group of people (in our case potential LPG subsidy users) so that they can be more carefully targeted. Knowing about the customers helps us find ways to deliver subsidies as well as to know where and when subsidies are useful.

Section 5.4 discusses Data Visualization as a tool for identifying subsidy recipients. Econometrics mostly focuses on average relationships and data visualization is especially effective at pointing out outliers so acts as a very useful compliment to summary statistical techniques. Humans are visual and it is often the work of seconds to identify patterns over time, over space, or over some other variable that would be extremely hard to identify from raw data. Data visualization therefore is an extension of traditional graphing that works well as a pedagogical tool. Although big data allows for very reliable graphical presentations, more work needs to be done to separate the parts of a map, for instance, that have strong statistical support from the parts that do not. Therefore if maps were to be used as a logical “proof” as economists are wont to do, there is still work needed for how to show this. Data ownership is an enormous issue – the data should belong to participants on both sides of a transaction, and for now data is monopolized in a way that is not socially optimal nor just. A set of principles are proposed in the paper for a just use of data.

The following principles should be guaranteed

Some return for those who collect data in line with cost of collection

Some ability for others to use data – prices subject to approval by the government

Some ability for governments to make use of data for public good projects

Ability for companies to preserve original data, that can be licensed.

Ability to reserve data that is specific to the company for competitive reasons.

Lewis (2016), Indian Statistical Institute

Chapter 6 Conclusions and Policy Recommendations

6.1 Introduction

With world energy prices likely to rise in the future energy subsidies are likely to become important for middle income countries again in the near future. Although many people are against commodity subsidies for efficiency reasons, they are likely to reemerge as they solve two common problems of middle countries. The first is that middle income countries can not easily track all the details of their citizens lives, especially their income. Therefore, the alternative, of providing a cash transfer to each poor person is difficult to accomplish.

Second, subsidies are as much a political response to higher prices as they are an economic response. Government's must appear to be solving their citizens problems (in democratic societies) or they will not be elected. Therefore a once and for all solution seldom has the same appeal to politicians as a series of subsidies. Economists must learn to work together better with politicians designing more short run solutions if they do not want the populist mistakes of the past repeated over and over.

Therefore, if subsidies are part of the daily life of a middle income country (and this paper argues, as a welcome addition) they need to be designed better, with clear targets and objectives. Issues addressed in this paper include the type of subsidy that would best be given, the amount of the subsidy to be given, instrumental alternatives to direct subsidies (subsidize something else such as small tanks that only the poor use) , econometric tools for tracking the use of subsidies, econometric tools for distributing subsidies and finally econometric tools to help identify and target specific groups. Almost all of these tools are developed for use with the widely available household economic survey available in many countries around the world.

Some alternatives – One alternative to tracking poor people better is to collect better data on everyone. That is one of the objectives behind recent efforts to promote demonetarization, most dramatically seen by recent events in India. Not only was most of the currency withdrawn from circulation overnight, but strict limits were imposed on cash that could be removed from banks, mobile phone paytm was set up to help poorer persons make payments electronically, and there will soon be high transaction fees at banks for people making transactions in cash. This paired with the widespread adoption of the Aadhaar card could potentially bring most of the transactions of a country of 1 billion people to be under the observation of a central government. (Admitted it is only a dream so far – India at 5% electronic transactions is one of the least electronically equipped countries on the planet).

Other countries have made moves towards demonetarization by removing larger banknotes from circulation (EU 500 and 200 Euro and Venezuela) are two cases with Switzerland (500 Swiss Francs) likely to follow soon. Two questions should arise from this often reviled idea. 1) Is it possible and 2) Is it desirable? The answer to the first should be answered by observing what happens in India where a lot will need to change. The answer to the second is also uncertain. There are many that feel strongly that the authoritarian loss of freedom of a centralized control of all finances are not

worth the benefits that it would confer in terms of managing the citizens, and incidentally, managing cash transfers and subsidies.
For the moment Thailand remains on the libertarian end of the spectrum, and the central government seldom tries to force its ideas on the population, rather preferring to persuade them or entice them.
As mentioned earlier in the paper, subsidies can be incredibly expensive when poorly designed, and popular ones can be very difficult to remove by an elected government, therefore they need to be designed with care. Hopefully some of the econometric tools suggested in this thesis could be of use in this effort. Especially care must be taken to design subsidies that 1) reach only the target group, or 2) are available in Equal Quantity to each person or 3) are designed to be self-selecting, so only those who truly need them will use them. These principles, plus the cash transfer principle are described in detail in the chapter on Targeting and Distribution above.

6.2 Conclusions and Policy Recommendations

6.2.1 Big Picture

The challenges remaining for Energy Policy in Thailand are many. Thailand still need to find a way to reduce its use of energy, as it has one of the worst energy to GDP ratios in Southeast Asia. As an energy importer this is expensive and risky. Many well conceived conservation policies were initiated under the thoughtful leadership of former Energy Ministry Piyasvasti Amranand but mostly they were not developed.

Thailand still needs to meet commitments to reducing global warming through NAMAs (Nationally Appropriate Mitigation Action.) Although there was some interest generated when it appeared that Carbon Credits could be profitable, little has been done to reduce Carbon emissions, and to help the world we live in (Limmeechockchai, 2016).

Electricity and Fuel supply is still mostly in the hands of large State-Owned Enterprises, with insufficient competition to increase efficiency and reduce price. Electricity is still mostly produced with natural gas which puts Thailand at risk of supply interruptions.

Alternative energy is far from sufficient to meet future energy needs, so Thailand needs to continue to be vigilant and be concerned with promoting alternative fuels and with energy security and energy integration.

6.2.2 Subsidies

This paper has selected just one issue on the complicated energy agenda by looking at subsidies, but it is an especially important direction for several reasons. Energy subsidies have a strong effect on the lives of ordinary Thai persons, and especially the ones with the hardest lot, and the least reason for hope. Well designed subsidies can make a real difference.

Energy subsidies, poorly done, can be very expensive. Thailand is a country with strong Macro policy and weak Micro policy, and there is an opportunity here to do better by using econometric and big data techniques to improve microeconomic policy.

The paper has looked at current and recent energy subsidy programs and investigated their success at reaching the poor, at reaching the elderly, the educated or uneducated, those who live in different regions. The advent of big data allows us to know much more about those who receive subsidies. In addition to looking at how many of each group receive benefits and how much they receive , some estimate of the share of the budget going to each policy is estimated.

The large data sets used in the paper comes from the National Statistics Office which has taken the lead in Thailand in making government data available both to the public and to researchers. The author is very grateful for their support in making the data obtained available to the broader academic community for academic research.

6.2.3 Summary of Results

In this analysis, free electricity is clearly the winner in terms of reaching its targeted recipients and being used primarily by the poor. In no other subsidy program does more than half of the benefit go to the targeted recipients. Nevertheless there are still unintended consequences to the way free electricity was designed that make it favor those in mountainous areas, those with small families and those with absentee family members. In addition, the cutoff level for free electricity is clearly too low to be optimal.

Subsidizing LPG is largely synonymous with subsidizing the middle class. The characteristics of households who use LPG are closely aligned with a middle class lifestyle. There is only a small overlap between those who receive free electricity and those who use LPG for cooking, which makes any link between the two policies problematic.

Subsidies of vehicles and transportation fuel have the least link to issues of poverty. If subsidizing any transport fuel could help the poor, it would be keeping the price of gasohol 91 cheap, as it is the primary fuel used for motorcycles. Even so, the vast majority of expense will go to wealthier households who drive much more. Perhaps it would be better to subsidize motorcycles instead? Benzene 95 is a luxury fuel, so that its usage rises rapidly as households achieve a high level of income. Diesel is not used at all by the poor, but may benefit the middle class who use it for pickups used in production, or more frequently by the wealthy. However, the 30 Baht per liter price cap set by the government on diesel seems to have been directed only at the commercial transport sector, with little awareness by, or effect on anyone else.

The first car policy seems to have been directed at supporting the domestic car industry, especially in the aftermath of the flood. Only the top two deciles appear to have been affected by the first car policy for automobiles. However, there does seem to have been an increase in the purchase of pickup trucks for middle decile groups. It might be worthwhile for the government to target this sector directly in the future to decrease production costs.

The ethanol support program, in the form of promoting gasohol is mostly an attempt to prepare the country for a future using alternative fuels. It is not cost effective at current energy prices, and it does not do a very effective job of reaching the poor, so the justification for the policy must be fuel availability in the future.

All alternative energy policies including ethanol and especially solar that are not linked to least cost economic rationality need to be closely monitored for economic reasonability as they have become a strong growth industry for transferring funds from taxpayers and utility users to the wealthy and well-connected in recent years. Although continuous efforts have been made to reduce perverse incentives, more needs to be done in this area.

In summary, if the primary purpose of the government is to subsidize poor people, it is not doing a very good job. Electricity subsidies are very effective at reaching the poor, but have been cut back. LPG subsidies mostly reached those in higher income groups and have very little effect on the household budget. The more useful goal of improving people's health by removing particles for the air is not being met.

6.2.4 Potential Solutions to Aid the Poor.

In this section, several possible solutions that can help in the design of subsidies are suggested. Engel curves provide a way to estimate the amount of energy that should be provided as lifeline support. Subsidizing capital goods that are used by the poor is a more direct way to reach a target group. Finally, the issue of aid for the elderly and income equality are a more direct way to decrease economic disparity.

1) Use Engel curves to bring the poorest households to a target level set by the 10-50% centile energy use. Suggested levels from section 7 are:
    a. Electricity 80 - 140 KwH / month
    b. LPG 3-5 Kilograms / month
    c. Gasohol 91 10-20 liters / month.

The manner in which to insure these levels remains to be determined, but it is useful to have a target level. Presumably, the easiest would be the same approach as is used for electricity.

2) Subsidize capital goods used by the poor rather than subsidizing fuel.
    a. Stoves – government could distribute gas stoves or encourage their use.
    b. Motorcycles – why a first car policy, when a significant expense for most of the poor are motorcycles?
    c. Pickup trucks
    d. Farm equipment

3) Other subsidy issues
    a. The subsidy for old people is clearly not sufficient at 600-840 baht a month. For a typical elderly couple who each receive 720 baht a month, this is less than 25% of their monthly costs for the poorest decile.
    b. Energy use should be increased for the poor and decreased for the wealthy. Probably this means higher taxes on cars and benzene95 and continued subsidies. It is not true that *everyone* should reduce energy use.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., & Tobias, J. (2012). Targeting the poor: evidence from a field experiment in Indonesia. *The American economic review*, *102*(4), 1206-1240.
Amranand, A. (2014). *(Energy) Between the Lines*. Bangkok, Thailand (in Thai).
Anand, R., Coady, D. P., Mohommad, A., Thakoor, V. J., & Walsh, J. P. (2013). The Fiscal and Welfare Impacts of Reforming Fuel Subsidies in India.
Arrow, K. J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, 265-290.
Arze del Granado, F. J., Coady, D., & Gillingham, R. (2012). The Unequal Benefits of Fuel Subsidies: A Review of Evidence for Developing Countries. *World Development, 40*, 2234-2248. doi:10.1016/j.worlddev.2012.05.005
Bacon, R., Bhattacharya, S., & Kojima, M. (2010). Expenditure of low-income households on energy. *Extractive Industries for Development Series, 16*.
Bangkok Post. (2017, Jan 11, 2017). LPG subsidies raised for price stability. *Bangkok Post*. Retrieved from http://www.bangkokpost.com/print/1177953/
Bank of Thailand. (2008). *2008 Inflation Report*. Retrieved from
Barnes, D. F., Khandker, S. R., & Samad, H. A. (2010). Energy Access , Efficiency , and Poverty How Many Households Are Energy Poor in Bangladesh ?
Chuangwilai, C. (2014). *Controlling Diesel Price in Thailand*. Senior Research Project, Faculty of Economics, Chulalongkorn University.
Scoring Pattern, Socio Economic Survey 2002-2003, (2002).
Dartanto, T. (2013). Reducing fuel subsidies and the implication on fiscal balance and poverty in Indonesia: A simulation analysis. *Energy Policy, 58*, 117-134. doi:10.1016/j.enpol.2013.02.040
Data Science for Developing Countries. (2017). Data Science foir Developing Countries Website. Retrieved from http://www.bigdata.econ.chula.ac.th/
Deaton, A. (1997). *The Analysis of Household Surveys : A Microeconometric Approach to Development Policy*: World Bank, John Hopkins Press.
del Granado, F. J. A., Coady, D., & Gillingham, R. (2012). The unequal benefits of fuel subsidies: A review of evidence for developing countries. *World Development, 40*(11), 2234-2248.
DFID. (2002). *Energy for the Poor, Underpinning the Millennium Development Goals* UK Department of International Development.
EGAT. (2017). *Electric Generation Authority of Thailand. Price of Electricity*.
EPPO. (2017). *Energy Policy and Planning Office, Energy Statistics, various tables,* .
Foster, V., Tre, J.-p., & Wodon, Q. (2000). Energy prices, energy efficiency, and fuel poverty.
Halff, A., Sovacool, B. K., & Rozhon, J. (2014). *Energy poverty: global challenges and local solutions*: OUP Oxford.
Hills, J. (2012). Getting the measure of fuel poverty Final Report of the Fuel Poverty Review of fuel poverty Fuel Poverty Review.
IEA OPEC World Bank. (2010). Analysis of the Scope of Energy Subsidies and Suggestions for the G-20 Initiative. 26-27.
IISD. (2013). *A Citizens' Guide to Energy Subsidies: Thailand*: International Institute for Sustainable Development.
IMF. (2013). Energy Subsidy Reform: Lessons and Implications. 68.
International Energy Agency. (2016). *World energy outlook*: OECD/IEA.

Developing an Energy Based Poverty Line for South Africa, IAEE_June_2011 (2011).
Krungsri Bank. (2011). Economic and Industry Watch.
Kuner, C. (2007). European data protection law. *Corporate Compliance and Regulation Oxford University Pres., Büyük Britanya*.
Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, 1-28.
Lewis, D. R. (2011). *Energy Policy for the Poor*. Paper presented at the Shanghai International Conference on Social Science, Shanghai, China.
Lewis, D. R. (2013a). *LPG Subsidies and Electricity*. Faculty of Economics Working Paper.
Lewis, D. R. (2013b). *Subsidies, Dedicated Smart Cards and Direct Benefit Transfer Programs.* Paper presented at the Singapore Economic Review, Singapore.
Lewis, D. R. (2014). Energy and ASEAN: An Experiential Introduction to Energy Policy [Workshop]. [Workshop and Materials]: Chulalongkorn University.
Lewis, D. R. (2016). *Data Visualization and Public Data Availability.* Paper presented at the International Conference on Big Data and Knowledge Discovery,, Bangalore, India.
Lewis, D. R., Lekfuangfu, N., & al., e. (2010). *Forecasting Migration Flows in the GMS*. Retrieved from
Limmeechockchai, B. (2016). Designing a National Policy Framework for NAMAs *Enabling Asia to Stabilize the Climate* (pp. 275): , Springer.
Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic Theory* (Vol. 1): Oxford university press New York.
MEA. (2017). *Metropolitan Electricity Authority, calculator, and historical announcements*.
Meyer, B. D., & Sullivan, J. X. (2003). *Measuring the well-being of the poor using income and consumption*. Retrieved from
Ministry of Energy. (various years) Energy Policy (in Thai). Ministry of Energy.
Ministry of Information and Communication Technology. (2006). *2006 Household Energy Consumption Survey*: NSO.
National Statistics Office. (2006-2013). *Thai Socio-Economic Survey, Whole Kingdom (SES)*. Bangkok: NSO.
Nikomborirak, D. K., P. (2014). Thai Energy Sector Reform: Whom to Believe? (in thai). *TDRI-Insight, June 2014*(June 2014).
Practical Action. (2010). *Poor People's Energy Outlook 2010*. Retrieved from http://practicalaction.org/poor-peoples-energy-outlook-2
PTT. (2017). *Petroleum Authority of Thailand Monthly Transportation Fuel Prices*.
Rao, N. D. (2012). Kerosene subsidies in India: When energy policy fails as social policy. *Energy for Sustainable Development, 16*(1), 35-43.
Salehi-Isfahani, D. (2011). Iran: Subsidy Reform amid Regional Turmoil *Brookings Institute*.
Sarntisart, I. (2011). *Income Distributional Analysis: Measurement and Economic Modeling (in thai)*: Chulalongkorn University Press.
Sukkumnoed, D. (2008). *Energy: Work that has Energy (in Thai)*: Green World Association.

Tennakoon, D. (2009). Energy poverty: Estimating the level of energy poverty in Sri Lanka. *Practical Action, Intermediate Technology Development Group. United Nations*.
Thai BOI. (2017). *Summary of Thai Electricity Pricing*.
Thai Meteorological Department. (2017). Statistics on Thai Temperature and Rainfall. https://www.tmd.go.th/en/climate.php?FileID=7.
World Bank. (2001). *World Development Report 2000/2001-Attacking Poverty*: The World Bank.
World Bank. (2010a). Subsidies in the Energy Sector : An Overview. *Background Paper*.
World Bank. (2010b). *World Development Indicators 2010*: Oxford University Press, USA.
World Bank. (2015). *World Development Indicators 2015*: Oxford University Press, USA.

REFERENCES

(2006-2013)

APPENDIX

VITA

Ph.D. Candidate (Applied and Ag. Economics), University of Wisconsin-Madison พ.ศ. 2541

M.A. (Agricultural Economics), University of Wisconsin-Madison พ.ศ. 2536

M.A. (Economics), University of California-Riverside พ.ศ. 2533

B.A. (Philosophy, Biology), Oberlin College, Ohio USA พ.ศ. 2527

EBA Lecturer 2002-Present, Faculty of Economics, Chulalongkorn University

Publications

2003 Book - The Long Trip Down the Mountain: Social and Economic Effects of Illegal Drugs in Thailand – UN Office on Drugs and Crime (UNODC) 181 pp

2017 Big Data Insights into the Distributional Effects of Thai Energy Policy 2006-2017, Southeast Journal of Economics, June 2017

2017 Financial Inclusion in Thailand (With Sukanda Lewis), Hokkaido University Economic Journal.

Research Projects

2013 Land Degradation in Thailand 2013 – Team Member - Food and Agriculture Organization of the United Nations (FAO) and M. of Agriculture, D. of Land Development

2009 Forecasting Migration Flows: the Relationship between Economic Development, Demographic Change, and Migration in the Greater Mekong Subregion (GMS) – Project Leader Asian Development Bank (ADB) and International Organization for Migration (IOM)

2004 Quarterly Jewelry Update distributed to 60 Thai Jewelry Companies - Q2:2002 - Q2:2004

2004 International Indicators of Public-Private Partnership Readiness – Online Tool and Workbook, UN ESCAP (With John Evans-Klock)

2002 Strategic and Competitive Position of the Thai Jewelry Industry. Thai Farmers Research Center

2000 Global Study on Illegal Drugs: The Case of Bangkok. United Nations Drug Control Programme (now UNODC). (With Nualnoi Treerat, Noppanun Wannathepsakul)

Website

2017 Data Science for Developing Countries, http://www.bigdata.econ.chula.ac.th/ Visualizations and data about Thai Agriculture, Poverty and Energy Issues

Conferences with or without proceedings

2016 Data Visualization and Public Data Availability, Indian Statistical Association, Bangalore, Published in Proceedings of International Conference on Big Data and Knowledge Discovery

1997 "The Effect of Industrialization on the Rural Thai Population: Tracking Migration Flows due to Industrialization", published in Proceedings of the 7th Tun Razak Conference, in Malaysia

2017 Big Data Time Series for Analysis of Economic Policies Using Social Economic Surveys, Conference Paper, Singapore Economic Review, Singapore Aug 2017)

2017 Quantile Regressions and Elasticities for Estimating the Demand for Residential Energy, Conference Paper, I. Conference Applied Economics and Policy, U of Malaya, Malaysia, Aug 2017

2016 Controlling and Managing Energy Subsidies, East Asian Economic Association, Bandung Indonesia

2014 Controlling and Managing Energy Subsidies, East Asian Economic Association, Thailand

2013 Probit Regressions and Data-Mining with Socio-Economic Data: Predicting who will use LPG Global Information and Management Symposium, Taipei, Taiwan

2013 Subsidies, Dedicated Smart Cards and Direct Benefit Transfer Programs, Singapore Economic Review Conference, Singapore

2011 Energy Policy for the Poor, Shanghai International Conference on Social Science, China

2006 Politics, Economics and the Price of Oil, Second Thai Economists Conference