ESTIMATING U.S. RESIDENTIAL DEMAND FOR FUELWOOD IN THE PRESENCE OF SELECTIVITY

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Ryan Michael Daly

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Agribusiness and Applied Economics

November 2014

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Estimating U.S. Residential Demand for Fuelwood
in the Presence of Selectivity

By

Ryan Michael Daly

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. David Ripplinger
Chair

Dr. William Wilson

Dr. Ryan Larsen

Dr. Tonya Hansen

Approved:

November 12, 2014
Date

Dr. William Nganje
Department Chair

ABSTRACT

Residential energy consumers have options for home heating. With many applications, appliances, and fuel types, fuelwood used for heating faces stiff competition in modern society from other fuels. This study estimates demand for domestic fuelwood. It also examines whether evidence of bias exists from residential homes choosing to use fuelwood. The use of OLS as an estimator will yield biased results if such selectivity exists.

Selectivity is addressed with a Heckman (1979) two-step procedure; bias in fuelwood demand estimation using OLS is reduced. Non-wood energy prices and income are major determinants of fuelwood demand. Geographical regions and urbanization confirm results from prior studies.

ACKNOWLEDGEMENTS

I'd like to thank my major advisor, Dr. David Ripplinger. His motivation through my hard work involving Energy Economics and market research is a major driving force in my determination to succeed in my studies and future career. Dr. Ripplinger's support from the beginning greatly influenced my ability to complete this study. Without the guidance or patience and understanding, this project would never have come to completion. I'd like to thank my thesis committee members, Dr. William Wilson, Dr. Ryan Larsen, and Dr. Tonya Hansen. Without their expertise, comments, contributions, and grand efforts in graduate studies in economics, I wouldn't have succeeded in the completion of my thesis. Lastly, I'd like to thank NDSU and the Agribusiness and Applied Economics Department for their financial support and educational opportunity in this research.

DEDICATION

Before engaging the reader, I sincerely like to thank my dear family. My lifelong friend and loving wife of 7 ½ years, Nancy, and our wonderful children, Aiden (3 years) and Olivia (1 year), without whom the motivation of this study would easily have dwindled into the night. Aiden greatly drove me to continue on to ensure success for myself in providing for my family. Olivia gave me additional heart and compassion to continue on and flourish in my studies. The two of you are what brings fire and life into our family. And Nancy, without your continued support (and the long nights of suffering, wondering if this will ever end), I most definitely wouldn't have returned to school to reeducate myself for the better of ourselves.

To my parents, Lyn and Tim Flieth, and Robb and Maria Daly. Your additional support and understanding of my determinacy to succeed continues to drive myself to excel in life. To my father-in-law and mother-in-law, Ron and Barb Martin, your exemplary support goes above and beyond what is ever required of an extended, yet extremely close, family. To my brothers and sisters, grandparents, aunts and uncles, cousins, nieces and nephews, your persistent kind words and acknowledgement of my hard work doesn't go without recognition. And lastly, to the many old and new friends and colleagues I've met through the last 5 ½ years, especially the last year. The few times we've spent together during this hard era in my life was well needed for fun and relaxation.

Ceteris paribus carpe diem

"Never regret - If it's good, it's wonderful. If it's bad, it's experience." ~ Anonymous

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

As energy prices rise and fossil fuel reserves diminish, U.S. households face challenges in economically and reliably heating their homes. This is in part due to switching costs and limited heating system alternatives (U.S. Energy Information Administration, 2008). With energy being a key determinant of livability, growth, and achieving a higher standard of quality of life within any developed economy, volatile prices and supply disruptions create uncertainty for future economic stability (Stern, 2011). Since residential homeowners have limited heating fuel options, with the majority being fossil fuel based, many simply default to the homes original heating source supplied when purchased without realizing the option exists to switch to another source. Even if the possible utility received from switching is higher than the current source, as consumers tend to maximize energy utility, the investment costs associated with the heating source implementation may not be fully understood (Quigley and Rubinfeld, 1989; Dubin and McFadden, 1984). This poses problems for firms trying to define strategies to market alternative wood-based systems. This study looks to estimate demand for woody biomass, also known as fuelwood[1], in the presence of selectivity. By incorporating selectivity correction, demand for fuelwood can be estimated with greater precision and accuracy without creating biased estimation results.

Residential homeowners encounter a variety of alternatives for home heating, with the initial choice being whether or not to utilize a specific fuel. When the homeowner decides to use a fuel, essentially analyzing or comparing utility obtained from using the fuel while taking into consideration overall operational and capital investment costs associated with implementing the fuel, bias and inconsistent results can occur if a researcher attempts to estimate this selected fuel demand

[1] Fuelwood, or wood used for energy, is used interchangeably with any reference to wood energy within this paper.

from a sampled population. These decision variables need to be factored into the model to identify determinants of demand for endogeneity within omitted variables.

The fuelwood market within the U.S. residential sector accounts for only 3.2% of all energy used (U.S. Energy Information Administration, 2010b). Thus, firms marketing fuelwood can benefit from this study, as it provides deep insight into a relatively small energy market share. Understanding factors of U.S. residential energy consumption is important as the sector consumes about 22% of all energy within the economy (U.S. Energy Information Administration, 2010a). Due to growing interest in renewable energy, fossil-fuel competing firms are entering the market to provide alternatives to fossil fuels. Fuelwood is used primarily for heating, and demand for these products, such as pelletized wood, is increasing in foreign markets for power generation (National Renewable Energy Laboratory, 2013). With the rapid increase in foreign interest of wood energy, U.S. and Canadian wood biomass firms have increased production by exporting to global markets, especially Europe, while retaining an interest in domestic markets. Initial forecasts of domestic interests in the fuelwood market come as the residential sector faces increasing prices for natural gas, propane, and electricity due to an unusually cold winter and unforeseen supply problems (U.S. Energy Information Administration, 2013c, 2014a). A possible solution to these heating problems is to have a secondary source of energy, such as fuelwood, to reduce potential economic impacts during these conditions (Song et al., 2012a, 2012b; U.S. Energy Information Administration, 2014a).

For firms to market to the residential sector, additional understanding of potential demand for fuelwood is important to competitively market their products and services. Biomass accounts for only 6.5% of all primary energy consumed in the U.S.[2], with wood derived biomass being 48.7% of that total (U.S. Energy Information Administration, 2010b). This is relatively small in comparison to fossil fuel sources.

[2] U.S. demand for wood energy as whole, all sectors, all uses.

We present this study on residential wood energy demand as a means to estimate the responsiveness of typical households in the relatively small wood energy market. The study should aid industrial wood-fuel appliance and fuel suppliers to make decisions regarding how to market their products and related services. Correcting for decisions associated with residential selectivity of energy use will correct for potential bias and inconsistent results if a OLS regression is used (Heckman, 1976, 1979). This study estimates whether fuel selection plays a role in residential wood energy usage and, if so, use a model to account for biased decision making. A few prior studies seek similar objectives, however no studies using recently available data on the U.S. residential sector exist. This study also contributes knowledge to academic institutes, government bodies, policy makers and fuel industries on how a relatively small energy market can be estimated in the wake of increasing concerns for energy use and security. Policies aimed at reducing emissions, increasing energy security, and reducing energy dependence should be made with more due diligence from this study. We begin the discussion by briefly exploring a background of residential energy demand.

1.1. Overall Residential Energy Demand

Since fuelwood has a small demand relative to the wide range of energy sources, we need to put into perspective overall residential demand for energy. There's a historical trend within the U.S. that shows the industrial sector uses more BTUs of energy relative to other sectors. When viewing historical changes from 1949 to 2009 in overall energy market share (Figure 1.1), however, we find a shift among the four primary sector categories: residential, commercial, industrial, and transportation. The industrial sector observed a dominant 46% market share post World War II in 1949. In 1949, transportation, residential, and commercial held 24.9%, 17.5%, and 11.5%, respectively.

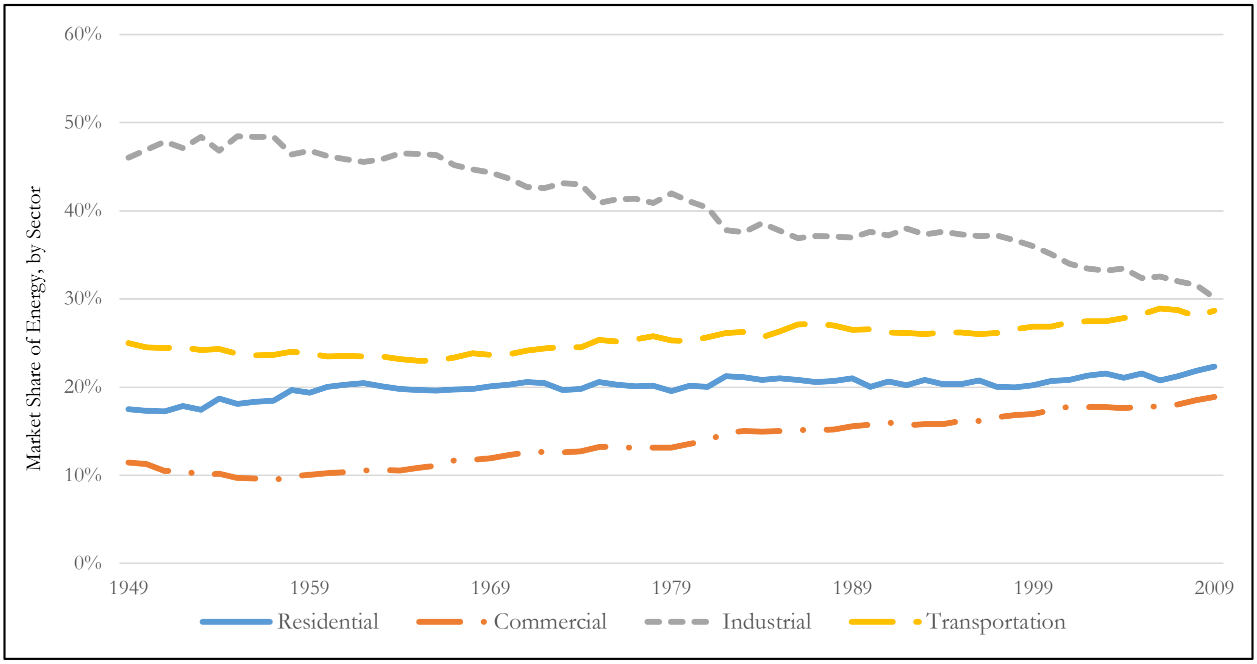

Figure 1.1. Percentage of Overall BTUs Consumed by Sector, 1949 – 2009.
Source: U.S. Energy Information Administration. Monthly Energy Review. (2010c, 2014b).

As the decades passed, the market shares balanced out towards a focal point. As of 2009, the industrial sector holds a 30.1% market share, an overall drop of 34.5% from 1949. The other three increased disproportionately at average growth rate of 15%, 27.5%, and 64.5% for transportation, residential, and commercial, respectively. We find here the residential sector's increase in market share was fairly large in size. If marketing fuelwood as an alternative to traditional fossil fuel energies is a firm's objective, viewing this historical rise in residential energy use needs to be understood as a basis for understanding the target audience. However, observing market share changes doesn't tell the whole story for the residential sector as finer details of which energy sources are used within the residential market provide more insight.

Data available from the U.S. Energy Information Administration indicates all sectors increased in their demand for BTUs of energy during this period (U.S. Energy Information Administration, 2014e). Focusing strictly on the residential sector, quantity of energy demanded has increased 252% since 1949. This sector now represents an overall 22% market share of 10.2

quadrillion BTUs consumed in 2009 within the U.S. Of this total demand, the residential sector's heating needs account for 59.2% of the total energy mix with appliances, electronics, lighting and air conditioning consuming the remaining 40.8% (U.S. Energy Information Administration, 2013b). In the aggregate, this is a strong, growing sector that's susceptible to the same market price and supply shocks, as well as energy and environmental policy concerns the other sectors must worry about as well. This can lead to interests in other potentially more reliable fuel sources if the shocks create enough concern for stability in the short- and long-term. A firm knowing what this aggregate energy mix is can gain additional insight into how to market woody-biomass products by showcasing how to increase heating stability by diversifying the home's energy mix with more than one source of available fuel (U.S. Energy Information Administration, 2014a).

Energy use presented in Figure 1.1 come from a variety of fuels. The U.S. Energy Information Administration lists two types of energy sources for primary consumption: fossil fuels and renewable energy. Looking at cumulative bar charts of the residential energy sourcing mix while focusing primarily on residential housing[3], Figure 1.2, we find the total demand for all energy sources appears to be fairly level between 1978 and 2009, with consumption of all primary energy sources remaining around 10 quadrillion BTUs. However, we observe electricity usage increasing greatly, most likely from an increase in wide spread electrical appliance installation and usage, such as refrigerators, televisions, and air conditioning units (U.S. Energy Information Administration, 2012b, 2012e), while demand for natural gas decreases and fluctuates for fuel oil, liquid petroleum gases, and wood. When viewing per capita energy consumption over time of U.S. residential homes, (Figure 1.3) we find the average yearly consumption trends downward, decreasing 34.78% over this

[3] Footnote 1 of table 2.4 within the 2009 U.S. Energy Information Administration RECS states the focus is on primary housing, which is not the same as entire energy consumption mix the residential sector offers and used for Figure 1.1's market share calculations.

period. Energy used for heating needs is also on a downward trend, as there is a 17.09% decrease in energy consumed for space and water heating from 1993[4] to 2009 (U.S. Energy Information Administration, 2013b).

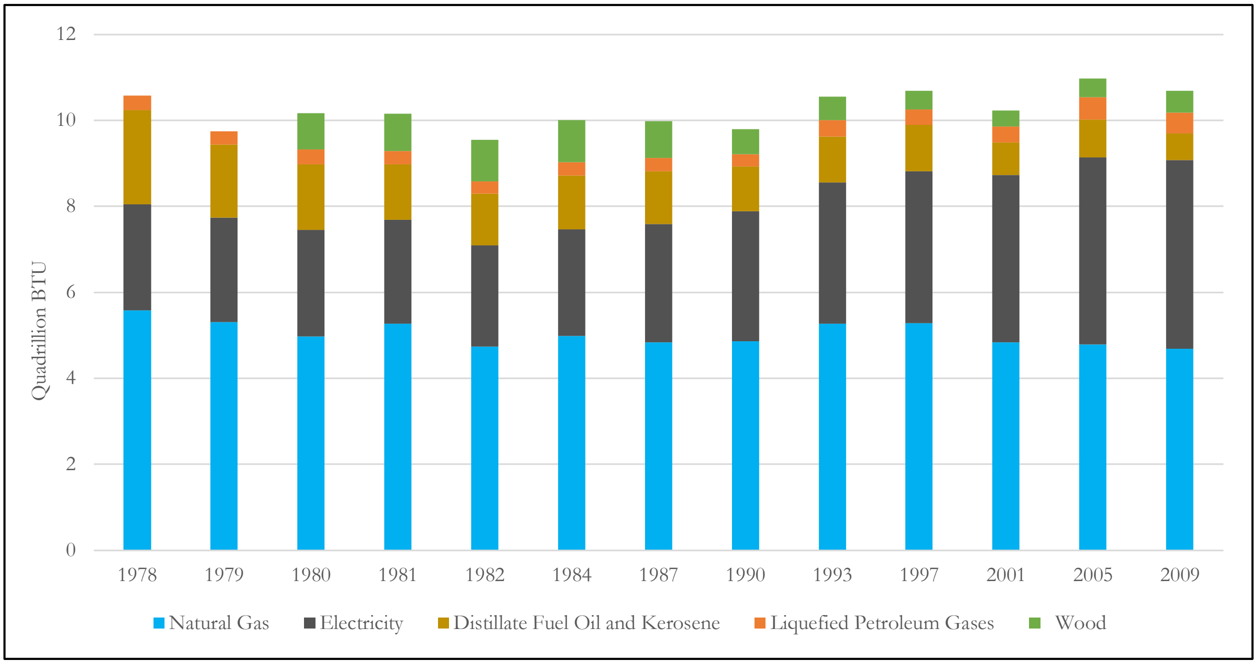

Figure 1.2. U.S. Residential Energy Consumption, Cumulative Total by Source, 1978 – 2009. Source: U.S. Energy Information Administration. Annual Energy Review. (2010d, 2012a).

With this decrease in heating demand among residential homes, a firm marketing fuelwood has reasons for concern as the market becomes tighter, i.e., the aggregate supply of heating energy demanded is shrinking. Thus, we find an increasing need for firms to understand what causes homeowners to switch energy sources, or decide to use one specific source to begin with. Seeing wood energy appeared to observe peak demand within the 1980 to 1990 decade, with another increase in consumption trending near the end of 2009, Figure 1.2, we look to explore residential

[4] Data for RECS years prior to 1993 are not in a user-friendly format, requiring extensive amounts of resources to gather, organize, and report. Please refer to prior RECS studies on consumption habits of residential homeowners.

responses to possible regressors on fuelwood demand as of the recent decade. We expand on the wood energy market over time in the next section.

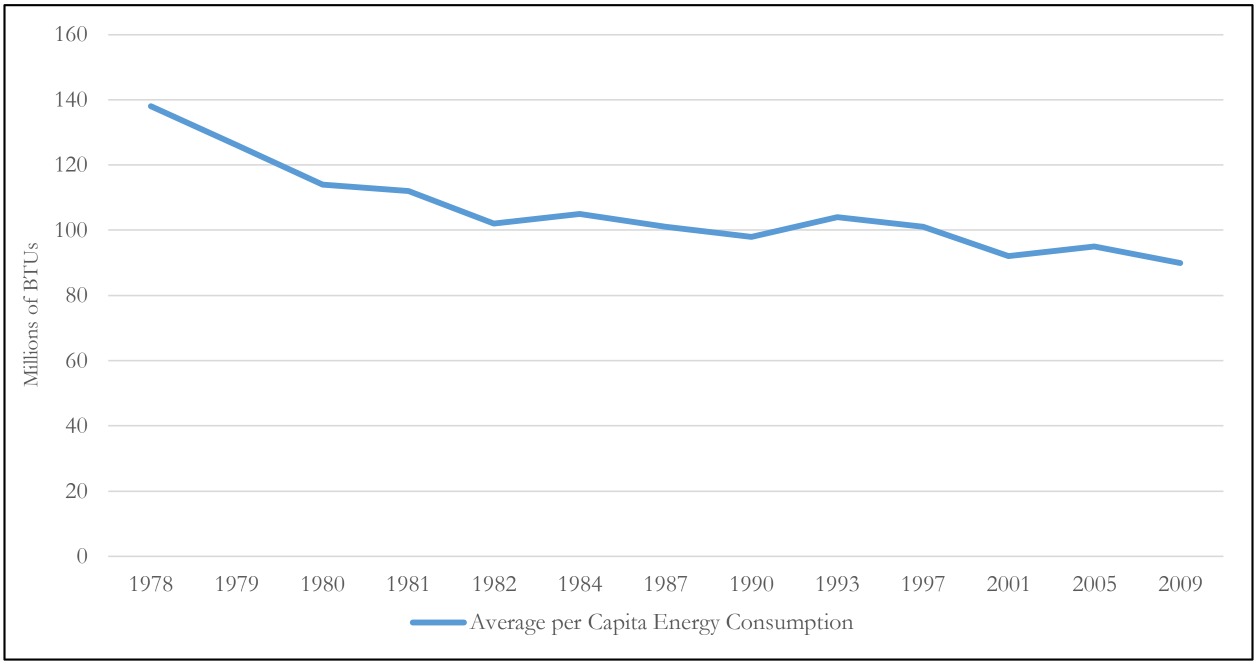

Figure 1.3. Yearly Average Energy Consumption Per Capita, Millions of BTUs, 1978 – 2009. Source: U.S. Energy Information Administration. Annual Energy Review. (2011, 2012a).

1.2. Residential Demand for Wood Energy

U.S. residential demand for wood-based energy is on a long-term, 65-year decline (Figure 1.4), with periods of shocks to demand appearing to coincide with the OPEC oil embargo in 1973 – 1974, and more recent shocks from natural gas, oil, and liquid propane supply/price shocks in 2007 – 2008. A closer look at recent years, however, raises the question of whether residential demand for fuelwood is increasing (U.S. Energy Information Administration, 2010f). Recent research on current demand for fuelwood has yet to be analyzed. Looking closely at Figure 1.4, demand appears to be trending upward during the last decade. Possible reasons for the recent increase include federal tax policies to promote green energy use, modern appliances that are more efficient than older units and directly release heat into homes, and heightened public consumer awareness of the impact energy has on climate change (Song et al., 2012a; Favero and Mendelsohn, 2014). It is also possible the upward trend is due to heightened awareness of environmental safety, climate change, energy and economic stability. Consumers may shift their energy preferences towards fuelwood sources thought to benefit society as a whole, as current usage of fossil fuels has implications to global warming due to the release of trapped hydrocarbons. Even the preference of supporting local economies and stabilizing energy security may be contributing to an increase in demand, as problems arise from an aging energy infrastructure and concerns of terroristic threats on general global energy supply. Whatever is causing the demand to increase, the need to estimate effects of general determinants of fuelwood demand is needed to further support the development and expansion of a wood energy industry.

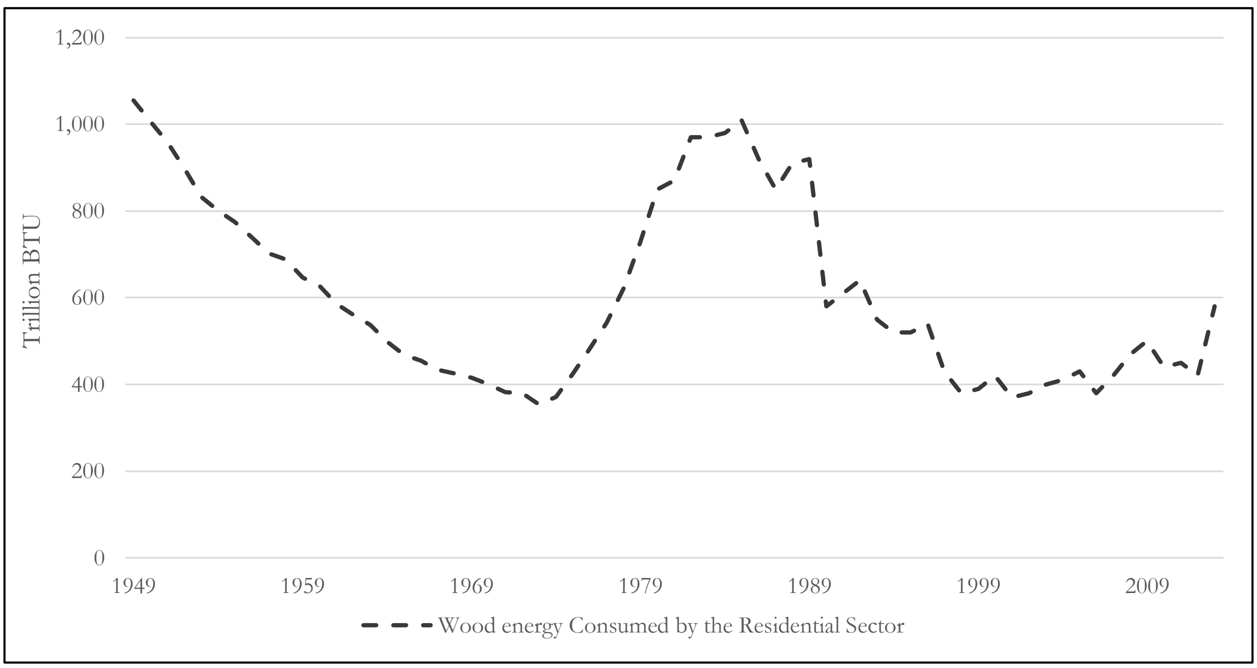

Figure 1.4. U.S. Residential Demand for Wood Energy, 1949 – 2012.
Source: U.S. Energy Information Administration. Monthly Energy Review. (2014c).

An example of possible wood demand increases from these causes is currently occurring in Europe. As traditional fossil fuel energy market prices continue to react in a volatile manner, many nations in Europe are increasing their demand for wood energy sources (U.S. Energy Information Administration, 2014a). This comes largely due to changes in state energy policies, aimed at curbing fossil fuel usage and dependence by imposing taxes, as well as dwindling local supplies. As such, they've turned to sourcing from external nations to satisfy their demand (U.S. Energy Information Administration, 2014d). We therefore examine U.S. residential household response to market prices and incomes to deduce whether the domestic market reacts in similar manner, as the stated possible reasons predict a way to obtain a reactive market pricing mechanism for firms in the fuelwood business.

We also look for possible reasons why U.S. residents would demand wood energy in the first place. Pre-existing conditions, such as the location of the home and its proximity to forests, is influential. A lack of urbanization within remote rural regions can also impact the decision of fuel usage. Both of these cases essentially leave the resident to seek self-sustainable sources like wood or ag-based residuals, or (if available) to enlist delivery of modern fossil fuels, such as fuel oil and liquid propane.

Availability and reliability of fuels used is also a possible reason for usage. We question this within the decision outcome as a means to ultimately choose the utility obtained from fuelwood over fossil fuels. Price is also a factor, as a less expensive fuel should prove economically viable during usage. Other determinants of use and choice are explored in Chapter 3 on the methodology behind selectivity as well.

1.3. A Concern for Energy Sources

It's well documented that wood energy was used centuries ago for heating and cooking needs (Yergin, 2012). As societies increased their economic wellbeing, shifts in energy resource usage trended towards more easily accessible, versatile, robust and reliable sources. Those sources stem from fossil fuels with excellent energy traits: rich in hydrocarbon chains, transportable, and easily adaptable to be utilized in different scenarios. Because of these characteristics, fossil fuel energy usage within the United States met approximately 83% of the total energy consumed in 2009

(U.S. Energy Information Administration, 2012b). U.S. residential households in that same year consumed about 22% of all energy produced (U.S. Energy Information Administration, 2010a); fossil fuels accounted for 90.7% of total primary energy residential homes consumed (U.S. Energy Information Administration, 2014e).

With this large of a proportion of energy relying upon fossil fuels, problems involving the supply and distribution of energy can easily cause stress to the economy as the proportion of energy supply is concentrated. Public and governmental agencies have developed energy policies for transitioning to renewable energy resources that are domestic and sustainable in the long-term (Office of Management and Budget, n.d.). When an advanced economy relies on a large proportion of energy usage for economic activity, serious implications occur from economic shocks that result when supplies are disrupted. The effects of the 1973 OPEC oil embargo bring distant memories of reliance upon foreign resources that controlled large proportions of the supply, creating concerns on energy dependence and a need to increase domestic production for supply security.

Recent natural disasters have created domestic constraints. In 2005, The Gulf of Mexico region experienced Hurricane Katrina, which shocked oil supplies to the United States. Hurricane Sandy in October 2012 severely damaged the upper-east coast's infrastructure, causing electricity and natural gas transmission and distribution problems. Recent man-made disasters have also contributed to supply issues as well. The April 2010 BP Deep Water Horizon disaster caused an estimated 4.9 million barrels of oil to flow for 87 days at a depth of 5,000 feet below sea level (U.S. Coast Guard, 2011). Furthermore, the 2014 winter brought unusually cold weather to regions where natural gas transmission lines run, propane usage being dependent upon for crop drying, and closure of the Canadian pipeline (U.S. Energy Information Administration, 2014f, . Multiple lines exploded during the winter, causing supply shocks to these cold regions (U.S. Energy Information Administration, 2014a). It is because of supply shock instances such as these that the desire for other, more stable energy resources exists. Fuelwood and other locally sourced biomass products provide a solution to fossil fuels. Having many firms provide regional or local wood and agriculture-based energy products creates a solution to those energy needs that are not as reliant on single sources, long distance transport, or result in environmental concerns if disaster strikes. Furthermore, it provides solutions to rural economic development to those who rely on delivery of fuel oil and propane.

1.4. A Possible Solution

These concerns help demonstrate the potential renewable resources have for society in providing both environmental and economic benefits. One key trait renewable energy sources satisfy, and fossil fuels fail to provide, is that they are renewable. The fact of a finite amount of fossil fuel energy is not debatable. Another trait that fossil fuel resources fail to satisfy is that they are not a sustainable form of energy. Merriam-Webster defines sustainable as a method of harvesting or using a resource so that the resource is not depleted or permanently damaged (*Merriam-Webster.com*, 2014). We now see an increase of pressure from society to shift resources toward renewable and sustainable energies. As awareness of the benefits increase we may lead ourselves to an increase in overall utility of energy use for both individual and society.

But this increasing pressure from society brings a deterministic issue of what single resource, or multiple resources, present better traits satisfying all energy characteristics (safe, portable, durable, dense, versatile, etc.). If fuelwood is the energy source to be used in satisfying heating requirements, appliances should also to be considered. Different types of fuelwood heating systems are available, such as heating stoves, forced air central systems, boilers, and fireplaces, which all have different considerations in regards to efficiency, longevity, and periodic maintenance and timely/continued manual feeding of fuelwood. These traits are all to be considered by a residence in the decision making process of using the source.

Even when we consider appliances, overall benefits are to be considered as well. Those individuals or special interest groups arguing the vast amount of reserves available, either from petroleum, natural gas, coal, or others, seek to bring attention to the lack of needing to switch energy sources now. Why switch when we have large resources of energy and create other economic concerns, such as switching and investment costs that may not be fully recovered for many years after implementation? The counter argument asks why wait to switch towards sustainable and renewable sources when no one knows the true cost of fossil fuels on the environment. Using fuelwood for heating easily removes the concern of releasing trapped carbon dioxide and other greenhouse gases associated with climate change into the environment, since it is considered carbon-neutral, versus fossil fuels that contain carbon trapped for millions of years. Also, as prior discussed, we've already experienced issues involving energy security from supply disruptions. If a finite source of energy is used to its end, then society must seek an alternative substitute or face potential economic security issues from complete reliance and dependence upon said depleted source. Essentially, why wait when we can begin a transition to a heating source that was once the only means of providing heat for a residential home?

This brief background leads to the possibility of biomass energy and its wood derivatives to be utilized as energy sources with a possible increase in near-term demand. Thus, if biomass and biofuels provide a few solutions for consumption requirements that fossil fuels don't, we need to understand potential demand by estimating it through correcting for selectivity. With today's available technology, the world's population cannot utilize fossil fuels without depleting the resource, nor can we replenish at a sustainable rate. A transition to biomass-derived energy for heating addresses home heating as part of this problem by being able to replenish what is used. Waste can be utilized as well, aiding in a recycling process and reducing environmental degradation. Song et al., (2012a) find the U.S. consumes only one-third of estimated annual, sustainable supplies.

We've just explored the background of residential demand for energy. The next section provides a layout of the research, support for conducting it, methods and data used in the research, as well as anticipated outcomes.

1.5. Overview of Research

In order to determine how to approach this fuelwood demand estimation, we look at using a Heckman two-stage estimation: what possibly consists of a selection criteria, and what determines how much fuel is used. With so many fuel options available for home heating needs, the residential household faces a choice of whether to utilize wood-based bioenergy in the first place. This study addresses selectivity, i.e., the ability of residential households to choose to use wood energy, which creates bias within demand estimation particularly within traditional OLS procedures. A possible reason for this decision outcome is due to fuelwood appliance availability, the specific chosen space to heat, or the type of fuelwood used. Furthermore, the use of other methods, such as the Tobit method (Tobin, 1958), doesn't include all observations in the estimation procedure (Arabatzis and Malesios, 2011). To solve sample-population bias, we use a two-step Heckman selectivity method to estimate U.S. residential wood energy demand in year 2009. Data obtained from the U.S. Energy Information Administration is used.

1.6. Problem Statement and Objective

To have a good understanding of how demand for energy will shift through time, continued studies on each segment must be completed periodically. The most recent study on U.S. fuelwood demand found was conducted on data that is almost a decade old. We therefore propose a study on residential wood energy demand using data from 2009, which is the most recent data available to date, as a means to provide updated information on how responsive the typical household is to the wood fuel market.

The objective of the research is to estimate U.S. residential demand for fuelwood while taking selectivity into consideration. Selectivity measures must be accounted for in order to appropriately correct for bias and inconsistency found from the decision making process (Heckman, 1979). Thus, our primary focus is to analyze the residential wood energy market decision process, correct for choice with a Heckman method to account for bias decision making, then estimate effects of determinants of how much fuel is used for heating needs. This study will provide insight as to how a market responds to an ever changing, and volatile, energy market.

1.7. Conceptual Framework

We believe the residential household faces a decision: whether to use fuelwood or not, generally determined by extra factors not necessarily directly impacting the quantity consumed other than a positive quantity. Because of this decision outcome, our theoretical framework revolves around a selectivity-correcting demand model. Research presented by Heckman (1976, 1979) provides us with the theory and method for incorporating a means to estimate decision-making omitted variables. Thus, a two-step selectivity correction model is developed in order to correct for this consumer-induced selectivity bias. As fuelwood users may observe, or exhibit, decision instances where consumption is predetermined, not correcting for such practice creates problems likely leading to downward-bias estimates from the omission of such factors (Heckman, 1979).

1.8. Methods and Procedure

Data from the 2009 Residential Energy Consumption Survey obtained from the U.S. Energy Information Administration is used. This information is the latest and largest survey on U.S. residential household demand for many energy uses. While it contains much of the data needed, it does not contain price paid for wood fuel. We focus on aggregate alternative energy prices to infer on demand responses, similar to Song et al. (2012b).

A simple two-step estimation procedure is used, as presented by Heckman (1976, 1979). It is designed to correct for selectivity bias within the demand estimation procedure through the use of an omitted variable term. This omitted term is estimated as a probable outcome for using fuelwood, the decision choice indicator. When this term is included within the second OLS step, there is a reduction of collinearity between the OLS demand equation and the missing variables and error term of the selection equation, the first step. We compare estimates of an OLS and Tobit (Tobin, 1958) model of the second-step demand equation without the correction term as well.

1.9. Research Questions

While the primary objective of this study is to estimate fuelwood demand by correcting for selectivity, the effects from changes in alternative energy prices on wood energy demand and household income effects on quantity demand are explored as well. The following three questions summarize the driving factors for this study to provide additional insight:

1.) Is fuelwood a normal or luxury good?

2.) What effect do energy prices have on residential wood energy demand?

3.) What impacts do climate and geographic location have on demand?

4.) Can we rely on these findings to accurately forecast potential demand?

1.10. Hypotheses

We explore possible outcomes of specific variables that drive the decision to use fuelwood and how much fuel is used. These outcomes provide insight in determining future forecasts for residential demand for fuelwood, aiding in fuel production decisions often made well in advance. We hypothesize the following outcomes:

1.) We look at the effects of income on fuelwood usage as well, allowing us to determine if it is a normal or luxury good. Fuelwood can be thought of both a necessity item and as a preferential aesthetic trait for home ambiance. Also, an increase in income doesn't necessarily lead to an increase in energy usage; an increase of income may lead to larger homes, however. We hypothesize as a household's income increases, the usage and quantity decrease making fuelwood a normal good. Further, we look to an increase in heating space to drive demand as well.

2.) Given the fuelwood market consumes 3.2% of overall demand, we look to see if alternative fuel prices are significant on the use of and quantity of fuelwood used. As non-wood energy prices increase, both the likelihood and quantity of fuelwood consumed increase.

3.) Historical climate data and geographic location can provide specific indications of fuel types used for residential homes. We believe historically colder climates increase the probability and quantity of fuelwood used, relative to other fuels. Geographic locations associated with cold regions, as well as wood resources, lead to an increase in probability and quantity as well.

1.11. Organization of Study

We explore this study in six chapters. The first chapter, just discussed, provides a brief history of residential energy demand and explored reasons to believe selection plays a role in deciding to use fuelwood as a heating source. We focused first on residential energy demand as a whole, then on wood energy demand. We also presented a cause for concern and the possible solution wood-based renewables have for our energy dependence. Chapter one concluded with an overview of this study's purpose, including the research questions and hypotheses to be tested.

We present Chapter Two next by looking at the literature on three specific areas. The overall demand for energy by the residential sector is viewed first as the backbone to understand selectivity issues. Then, we move onto seminal wood fuel research. These two topics are similar but have different objectives as we observe that wood energy is seldom studied relative to fossil fuel based energies (or other solar/wind/ethanol renewables). Finally, the third part of chapter two looks into methods other studies use for energy demand estimation. This is where we emphasize selection criteria methods that correct for omitted variables, particularly the Heckman (1979) two-step procedure.

The theory behind Heckman's two-step procedure is presented in Chapter Three, accompanied by the empirical application. The theoretical aspect behind the correction procedure builds the foundation behind the reason to correct for residential energy selection. At the same time, we briefly discuss data usage then we apply the method to this study's empirical reasoning. This gives us the ability to understand the underlying econometric procedure so other areas of similar interest can have this method applied to as well.

A detailed look at the data source utilized for our research is found in Chapter Four. Very few public sources exist that provide such microeconomic data in detail. We are fortunate enough to have data available from the U.S. Energy Information Administration's (EIA) 2009 Residential Energy Consumption Survey (RECS) dataset. It is the most recent available source to date. Thus, we have the ability to update another study (Song et al., 2012b), which greatly influenced this study.

Chapter Five and Six close the study. Results are found in chapter five, where we look at those factors found to be significant and focus on the selectivity correction term. The variables identified within the research overview below (non-wood energy price and household income levels), are also emphasized while presenting the entire findings. Finally, chapter six looks at the study's conclusions and implications. Suggestions for further studies that can improve and complement our findings and conclusions end the paper. In the next chapter, we look at prior literature on the subject matter.

CHAPTER 2: REVIEW OF RELATED LITERATURE

The purpose of this paper is to estimate domestic residential fuelwood demand. There are several past and recent studies on U.S. residential fuelwood demand estimation, but few consider selectivity on recent available data. This section begins by reviewing demand estimation when selectivity is evident. We then shift focus to fuelwood demand studies. Two different views are itemized as they aid in identifying specific attributes required for understanding the implications involved with selectivity: (1) identifying determinants leading to the decision of a residence choosing to use fuelwood, and (2) identifying determinants leading to a residence consuming a quantity of fuelwood.

In reviewing these studies, reasons why other methods and variables are not a good fit for this study are discussed. As we reason why correcting for selection is required for fuelwood demand estimation in the residential sector, an understanding of what potential market demand exists should be uncovered for fuelwood firms seeking to market their products. The chapter ends with a discussion of support for the use of a selectivity correction demand model.

2.1. Demand Estimation with Selectivity

The Tobit model (Tobin, 1958), provides a way to estimate possible unobserved, e.g. latent, demand. Tobin quotes *Charlotte's Web* within his introduction describing the issue behind trying to estimate something that doesn't exist, "less than nothing" (p. 24). This latent demand can occur for many reasons. For example, goods may be unattainable due to geographical location or state control of imports and exports. Because of instances like this, survey data on demand for goods can exhibit a common gathering of observed variables near one or both ends of extreme values. Tobin uses luxury goods as a simple example, as many lower income individuals generally desire and/or have preference for those goods. Without the level of disposable income being available, the individual is unable to obtain said luxurious goods. The Tobit modeling procedure estimates what possible demand exists if the good or service was attainable to the consumer in the first place. While this method provides an opportunity to estimate what demand exists within the fuelwood market, recall from the introduction that the market share for fuelwood is approximately 3.2% of the overall energy market within the residential sector, as most of the population doesn't consume fuelwood. Thus, we find it difficult to imagine the entire population may demand fuelwood if it were "attainable". We argue that once the choice has been determined to consume fuelwood, a bias effect on the quantity consumed will exist. A set of factors possibly exist influencing the decision to demand in the first place, which need to be accounted for in demand estimation to avoid bias results (Heckman, 1979).

Tobin's seminal work led to Heckman's (1976, 1979) two-step procedure, as well as Dubin and McFadden's (1984) extended work on multiple choice two-step estimation. The Heckman two-step selectivity procedure establishes a method to correct for bias from omitted decision variables in dichotomous choice outcomes. Heckman (1976, 1979) proposes a vector of decisive factors and their associated non-serially correlated errors lead to the end decision to consume a good; he argues choice decisions are correlated, which can lead to a lack of independence between the variables that cause biased results if the factors are omitted in an OLS procedure. Without such correction, the potential of omitted variables due to sampling errors within the population result, causing sample selection bias problems within regular regression fittings. In essence, a sampled population indicates a quantity demanded x_{1i}, which may do so from predetermined preferences. Other observations go without demanding x_{1i}, driven to its alternative x_{2j}. Due to the inability to demand from choices within the selection phase, sample bias results in zero demand for x_{1i}.

As a means to correct such bias, Heckman considers a generalized two-stage two-choice outcome model. To do so, a simple two-stage correction method solves for the resulting specification error that traditional OLS would bring. A choice outcome on the entire sample estimated by a probit regression yields the probability of choosing y_i, which then estimates a correction term deemed λ_i to be inserted within an OLS regression run on the subsample y_{1i} (Heckman, 1979). This method has been adopted and adapted into many different topics procedures (Arabatzis and Malesios, 2011; Hardie and Hassan, 1986; Mackenzie and Weaver, 1986). The procedure is simple, straight-forward, and included within nearly all econometric software available to date. Because of this ease, we use this procedure to estimate regressor effects on fuelwood demand while correcting for selectivity of the end quantity amount consumed. We discuss in further detail this methodology in Chapter 3, but Heckman's work is further extended within other studies completed after his work. We continue with other demand procedures and why they aren't used below.

Dubin and McFadden (1984) extend Heckman's method by incorporating mutually exclusive multinomial logistic estimation as the first stage in selectivity correction on 1975 electricity consumption data obtained from the Washington Center for Metropolitan Studies. By incorporating multinomial logistical probabilities for a selected choice, therefore correcting for such potential bias within a polychotomous selection outcome, the subsample of data on that particular usage will yield unbiased and consistent OLS regression results. Without correcting for such selection bias, estimated results are statistically inconsistent and compromised (Dubin and McFadden, 1984). They argue the consumer may observe multiple mutually exclusive decision outcomes, where each outcome is also not dependent upon each other following an assumed IIA process. Since residential homeowners do observe multiple options for satisfying their heating needs, this procedure gives more substance within theory and application. This method is also found to be more robust to Heckman's method (Dubin and McFadden, 1984; Bourguignon, Fournier, and Gurgand, 2007). This method, however computationally challenging and not found within available software packages, is provided within a user-written program for Stata by Bourguignon et al. (2007).

An assumption is made, however, within the Dubin and McFadden method: multivariate normal distribution of the error term found in the Heckman choice model. If the disturbance term doesn't follow this assumption, further bias of the estimates result (Lee, 1982). While the topic of study is not of direct relevance to this thesis study, Lee looks at a method of correcting such potential outcome by relaxing this assumption (1983).

With both Dubin and McFadden's, as well as Lee's methods for correcting bias available, questions arise regarding which is more robust and consistent given their assumptions. To aid in determining which model is sufficient for the application, Schmertmann (1994) presents a clarified review and Monte Carlo study on the two methods to display which should be used. He finds inefficiency within the Generalized Heckman procedures (GH) possibly due to multicollinearity, and an inability of the Lee procedures to be robust when assumptions are violated within the data generating process (Schmertmann, 1994). Vella (1998) provides another review of prior studies on selection bias correction methods. Further, Vella explores the history of sample selectivity bias lineage through time, beginning with Heckman (the quintessential "godfather" of selection correction, if you will), as the years following the 1979 work on the probit correction led to many studies that extended his work (Vella, 1998). As with Schmertmann's work, Vella provides clarity and examples on the methods without divulging specific properties or efficiency comparisons (Vella, 1998). These two reviews greatly aid in a means to clarify prior work, the benefits and deficiencies of the methods proposed, as well as a means to incorporate the work with modern computing power and data storage capabilities.

Another estimation procedure is proposed by Dahl (2002), as a method to cope with the infeasible or restrictive means of prior work from imposing assumptions on the dimensionality aspect of the probability outcome from multiple decisions. Dahl adds parametric and semiparametric models to reduce this dimensionality, which results in a relaxation of the IIA property typically found within a multi-conditional outcome model (Dahl, 2002). Dahl provides this as a means to modify Lee's (1982, 1983) method in order to solve U.S. workers willingness to enter the job market effect on wages and returns to higher education. Dahl concludes the method finds evidence of correction within the issue behind "mobility" of a worker's willingness not to just enter the market, but also move within the country in seeking employment and higher wages. While the topic this method is applied to is of no relevance to this study, it is easily transferable into any scenario.

Currently, the Tobit and Heckman models are the two simplest models to use on a surveyed population on fuelwood demand. We've also briefly discussed the use of Heckman's two-step selectivity correction procedure for this study. While a traditional multivariate OLS regression yields the best, linear, and unbiased estimator if all conditions are met, it poses potential issues of bias and inconsistency of the estimators when the sampled data is incomplete or not representative of the true population, i.e., selectivity of survey data is present (Heckman, 1976, 1979; Dubin and McFadden, 1984). This may come from latent demand, where demand exists but isn't measureable due to uncontrollable circumstances, or selectivity of demand as a result of previous decisions leading to the use and quantity demanded. We now discuss other studies on other energy estimation procedures.

2.2. Other Energy Demand Studies

An earlier U.S. study conducted by the RAND Corporation on the growing concerns over residential energy demand and responses to price and energy substitutes questions prior literature on consumption and substitution responses to price changes (Anderson, 1973). Anderson provides this study to build knowledge on energy demand in an era where little records existed on the economics of energy efficiency, conservation, or substitution. Looking back, this era faced pressure for an increase in energy efficiency and a reduction in emissions over growing concerns of the negative effects of energy on the environment, hence the need to determine residential responses to energy price changes as government questioned energy policies (Anderson, 1973). Anderson uses data on the 50 states from 1960 to 1970, focusing on different primary energy types (coal, natural gas, propane, fuel oil, electricity, and kerosene), and on modeled determinants such as capital costs (price of appliances), household income, number of members residing within the household, being located in an urban population, structural type of the building, as well as temperatures during the summer and winter seasons. Two equations, supply and demand, are conducted independently and simultaneously in order to find own- and cross-price elasticities.

The importance of understanding the meaning of own-price elasticity is described by a means of a "conservation" measure to price, i.e., the residence energy conservation response to increases in energy prices (Anderson, 1973). Further, consideration of the substitutability between energy options is presented, as electricity likely isn't a substitute for oil-based energy within this era (Anderson, 1973). Anderson confirms indirectly estimated own-price elasticity estimates being negative, while cross-price on unlikely substitutes are also near zero, as hypothesized. An issue within this study is its inability to involve more modern estimation techniques, let alone data. Given the lack of knowledge and studies, possible selectivity issues aren't addressed fully as one would expect (as discussed in detail below). The summary of this literature isn't here to point out the fact that it is dated, but rather to include prior historical findings and concerns regarding energy pricing effects on consumption habits, both own- and cross-pricing effects. As such, we investigate the inclusion measures of income, geographical location, size, and climate within our model. Energy prices are an area of interest and will be discussed below.

To expand on energy demand modeling, literature is widely available (Jebaraj and Iniyan, 2006; Gately and Huntington, 2001; Kialashaki and Reisel, 2013; Song et al., 2012b; Kaza, 2010). Jebaraj and Iniyan (2006) provide a lengthy review of different types of models utilized within research prior to 2004. They find a wide range of energy models within the literature including forecasting, supply-demand, optimization, emission reduction, and decision-making computer simulations. Many models are found focusing on biomass energy, with an emphasis on location-based logistics, sustainable forecasting, and impacts of emissions. Gately and Huntington (2001) researched energy and oil demand among 96 OECD and Non-OECD energy countries. They found price and income elasticities were higher than other research, and Non-OECD countries have a higher long-run income elasticity than OECD countries.

Research on residential energy consumption has been studied in detail recently. Kialashaki and Reisel (2013) examined U.S. residential energy demand models to forecast changes from different growth scenarios. Their use of artificial neural network (ANN) and multiple linear regression (MLR) techniques found differing results, as the ANN model included recent economic recession impacts, lowering the forecast versus the MLR model that included regression trends. They believe residential consumption of energy will continue to decline in the wake of emission reduction concerns in the future.

Another study on U.S. energy demand is provided by Kaza (2010), who recognized the importance of residential energy consumption, as the aggregate impact it has on overall energy consumption amounts to about 22% of all U.S. energy consumed in 2008. Kaza used a quantile regression analysis in order to find specific factors that impact residential energy consumption the most across different ranges of variable measures. The focus of this research was on 2005 data from the U.S. Energy Information Administration's Residential Energy Consumption Survey (RECS), however, since the survey is for a single period, Kaza included trend results from prior year surveys.

The findings indicated energy consumption for residential space heating is on a general trend decline. Additional results indicate that rural households consume less energy for heat, heating square footage has a large impact on consumption relative to other variables, and price for energy has an inverse effect on consumption (price increase, consumption decreases). Newer homes are found to consume less energy for heat, likely due to efficiency and energy weatherization techniques.

The quantile regression analysis on energy consumption analyzes effects at different levels of impact rather than mean averages that traditional regressions estimate. The reasoning behind this is the variance associated with traditional variables (household size, geographical location, spacing of homes, number of appliances, etc.), doesn't prove to provide a beneficial analysis when mean averages are utilized in consideration of conservation policies (Kaza, 2010). Being able to analyze effects on the whole distribution of energy consumption, rather than the most-likely estimate, brings a greater understanding of policy impacts to other levels of possible observations. The estimated results indicate the size of a house is the greatest influencer on energy consumption habits. Geographic location, hot versus cold regions, is also a critical variable in energy consumption. The quantile results, focusing on the upper and lower 10% and 90% tails, indicate those of higher consumption amounts respond the greatest to increases in heating needs. Suggested policy analysis includes researching those who consume at the higher quantiles. This implies more efficient policy developments aimed at conservation to those who consume the most energy. None of these prior studies include means to measure, implement, nor correct for decisions the consumer can make between fuel choices and the bias it can create. While the quantile method is robust for Kaza's (2010) study on various energy options, we argue it doesn't hold when data appears to be censored or minimal relative to the entire sampled survey. Observation of fuelwood usage appear less than 12% of the total 2009 RECS surveyed individuals, making it difficult to estimate across several different ranges and retaining statistical significance for the entire population. Furthermore, there currently doesn't appear to be a method to incorporate selectivity correction within the procedure.

2.3. Residential Demand for Wood Energy

Seminal work on U.S. wood energy demand was conducted during a period of increasing demand due to rising costs of petroleum-based heating fuels (Hardie and Hassan, 1986; Mackenzie and Weaver, 1986). Residential demand for fuelwood increased drastically, nearly doubling in some regions. Hardie and Hassan (1986) conduct a two-step Heckman (1976, 1979) selectivity model for five U.S. regions, while using estimated average market fuelwood prices for those homes that provided survey responses of not exchanging currency for fuelwood within the 1980 EIA RECS data (Hardie and Hassan, 1986). This is done to avoid excluding fuelwood demand observations during the estimation procedure. The procedure uses heating space in square feet, heating degree days, number of household members, prices of both wood and non-wood energies, and type of heating system. The authors use similar measures to the variables just discussed, plus dummies for main types of fuel and regions within the consumption model. The use of the Heckman procedure comes at a time when the method is relatively new in providing insight for what was a small market during U.S. history. The results find cross-price elasticities for non-wood energies that indicate demand is fairly responsive. The authors suggest the choice model may perform better with better rational and specification, perhaps by having a survey designed with questions of choice involved. The authors' model does not find statistically significant income effects on consumption.

Mackenzie and Weaver (1986) focus their 2SLS household production model, which is based on Heckman's (1976, 1979) procedure, on estimating fuelwood demand within the state of Rhode Island. During this period, Rhode Island had a high population density, large forestry geography, and an apparent dependence of local consumption of fuelwood production. The authors use a telephone survey to obtain data on fuelwood consumption, head of householder's age, type of dwelling, non-wood energy prices, head of householder's education, acreage, income, estimated heating requirements, and how much forestry is within the vicinity. The survey provides opinion and objective based questions as well, including maintenance of burning wood, enjoying the handling of wood, aesthetics of burning wood, and how many years they've been burning wood. The survey finds average incomes of wood burning households to be higher than non-burning households. They also find fireplace users spend more on heating than non-burning homes, whereas users of stoves or furnaces burning wood spend less than non-burning households. The results indicate households burning for aesthetic purpose use less fuelwood with even lesser efficient appliances (Mackenzie and Weaver, 1986). Of use for this study, the authors estimate a negative Heckman (1979) lambda coefficient.

The next two studies—the most recent and relevant influence on this research—are provided by Song, Aguilar, Shifley, and Goerndt (2012a, 2012b), and focus on U.S. residential consumption of wood energy. These studies provide a framework behind determinants of U.S. residential fuelwood demand, along with historical trends leading to 2009.

Song et al. (2012b) use a Tobit model for latent demand on the U.S. Energy Information Administration's 2005 RECS data on wood energy consumption. The intention of the study is to address recent public attention for sustainable and renewable energies like fuelwood by focusing on developing eight separate models from the four U.S. Census regions, differentiated between rural and urban areas. Since the 2005 RECS doesn't provide wood prices, nor individual levels of incomes, a composite vector of non-wood energy prices is used to determine cross-price relationships, and a focus on mean incomes from categorical income variables provides a similar measure outlook. Other factors such as date of home construction and number of residences are used in addition to climate effects over the 2005 heating season.

In their analysis, the Tobit model of latent quantity of wood demand is defined as:

$$WOODij = max(0, WOODij^*), \tag{2.1}$$

with $WOODij^*$ defined as:

$$WOOD_{ij}^* = f(PNW_{ij}, NHM_{ij}, AGHH_{ij}, HINC_{ij}, HINC_{ij}^2, THSQ_{ij}, AHDD_{ij}, TOWN_{ij}, SUBURB_{ij}, RURAL_{ij}, \textstyle\sum_l{=}1,3,5,6,8 * D_l, \eta_{ij}). \tag{2.2}$$

Here, wood energy is consumed when a threshold level above zero quantity is observed, $WOODij^*$, or the resulting quantity is zero. Once it is observed, residential household consumption of wood is a function explained by a vector of non-wood energy prices *(PNW)*, number of household members *(NHM)*, age of the head of household *(AGHH)*, both the household income level and its squared value to measure change *(HINC and* $HINC^2$*)*, total square footage of the residence heated *(THSQ)*, average heating degree days *(AHDD)*, a set of dummies to describe whether the residence is within a town, suburb, or rural community with the exclusion of the city classification to avoid the dummy variable trap *(TOWN, SUBURB, and RURAL)*, and a final set of dummy variables to classify geographical location within the United States as classified by the Census Bureau's divisions *(*D_l*, with* l*=1 to 9)* (Song et al., 2012b). The model excludes dummies 2, 4, 7, and 9 to consider base geographical locations (four divisions in the census classification system, nine regions total). Region *i* and area *j* (urban or rural) are geographical descriptor statistics for those sampled within the survey.

Song et al. (2012b) find it can be inferred that rural homes operate as self-sustainers when it comes to wood burning for heat (70% of the respondents within the 2005 RECS are rural homes; urban homes are easier to network with public energy utilities). The results indicate rural locations use more wood energy than urban, colder regions utilize more wood energy, and price and income have the expected negative and positive economic effects on consumption habits, respectively. They were unable to determine own-price elasticity effects from wood energy due to the survey not having fuelwood price information. Furthermore, fuelwood is often found to be obtained at a very low cost, if not free on available owned land, thereby making own-price estimation possibly trivial (Song et al., 2012b).

A criticism of this study, we argue, is exclusion of the individual household making the predetermined decision to consume fuelwood. There is no indication of effects of other possible factors of choice leading to demanding a quantity of fuelwood. The Tobit (1958) model traditionally doesn't involve this step, but truncates the data as missing observations. The Heckman two-step correction method (1979) does correct for endogenous selectivity and includes all observations even when they indicate no demanded quantity, thereby treating the observation as a true zero demand (Heckman, 1979; Arabatzis and Malesios, 2011). Furthermore, the authors don't consider whether equipment capable of burning the fuel is available, nor the different spaces or uses the household could use the heat source. They don't include a dummy for type of fuelwood, as the majority of demand is found from logs over scraps or pellets. The authors could have discussed the possibility of using a dummy indicator for the use of pellets beyond statistics of the data.

Lastly, the use of a Tobit method doesn't necessarily fit the purpose of their paper. The authors seek to estimate residential demand for fuelwood by using a truncation method that estimates latent demand, i.e., demand that isn't observed due to unobservable and impassible market entry costs. The paper estimates latent demand in lieu of actual demand by essentially treating observations of no fuelwood demand as missing observations. The paper should use a selectivity correction method to determine why residential households choose to demand fuel and correct for this assumption of the Tobit method.

Another paper by Song, Aguilar, Shifley, and Goerndt (2012a), published at the same time as the other study, looks at factors affecting U.S. residential wood energy consumption over the years of 1967 to 2009 in order to model the effects those factors have on demand. They also inspect government policies that may have an impact on increasing the demand for residential wood energy consumption. When researching wood energy, it is important to observe which targeted markets have the greatest proportional impact on the market. The researchers focus on the residential sector, as it consumed roughly 23% of the U.S. wood energy in 2010 (Song et al., 2012a). Annual data from 1967 to 2009 are from a few public sources including the U.S. Census Bureau, Bureau of Economic Analysis, Bureau of Labor Statistics and the Energy Information Administration. The authors use an error correction model (ECM) to estimate both long and short-run coefficients that have an influence on residential energy demand. The ECM method provides insight into long and short-run responses of variables over time. Since we aren't estimating fuelwood demand over time, just for the survey year 2009, an ECM model isn't appropriate for our application. Coefficients in question are wood energy consumption, price of non-wood energy, number of houses occupied, annual income per-capita and annual heating degree days. The authors argue the inclusion of non-wood energy sources, which are considered substitutes, along with wood prices may lead to inflated variance and multicollinearity (Song et al., 2012a; Greene, 2003).

Their econometric results on factors that affect wood energy the most indicate both income per-capita and number of occupied houses are highly correlated with an estimated correlated coefficient of 0.95, thus necessitating the use of a reduced form; the authors removed both of these coefficients in question resulting in an increase in the adjusted R^2 and *F*-statistic (Song et al., 2012a). After the authors removed the two insignificant and highly correlated variables, the remaining results indicate U.S. wood energy consumption is sensitive to price changes in non-wood energies. However the sensitivity tests from this result do not prove a large impact on the overall consumption of wood energy within the 1967 to 2009 time period.

The Song et al. (2012a) study provides many items of use for this research. The finding of fuelwood consumption being sensitive to non-wood energy prices provides a means to include a price measure within the demand estimation. This is beneficial as the 2009 RECS doesn't include fuelwood prices, just like the 2005 RECS used for the Song et al. (2012a, 2012b) studies. The inclusion of significant division variables is applied in our study as well. We argue, however, the separation of observations into eight different models dilute the small impact non-wood price has on demand. Therefore, we correct for Song et al. (2012b) geographical and urban differences through mutually exclusive dummies in our model. Since we estimate with all observations in a single equation, selectivity within the survey can be estimated as a single variable; we are able to easily include a correction term for selectivity of fuelwood use over non-use without diluting observations into smaller sample sizes.

Two variables of particular interest for this study, prices and wages, are found to have interesting outcomes within the ECM model. Song et al. (2012a, 2012b) find in a relatively short period, two years or less, that a homeowner doesn't increase consumption of wood energy products if the price of non-wood energy factors increase. This is counter-intuitive to general economic theory for seeking substitutes of normal goods facing an increase in prices, but Song et al. (2012a) argue a good point; a homeowner's decision to consume energy is generally predetermined prior to market price changes, i.e., the homeowner already decided to consume the good even though external market forces would otherwise create pressure to seek a substitute. The authors also estimate wood consumption to continually decline in the long run, even with fossil fuel energy prices historically increasing (Song et al., 2012a). Wood is generally cheap, if not free, but the increase in non-wood prices appears to have no significant effect on consumption habits. When it comes to wages, furthermore, they determine higher wages don't appear to have a significant impact on the decrease in consumption either, suggesting wood energy is an inferior good (Song et al., 2012a).

A study conducted on residential demand for fuelwood within Greece (Arabatzis and Malesios, 2011) uses three methods of demand estimation to establish determinants of choice upon demand. Their use of GLM, Tobit, and Heckman models aid in determining what variables are not only statistically significant, and find that the Heckman procedure provides estimates of greater reliability. This is due to the procedure not excluding observations where wood isn't demanded, which by involving all observations within the probit estimation, correction of selectivity bias is performed (Arabatzis and Malesios, 2011; Heckman, 1976, 1979). Their Heckman (1979) estimation yields a negative lambda coefficient like the prior study by Mackenzie and Weaver (1986).

A more recent French study by Couture, Garcia, and Reynaud (2012) on residential fuelwood consumption seeks to model fuelwood as a secondary backup source of home heating by correcting for an endogenous choice using the Dubin and McFadden (1984) method of first stage multinomial choice outcome correcting. They argue the lack of studies from developed countries on fuelwood demand warrants a study using multiple energy sources for home heating needs, especially since those countries have the ability to obtain multiple fuels for heating. They believe diversifying beyond fossil fuel energy is a solution for reducing exposure to market price and supply volatility in the near future, as oil supplies begin to dwindle. Through the use of survey data collected from 2004 – 2005, they find a negative demand response to income when fuelwood is used as the primary heating source.

The first part of their estimation procedure determines whether an endogenous choice has a significant impact on fuelwood demand. The authors use housing characteristics found within the data, age of head of household, altitude, a gas connection, age of building, size of space to be heated, homeowner's profession level, and equipment that uses wood, along with other opinion-based questions on topics like renewable energy and whether fuelwood use is in decline. Of these variables, they find availability of equipment, gas connections, and modern homes being statistically significant across all choice outcomes. Price (obtained through a hedonic pricing model) and income are found to be significant for all but those instances where fuel oil is used as the main heating source. The authors do note the issue of the DMF method, as it requires the strong assumption of each choice being IIA of each other, and by using the Hausman test for each case they cannot reject the null hypothesis of the estimated outcomes being truly independent of each other if an outcome wasn't available at the time of decision. This brings into question the validity of using the DMF procedure, another reason why we choose not to use this method beyond the complexity of programming within Stata. However, we do use the results as possible indication and insight for our hypothesis testing.

This does lead to their conclusion of fuelwood being observed as a necessity over a luxurious good; lower income households are more likely to use fuelwood over wealthy homes. They also find price of fuelwood has a probability effect on choice, as well as age of head of householder and building type having significant effects on choice. Consumption is found to vary across uses, as a negative price elasticity of -0.42 is estimated for those who use fuelwood in a main heating space; fuelwood used as a backup source is found to be insignificant. The study suggests the use of other prices to further determine substitutability effects form non-wood energies. This study provides key hypothesis ideas for similar anticipated outcomes within the U.S. residential setting. Due to strong assumption of the IIA hypothesis, we find it difficult the choice of energy isn't strongly linked to substitutes or alternatives. Furthermore, the complexity of using the DMF method and time limitations of this study to not conduct our own survey for incorporating such complex choice outcomes, we suggest an extended study of this topic for any possible additional insight through the use of the DMF estimation procedure, use of the Hausman test for IIA, or other possible tests to determine validity of the method and data.

Another final wood energy study we discuss focuses on cost and environmental impact from switching traditional wood sources for wood pellets for residential heating needs in Canada's British Columbia (Pa, Bi, and Sokhansanj, 2013). This is a concern, as low efficiency and high emission characteristics of current residential appliances contribute approximately 10.5% of residential heating needs (Pa et al., 2013). Furthermore, state policy and local firms are promoting the positive effects of pellet stoves and furnaces in order to increase sustainability along with general energy efficiency and lowering localized emissions of generating heat from wood. Through a streamlined life cycle analysis, the authors are able to determine overall effects of health concerns, capital investment replacement costs, as well as efficiency differences from switching. Results indicate a decrease in wood energy consumption of about 53.7% due to an increase in modern appliance efficiency, as well as a reduction in environmental emissions between 17% and 95% depending on different scenarios. The switch to wood pellets are estimated at providing a lifetime cost savings of around $749 million. Savings from a reduction in costs associated with emissions reduction are estimated as well, however irrelevant for this research on demand estimation. The researchers determine that savings are not experienced for all users, as the type of appliance and installation location effects the returns and payback period from the investment in a pellet stove. The application of consumer choice on appliance and space usage needs to be accounted for in order to accurately estimate the quantity of wood energy demanded. Doing so without this adjustment may prove inconsistent and bias results (Heckman, 1979; Dubin and McFadden, 1984). From this study, we include pellets as an estimator to estimate average effects on demand if pellets are used over logs or other goods to confirm if consumption is lower.

2.4. Literature Review Summary

We just explored prior studies on wood energy demand within the U.S. These studies offer insight into how to econometrically include determinants of demand (Song et al., 2012a, 2012b; Couture et al., 2012; Arabatzis and Malesios, 2011). We find the methods Song et al. (2012a, 2012b) use, the ECM and Tobit models, don't completely answer a residential homeowner's influence on decision to use fuelwood in the first place. Without estimating factors that determine the decision outcome, biased estimates of regressors result (Heckman, 1976, 1979). The Dubin and McFadden (1984) multinomial method appears to work for answering multiple fuel options the residence face, but the complexity of it makes it difficult to use with available data[5].

Being able to study the impacts consumer selectivity will aid in helping to establish effects of market prices and temperatures, while verifying if income confirms wood energy is an inferior good. This enables biomass firms to target appropriate customers. As foreign demand continues to increase in light of policy and diminishing resources (U.S. energy Information Administration, 2014d; Peng et al., 2010; Spelter and Toth, 2009), and domestic energy security is of public concern (U.S. Energy Information Administration, 2014a), domestic wood energy firms can shift their focus on how to optimize their operations through appropriate energy tax policies and lobbying efforts (Song et al., 2012a).

[5] The STATA program provided by Bourguignon, Fournier, and Gurgand (2007) provides a means to estimate multiple choice outcomes. However, the application doesn't appear to work with the EIA RECS (2009) data nor model developed within this paper.

CHAPTER 3: THEORY AND METHODS

This study seeks to estimate U.S. residential demand for fuelwood in the presence of selectivity. To understand the method behind the model developed, we briefly discuss Chapter 4 on data now and cover it in further detail afterwards. An existing survey dataset, provided by the U.S. Energy Information Administration's 2009 Residential Energy Consumption Survey (RECS), essentially drives the method and model of demand estimation in this study. Limited observations indicating use of fuelwood (less than 12%) and few questions of direct relevance to fuelwood usage are primary limitations of the dataset.

The data generating process of a survey relies on the researcher's questions and expected responses. When a researcher conducts a survey of households to estimate potential demand for a product with demand that is relatively small, such as fuelwood, there exists a higher possibility of estimation problems that may lead to bias. The consumer of fuelwood, or any good for that matter, can also induce selectivity bias when deciding to use the good. Since this survey seeks to answer questions regarding the population as a whole from a small subset on consumption habits, the thin market share fuelwood consumes, and a surveyor looks at an even smaller sample of the total population, there exists the possibility of not obtaining an unbiased estimate of the demand model (Heckman, 1976, 1979; Dubin and McFadden, 1984; Vella, 1998; Bourguignon, Fournier, and Gurgand, 2007). This results from either individual or researcher selectivity, which is concerning as traditional OLS estimation doesn't account for non-users of fuelwood (Arabatzis and Malesios, 2011). Firms looking to estimate market potential won't be able to obtain unbiased estimates, leading to incorrect production quantities and market placement of their products. Heckman's (1979) two-step selectivity correcting procedure reduces this bias. Here we explore the theory and methods behind Heckman's procedure applied to our empirical model.

3.1. Theory: User Induced Selectivity

Prior studies exhibit a pattern for estimating residential fuelwood demand, centered on correcting for selectivity. This may come from many factors, those of which likely are predetermined aspects of home building traits and/or personal preference for comfort and convenience (Quigley and Rubinfeld, 1989). A method to estimate and correct for predetermined decision making is provided by Heckman (1976, 1979), through a two-step estimation procedure. As we face fuelwood demand estimation as a whole and not at a comparison level against alternatives, the theoretical framework of this research fits Heckman's (1979) model, who summarizes the situation we explore in a single word: selectivity. When a consumer is led to use a specific product or service, in this case fuelwood, there exists an opportunity for self-selection induced bias to occur; the quantity of fuelwood demanded in a given residential heating season is likely decided upon by direct or indirect factors influencing the residence to use a quantity greater than zero.

For this to occur, we experience two possibly almost simultaneously determined decisions, on the selection of fuel and the quantity to be used. For instance, when the cost of entry for consumption of a good is already achieved, whether this requires an investment or fee, the consumer then already made the decision to consume the good in a quantity greater than zero. Similarly, the cost of entry may essentially be zero (the home may have the appliance installed or the fuel is available at marginal costs near zero), but conveniences of goods with higher marginal costs may bring personal preference and greater utility causing the consumer not to consume the good, yielding zero demand. In either case, to demand or not demand, the factors influencing a decision to consume must not be omitted or bias estimates will occur (Heckman, 1979).

The bias results from the correlation of the omitted variables and associated errors, and the error terms of the demand equation. A solution for this dichotomous decision is theorized and addressed by a vector of variables presented by Heckman (1976, 1979) within a probit selection model. An estimate of the probable outcome for each observation leads to an inverse of Mills ratio calculation, which is used as an instrument for the omitted variables leading to decision making. Heckman describes the instrument as, “a monotone decreasing function of the probability that an observation is selected into the sample,” (p. 156). Furthermore, by using an estimate of the Mills ratio instrument, denoted as $\hat{\lambda}_i$, one can regress the probit equation to acquire the omitted variables’ coefficient estimate from the errors of the two equations, denoted $\frac{\sigma_{u\eta}}{\sigma_\eta^2}$, where u is the error term of the demand equation without correction of the omitted variable term and η is the error term of the probit selection equation. This correction term solves the downward bias that would occur without inclusion (Heckman, 1979).

3.2. Econometric Problem

Let d_{1i} represent the decision outcome of whether to use fuelwood for heating needs within a residence, where i represents the individual household surveyed on question use of fuelwood = 1, or no fuelwood = 0. This outcome is influenced by a set of decisive regressors, assumed to be endogenous to the quantity of fuelwood demanded, examined by the following equations similar to what Heckman (1979) presents:

$$y_{1i} = X_{ki}\beta_k + u_{1i}, \quad k = 1 \dots \text{N}, \tag{3.1}$$

$$d_{1i} = Z_{ji}\gamma_j + \eta_{1i}, \quad j = 1 \dots \text{M}. \tag{3.2}$$

Here we view a general demand equation (3.1), where the quantity of fuelwood, y_{1i} of individual i from the sampled survey on a question regarding total fuelwood quantity demanded = 1, with exogenous variables X_{ki}, of *k* among N total variables influencing the total quantity to be demanded. We also have a decision equation (3.2), for decision outcome $d_{1i} = 1$ to use fuelwood, $d_{1i} = 0$ to not use fuelwood, with exogenous decision regressors Z_{ji}, of *j* among M different variables for all observations i = 1 … I. Together, resident i will demand fuelwood when $d_{1i} = 1$, resulting in a quantity demanded, y_{1i}. Unknown parameter γ_j to be estimated influences the decision to use fuelwood for heating with the quantity of fuelwood in turn is influenced by unknown parameter β_k to be estimated.

The choice of d_{1i} occurs when its expected utility exceeds that of all the other alternatives, i.e., a residence will use wood energy once their indirect utility (equation 3.3) exceeds that of the utility and costs of all other alternative energy sources that are available at the time and location (Dubin and McFadden, 1984; and Vella, 1998):

$$d_{1i} = 1 \text{ when } E(V_i|d_{1i} = 1) > E(V_i|d_{1i} = 0), \tag{3.3}$$

$$d_{1i} = 0 = Z_{qi}\gamma_q + \eta_{qi}, \quad q = 1 \ldots \mathrm{R}, \tag{3.4}$$

where Z_{qi} is a vector of regressors that impact the decision for residence i not to demand fuelwood energy over alternative substitutes (3.4).

To extend this theory into fuelwood, Z_{ji} is a vector of decision regressors influencing the decision to consume fuelwood, $d_{1i} = 1$, i.e., those variables which influence the decision making process to use fuelwood for heating purposes of other alternatives, (3.4). Thus, $d_{1i} = 1$ is a selection rule to determine whether an individual residence i creates an observation of quantity demand for $y_{1i} > 0$. Once this is observed, the measure of β_k in (1.1), determined by regressors X_{ki}, influences the quantity of fuelwood demanded by household i. Finally, both u_{1i} and η_{1i} are residual error terms for the demand and decision model, respectively. Since the two equations aren't directly involved within the other equation, equation (3.2) is said to be omitted within (3.1) creating an endogenous relationship of correlation between these two error terms, thus selectivity bias exists (Heckman, 1979).

Let's explore Heckman's theory behind these errors. Because the predetermined choice to use fuelwood causes an endogenous relationship between X_{ki} and η_{1i}, i.e., $E\left[u_{1i}\middle|X_{ki}, Z_{ji}\right] \neq 0$, we observe Z_{ji} as a vector of omitted variables influencing the probability of the demand equation being greater than 0. When discussing the entire population from the sampled population, observations of y_{1i} exist for only a subsample; omitted dependent variables exist, i.e., data on y_{1i} occur for $I_1 < I$ (Heckman, 1979). Similar to what Heckman discusses, the decision to limit demand of fuelwood energy to a very limited number of the sampled population yields biased results in equation (3.6):

$$E(u_{1i}|X_{ki}, d_{1i} = 1) = E\left(u_{1i}\middle|X_{ki}, \eta_{1i} \geq -Z_{ji}\gamma_j\right). \tag{3.5}$$

$$E(y_{1i}|X_{ki}, d_{1i} \geq 0) = X_{ki}\beta_k + E\left(u_{1i}\middle|\eta_{1i} \geq -Z_{ji}\gamma_j\right). \tag{3.6}$$

Furthermore, for simplicity and generality reasons, $Z_{ji}\gamma_j$ can be expressed as a vector of decision variables, equation (3.7):

$$\Gamma = \{z_{1i}\gamma_1, z_{2i}\gamma_2, z_{3i}\gamma_3, \dots, z_{ji}\gamma_j\}. \tag{3.7}$$

To correct for this bias without a loss in efficiency, due to restricting estimation to only those observations where $y_{1i} > 0$ (observations that consume fuelwood), Heckman's (1979) two-step estimation procedure involves both equation (3.1) and (3.2) by: (1) the probability of observation *i* consuming fuelwood (d_{1i}) occurring through maximum likelihood estimation, and (2) using "the estimated values of the omitted variables … as regressors so that it is possible to estimate the behavioral functions of interest by simple methods", (Heckman, 1979), i.e., a standard OLS regression of equation (3.1).

In order to conduct this method, Heckman (1979) discloses a few properties and assumptions. First, the joint densities of u_{1i} and η_{1i} follow a bivariate normal density, we are then able to obtain the inverse of Mill's ratio, λ_i. This is shown by Bourguignon, Fournier, and Gurgand (2007) (herein referred to as BFG) and their generalization of the Heckman (1979) model:

$$E(u_{1i}|d_{1i} < 0, \Gamma) = \int \int_{-\infty}^{0} \frac{u_{1i} f(u_{1i}, d_{1i}|\Gamma)}{P(d_{ji} > 0|\Gamma)} d(d_{1i}) d(u_{1i}) = \lambda(\Gamma), \tag{3.8}$$

where BFG denote, "$f(u_{1i}, d_{1i}|\Gamma)$ is the conditional joint density of u_{1i} and d_{1i}" (Bourguignon et al., 2007). Furthermore, Heckman (1974) proposes u_{1i} and η_{1i} to be independently and identically distributed $N(0,\mathbf{\Sigma})$, with assumption (A.1):

$$\Sigma = \begin{pmatrix} \sigma_u^2 & \sigma_{u\eta} \\ \sigma_{u\eta} & \sigma_\eta^2 \end{pmatrix}, \tag{A.1}$$

where u_{1i} and η_{1i} being independent of Γ (Bourguignon et al., 2007). Obtaining the inverse Mill's ratio creates the ability to involve decision criteria for consuming wood energy, thus under these assumptions removing the endogeneity of omitted variable bias from predetermined choice decisions and involving all observations within the 2009 RECS survey. With (A.1) and (3.6), we are further able to formulate "the conditional expectation of a truncated random variable" (Vella, 1998),

$$E(u_{1i}|\Gamma, d_{1i} \geq 0) = \frac{\sigma_{u\eta}}{\sigma_\eta^2} \left\{ \frac{\phi(\Gamma)}{\Phi(\Gamma)} \right\}, \tag{3.9}$$

where ϕ and Φ are the probability and cumulative density functions, respectively, of the standard normal distribution for function Γ (equation 3.7). Because of these relationships and assumption (A.1), we are able to "capture the selectivity process" (Vella, 1998). Through a Probit estimation process, unknown selection parameters γ_j and σ_η, the independent variables that influence a residence decision on using fuelwood, $d_{1i} = 1$, and its associated standard deviations within the sample are able to be obtained (Heckman, 1979; Vella, 1998). Parameter $\left\{ \frac{\phi(\Gamma)}{\Phi(\Gamma)} \right\}$ is also known as the inverse Mill's ratio, λ_i. We then are able to include this term as a regressor, in conjuction with a new error term, w_{1i}, to obtain the selection corrected OLS theoretical model (Heckman, 1979; Vella, 1998; Bourguignon et al., 2007):

$$y_{1i} = X_{ki}\beta_k + \frac{\sigma_{u\eta}}{\sigma_\eta^2}\hat{\lambda}_i + w_{1i}. \tag{3.10}$$

3.3.1. Empirical Model: Selectivity

This section is devoted to applying the theory and problem just discussed towards analyzing U.S. residential demand for wood energy. Recall for this discussion that we are using the U.S. Energy Information Administration's 2009 RECS dataset, which essentially drives this model choice due to the underlying data generating process. It's important to note this, as well as some details on how the variables are derived in order to present the empirical model. From the theoretical model (3.10), we bring forth derived variables y_{1i}, X_{ki}, and β_k of the demand model equation (3.1), as well as derived variables d_{1i}, Z_{ji}, and γ_j of the decision model equation (3.2). We explain the application of those items which are considered determinants of residential wood energy demand.

We first look at what items are to be considered for predetermined choices when deciding whether or not to use fuelwood, i.e., $d_{1i} = 1$ indicating the residence decided to use wood energy as a heating source. Then, we follow with those variables, which influence the overall quantity of wood energy demanded, within the final outcome of equation (3.10). Finally, we look at those relevant econometric tests to determine validity of variables and overall model fit, as discussed within Heckman (1979), Vella (1998), and Van der Klaauw and Koning (2003).

Table 3.1. Description of Selectivity Variables, Justification, and Predicted Signs.

Variable (z_{ji})	Description	Predicted Sign Effect on d_{1i}	Explanation
NWP	Composite measure of the averaged price of non-wood energy per million BTUs.	+	An increase in the price of non-wood energy, the probability of using alternatives, such as wood, increases.
INC, INC^2	Gross income of household; squared value of gross income, in thousands of dollars.	–	An increase in the household's gross income leads to an decrease in probability of wood usage.
HDD30YR	30 year average heating degree day: average of sum of total below base 65˚F.	+	Take into consideration average climate temperature effects on energy choice.
HOMEAGE	De-trended series for age of home; HOMEAGE = 2009 – YEARMADE.	+	A homeowner is thought to view the overall age of their home as an input for energy use type. For instance, younger homes likely use modern or updated efficient appliances have better weather insulation.
DIVISION	Dummy indicator for U.S. Census Division Omitted base: Divisions 1; Division l =		These regions are located in areas where warmer climate is known, and use of wood is thought to not be prevalent among residential households.
	1: New England (CT, MA, ME, NH, RI, VT)	base	Those regions in close proximity to areas where colder climate is known is thought to lead to an increase in probability of using fuelwood.
	2: Middle Atlantic (NJ, NY, PA)	+	
	3: East North Central (IL, IN, MI, OH, WI)	+	
	4: West North Central (IA, KS, MN, MO, ND, NE, SD)	+	Those regions in warmer climates likely don't have heating needs, thus lead to probability in lower wood use.
	5: South Atlantic (DC, DE, FL, GA, MD, NC, SC, VA, WV)	–	
	6: East South Central (AL, KY, MS, TN)	–	
	7: West South Central (AR, LA, OK, TX)	–	
	8: Mountain North Sub-Division (CO, ID, MT, UT, WY)	+	
	9: Mountain South Sub-Division (AZ, NM, NV)	–	
	10: Pacific (AK, CA, HI, OR, WA).	–	
OWN	Dummy indicator for home ownership; 1 = yes, head of household owns the home, 0 = no, otherwise.	+	Home ownership is thought to lead to an increase in the probability of using wood energy, as it provides both energy utility and aesthetic appeal. This is thought to occur within owned homes.
SINGLEFAM	Dummy indicator for type of residence; 1 = single family dwelling, 0 = other, including mobile and apartments.	+	Base case those residences residing within apartments or mobile homes, which probably don't utilize wood energy.
RURAL	Dummy indicator for rural classification; 1 = home is within a Rural area; 0 = home is within an Urban area.	+	A home located within an rural area probably uses locally sourced wood energy for heating needs, as it provides a means to self-sustainability. Also, urban utilities like natural gas and electricity may not be available.
HOUSE-HOLDER_RACE	Dummy indicator for Housholder's ethnicity; Omitted base: (7) – 2 or more selected; Race r =		The householder and family ethnicity may play a role in family traditions that relate to the use of fuelwood. Correcting for bias from ethnicity leading to selectivity.
	1: White Alone	+	
	2: Black or African/American Alone	+	
	3: American Indian or Alaska Native Alone	+	
	4: Asian Alone	+	
	5: Native Hawaiian or Other Pacific Islander	+	
	6: Some Other Race Alone	–	
	7: 2 or More Races Selected	–	

SOURCE: U.S. Energy Information Administration, Residential Energy Consumption Survey (2009).
NOTE: Please see Chapter 4 for details on these variables and derivations.

A residential home is typically faced with the decision to use a specific fuel or heating appliance, generally to maximize utility obtained from the heat (Dubin and McFadden, 1984; Quigley and Rubinfeld, 1989). The source providing both the specific applicable use and the highest utility are chosen over all other alternatives. This choice may or may not easily be able to be reversed after being made, which can create an implication to the user in a form of a sunk cost leading to continued use regardless of what market forces dictate in the short term. Here we look at those variables that influence the decisive factor of whether or not to use wood as a fuel. We assume the choice is a dichotomous decision; to use fuelwood or not. Table 3.1 details those variables, Z_{ji}, where *j* = 1, … , M, signifying the actual variable to be considered exogenous regressors of selectivity for fuelwood, $d_{1i} = 1$. This table details the variables, descriptions, predicted signs, and explanations of model inclusion. Since the data is of individual observations and essentially is a probit model, maximum likelihood estimation is conducted (Pindyck and Rubinfeld, 1998). Here we look to estimate those coefficients that impact the decision to consume wood energy, $\hat{\gamma}_j$. Keep in mind this corrects for the probability of choosing to consume fuelwood over not to consume, and interpretation of coefficients is to be taken carefully as it is difficult to determine the overall magnitude the regressor has on the probability of demanding wood energy. We use this estimation to correct for the error term bias found within a regular OLS estimation, as previously discussed and presented within equations (3.1) through (3.10). These are the variables which we use to estimate the selectivity correction term, $\frac{\sigma_{u\eta}}{\sigma_\eta^2}\hat{\lambda}_i$.

For the selectivity model, we observe the overall positive or negative effect the variable has on the probability of choosing fuelwood, indicated by USEWOOD = 1 within the 2009 RECS dataset. Non-wood energy prices (NWP), describe prices faced by the residence and derived from other variables within the RECS dataset. NWP is an averaged composite measure of non-wood alternative energy prices per million BTUs, similar to the method used by Song et al. (2012b). Within the 2009 RECS data, total energy usage in thousand BTUs (TOTALBTU) and total expenditures in whole dollars on energy (TOTALDOL) are provided for each observation. Both exclude wood or other biomass energy within the calculations. The U.S. Energy Information Administration doesn't include wood energy prices or expenditures elsewhere, either; only the estimated quantity consumed of fuelwood is available, our dependent variable within equation (3.1). To ease this analysis, we simply divide the TOTALDOL data by TOTALBTU, and then divide by 1,000 to get the observation's price of non-wood energy in million BTUs. We believe an increase in the composite price of non-wood fuel alternatives (NWP) leads to a probable increase in fuelwood usage, as users of traditional fuels may seek alternative sources when market prices increase.

Next, we control for geographical location, including where colder regions are prevalent through division indications of survey respondents with a mutually exclusive dummy (DIVISION). The probability of a residence to choose wood is believed to increase as well in specific regions, as indicated within Table 3.1. Furthermore, controlling for effects of rural locations and the 30 year average historical heating degree days 65°F, (RURAL) and (HDD30YR), respectively, we believe they too have positive effects on the probability to choose wood energy. These are viewed as means to correct for possible selectivity determinants based on available energy market utilities and home location, as well as historical average winter temperatures that lead to demanding a heat source.

We look further at housing characteristics to estimate influential effects on the chosen outcome. Home ownership (OWN) is considered within the model, as it possibly effects the amount of responsibility, risk, and convenience the homeowner is willing to take. A positive measure here indicates the owner is more likely to use wood, even though it requires more frequent service and maintenance, relatively speaking to other modern energy sources. Also, it may measure the fire risk associated with burning the wood, although that is subjective and beyond the scope of this study.

Other decision variables to consider are those of dwelling traits with age and type. Older homes, dating to the earlier 20th century, likely used wood as a heating means over natural gas, fuel oil, liquefied petroleum gas, and electricity. HOMEAGE is the de-trended age of home construction. Younger homes are viewed to use local, perhaps more modern, utility energy sources for heating needs, whereas older units are considered to have heating stoves and fireplaces from the early 20th century, which are traditionally found in the older homes of the era. In addition, this includes a means to involve efficiency of newer homes and wood appliances and their likely impact on wood energy use as well. As the variable HOMEAGE increases incrementally from zero, indicating 2009 as the year of construction, the probability of demanding wood energy increases.

Finally, three other demographic and housing characteristics are hypothesized to affect choice in fuelwood use are included. If the residence indicates it is an attached or detached single family home (SINGLEFAM), the probability of using wood is predicted to be positive. The gross household income (INC) is believed to yield a negative probability on choosing fuelwood. As incomes increase, the home may tend to use fossil fuels for convenience due to work-time factors. Also, the homeowner's ethnicity (HOUSEHOLDER_RACE), is believed to predict probable usage from family traditions for cooking and heating. All are believed to be positive relative to the base case of 2 or more ethnicities chosen, as the number of observations may be small relative to the overall sampled size.

3.3.2. Empirical Model: Quantity Demanded

Once the residence decides to use wood, a quantity, y_{1i} measured in million British thermal units (BTUs), is determined by a set of the same, similar, or other variables, X_{ki}, where $k = 1, \dots, N$ different variables. Table 3.2 details the variables, descriptions, predicted signs, and explanations of model inclusion. Here we look to estimate those coefficients that impact the quantity demanded, $\hat{\beta}_k$. Unlike the estimates of the selectivity correction model, we are able to interpret the signs and magnitudes in accordance with traditional OLS estimations. Lastly, there must be differences between included variables of both steps to avoid excessive multicollinearity (Heckman, 1979; Bourguignon, Fournier, and Gurgand, 2007; Couture, Garcia, and Reynaud, 2012; Mansur, Mendelsohn, and Morrison, 2008).

Table 3.2. Description of Demand Variables, Justification, and Predicted Signs.

Variable (x_{ki})	Description	Predicted Sign Effect on y_{1i}	Explanation
USAGE$_i$	Categorical: Describes fuelwood usage; (see Table 4.6 for details) Pellets vs Logs, Main vs Secondary heating space, Fireplace vs Heating Stove.	~	Measures the additional, or reduction, in the quantity of wood demanded based upon different uses for wood energy; pellets expected to use less, heating stove more, and main space more.
NWP	Composite average of non-wood energy prices (ng, fo, elec, kero, lpg, etc.).	+	Price of wood energy is not reported within the survey. Use of composite non-wood energy prices provide means to estimate substitutable energy. (Song, et al., 2012a, 2012b)
INC	Estimated income level, based upon mid-point of income categorical range within the RECS survey, in thousands of dollars.	+	Measures the additional, or reduction, in the quantity of wood demanded based upon increase or decrease in income; aids in determining if wood energy is to be considered a luxury or necessity item (based upon income classification/range).
DRAFT	Dummy, indicates the respondent feels home is mostly, or all the time drafty.	+	A drafty homes require more energy to heat space.
TOTHSQFT	Total square feet of heating space.	+	A larger home requires more energy to heat space.
HDD65	The 2009 heating degree days, sum of total below base 65° Fahrenheit.	+	Measures the percentage increase, or decrease, in wood energy required for heating needs as the outside temperature increases.
RURAL	Dummy, 1 = home is within a Rural area, 0 = home is within an Urban area.	+	Measures the increase, or decrease, in the quantity of wood demanded based upon the residence location within an urban or rural.
OWN	Dummy, home ownership, 1 = yes, 0 = no.	+	Measures the increase, or decrease, in the quantity of wood demanded based upon the residence is owned by head of household.
PRIMEDUC	Dummy, 1 = head of household has K-12 education or less; 0 = otherwise.	+	Measures the increase, or decrease, in the quantity of wood demanded based upon the head of household having an education greater than primary K-12.
HHAGE	Head of household age in years.	–	Measures for possible effects aging has on the consumption of wood energy; as the head of household ages, use of other "easier" heating methods leads to decrease in wood energy demanded.
HOMEAGE	Age of home; HOMEAGE = 2009 – YEARBUILT.	+	As the age of the home increases, the quantity of wood energy demanded decreases due to lesser installations, more efficient appliance, as well as other heating choices available.

The first demand variables discussed here are those that are also included within the choice model, as they are believed to also have a direct effect on the quantity of wood energy demanded as well as choice of use. Non-wood prices (NWP), size of space to be heated (TOTHSQFT), being located in a rural region (RURAL), geographical location (DIVISION), age of home construction (HOMEAGE) and home ownership (OWN) are hypothesized to have the same effects as the selectivity model.

Next, we investigate the general use of wood energy and its impact on the quantity demanded through the mutually exclusive categorical dummies, USAGE$_d$, summarized in Table 3.3. This indicates which appliance, household space as main or secondary space heating, and the type of wood used. Wood pellets and logs are categorized with appliance type and main or secondary heating spaces. Since the Heckman (1979) is a selectivity two-stage estimation procedure, observations indicating no use of wood energy are dropped in the second stage. The first step, however, includes all observations to calculate the probit estimated lambda coefficient. This, in turn, involves all observations within the second OLS step. The base category includes those observations that indicate other wood types, wood burning appliances, and uses.

Table 3.3. Frequency and Percent of Observations for Wood Energy Usage.

$USAGE_d$	Description	N Obs. 12,083	Percent of sample	Percent of wood users
ι = nowd	No wood used; dropped from step-two Heckman	10,649	88.13	N/A
pelmfp	Main fireplace uses pellets	3	0.02	0.21
pelmhs	Main heating stove uses pellets	36	0.30	2.51
pelsfp	Secondary fireplace uses pellets	9	0.07	0.63
pelshs	Secondary heating stove uses pellets	38	0.31	2.65
logmfp	Main fireplace uses wood logs	38	0.31	2.65
logmhs	Main heating stove uses wood logs	169	1.40	11.79
logsfp	Secondary fireplace uses wood logs	623	5.16	43.44
logshs	Secondary heating stove uses wood logs	213	1.76	14.85
oth	Base: Other wood types and other wood burning appliances	305	2.52	21.27

The effects of income (INC) are explored, as we anticipate those with lower incomes view wood energy as a necessity, whereas those with higher incomes use it as a luxurious, aesthetic object. We derive, similar to the method used by Song et al. (2012b), a generalized mid-point income measurement for each observation in thousands of dollars, since the EIA doesn't disclose actual household income amounts, only a series of 24 categorical ranges between less than $2,500 and greater than $120,000. The midpoints are found for each observation, rounding down, to attach a measured income level rather than a categorical factor. Newer homes are also viewed to have greater insulation and more efficient appliances (that use wood, if any), indicated by a dummy for those who feel the home has drafts (DRAFT), and also within the age of the home (HOMEAGE). Like the climate regressor of the selectivity model, the percent change in total heating degree days in 2009 for the residence (HDD65) looks at the current consumption rate for the current weather season.

Finally, we consider a two regressors describing homeowner characteristics. HHAGE is age of the head of the household in years and PRIMEDUC is a dummy that indicates the head of household has at least a K-12 primary education, zero indicates no education or didn't finish primary schooling. Age is predicted to be negative. As the homeowner ages they are possibly not going to be able to handle the self-maintenance of wood appliance ownership. Lastly, education of the homeowner is viewed to be positive, as the overall effects are thought to increase overall environmental and market pricing awareness on energy consumption. These variables condense to model (3.11):

$$Y_i = a_0 + \Sigma X_k \beta_{ki} + \frac{\sigma_{u\eta}}{\sigma_\eta^2} \hat{\lambda}_i + w_i, \tag{3.11}$$

which is our empirical model in question to be estimated, ultimately describing the expected individual demand for fuelwood when the decision is made in equation (3.12):

$$E(y_i | d_{1i} = 1) = a_0 + \Sigma X_k \beta_{ki} + \frac{\sigma_{u\eta}}{\sigma_\eta^2} \hat{\lambda}_i + w_i. \tag{3.12}$$

Our model is developed from a set of traits aimed at not only describing specifics of quantities demanded, but also those that influence the decision in the first place. Endogenous decisions, essentially external to influencing the quantity to be consumed, are incorporated to correct the potential bias that would otherwise occur within an OLS procedure.

3.4. Programming and Post-estimation Verification Tests

We apply our model through the econometric program found within Stata. This module allows us to apply the Heckman (1979) two-step estimation procedure for correcting sample selection bias, previously described within U.S. residential wood energy demand. As Heckman states, "given its simplicity and flexibility, the procedure outlined in this paper is recommended for exploratory empirical work" (Heckman, 1979). This allows us to determine if there is evidence of self-selection among residential wood energy demand within the United States. The procedure provides probit estimates for the selection criteria, the bias-adjusted OLS demand equation, as well as the computed lambda value (derived from the estimated rho and sigma values). It is this lambda value that is needed to correct the selectivity bias previously discussed.

Post-estimation tests are fairly straightforward for this method. Standard *t* tests apply to the coefficients. Conveniently, the software conducts an asymptotic relationship test among all of the included variables. The first step uses a Pseudo R^2 calculation, similar but different to traditional R^2 measures, to determine overall probit model fit. Furthermore, the second step uses a Wald χ^2 statistic with the lambda coefficient correction term to determine full fit of the entire model. We also compare results of the selectivity model against a traditional OLS and Tobit (Tobin, 1958) models, which doesn't incorporate regressors of the decision correction model.

3.5. Model Assumptions and Limitations

The Heckman (1979) two-step model has a single strong assumption, that the errors are independent and identically jointly distributed $N(0, \boldsymbol{\Sigma})$ (See (A.1) on page 38). This has widely been criticized and tested (Vella, 1998; Bourguinon et al., 2007). This research imposes this restriction and merely presents the method as a tool to indicate a suggestion for possible evidence of further analysis.

3.6. Summary

This study estimates demand for fuelwood within U.S. residential households, an energy source that is relatively small to other modern sources. We use the most recent data available, discussed further within the next chapter, which seeks to provide additional knowledge beyond current literature on estimates of price and income effects on demand. Furthermore, this data provides insight on updated information prior recent studies didn't utilize, as it wasn't available at the time (Song et al., 2012a, 2012b; Kaza, 2008, Hardie and Hassan, 1986, Mackenzie and Weaver, 1986). We found within the literature limited studies of selectivity correction for fuelwood demand estimation, which we've argued within this chapter is possible given the relative small demand apparent within available data in modern day developed economies, such as the United States.

This chapter looks at the methods behind the Heckman (1979) two-step selectivity correction procedure utilized in our model (see equation 3.11). The use of this method aids biomass firms and policy makers in correcting for residential selectivity of heating fuel. Thus, when correcting for selectivity, unbiased estimates are provided. As Heckman (1979) states, this method is a simple means to determine if evidence of selectivity exists. Further studies are suggested post results. The next chapter looks into the details behind the data, which should aid in further understanding of the method and model driving this study.

CHAPTER 4: ANALYSIS OF DATA AND VARIABLES

This section goes into detail on the data used within this study. In order to understand why we impose a selection model for demand estimation, it is important to understand what data is available to conduct a demand study on wood energy. We split the information about the data into three sections, beginning with detailed information on the 2009 EIA Residential Energy Consumption Survey. We use information from this section to determine which available variables are appropriate to use within the modeling, or if there needs to be derivations based upon specific survey questions. The second section summarizes statistical information about the survey leading to the observation's probable choice outcome, i.e., the first step selectivity correction model. This part is important as it estimates a term to correct for omitted variables caused from the sample selection error process (Heckman, 1979). Variables chosen specifically fall within results from prior studies and those from economic theory, as discussed.

Section three of this chapter focuses on the variables found within the demand estimation model. These variables directly influence the quantity demanded, once the observation's probable decision to choose fuelwood as an energy source is incorporated from the first step. As in the first step, we include variables from results found within prior studies and those considered determinants of energy demand. We summarize the data used within the estimation procedure and present assumptions on the data in the last section.

4.1. 2009 RECS Survey

The study uses a data set comprised of U.S. residential demand for energy consumption. It is conducted and maintained by the United States Energy Information Administration (U.S. Energy Information Administration), titled the 2009 Residential Energy Consumption Survey (RECS). It contains information regarding residential consumption for the year 2009, surveyed across 16 states:

Arizona, California, Colorado, Florida, Georgia, Illinois, Massachusetts, Michigan, Missouri, New Jersey, New York, Pennsylvania, Tennessee, Texas, Virginia, and Wisconsin. These states are grouped within four regions as defined by the U.S. Census Bureau: Northeast, South, Midwest, and West (U.S. Census Bureau, n.d.[a]). These regions are further divided into ten divisions, also defined by the U.S. Census Bureau: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, and Pacific, with the Mountain region divided further into North and South subdivisions giving a total of 10 divisions (U.S. Census Bureau, n.d.[b]). This survey has six forms, detailed in Table 4.1, with 12,083 respondents representing the 113.6 million U.S. residential households as a whole. This survey was administered completely through electronic means by Computer Assisted Personal Interview (CAPI) (U.S. Energy Information Administration, 2009a).

Table 4.1. 2009 U.S. EIA Residential Energy Consumption Survey Forms.

Form	Title
457-A	2009 Residential Energy Consumption Survey, Household Questionnaire
457-C	2009 Residential Energy Consumption Survey, Nationwide Survey on Household Energy Use, Rental Agents, Landlords, and Apartment Managers Questionnaire
457-D	2009 RECS Propane (Bottled Gas or LPG) Usage Form
457-E	2009 RECS Electricity Usage Form
457-F	2009 RECS Natural Gas Usage Form
457-G	2009 RECS Fuel Oil and Kerosene Usage Form

Source: U.S. Energy Information Administration. 2009 Residential Energy Consumption Survey, Survey Forms. (2009b).

The 2009 survey year, and resulting data set, is chosen for this study because it is the latest, most up to date U.S. Energy Information Administration RECS release available. It is the largest REC survey to be conducted by the U.S. Energy Information Administration, with over 2.75 times the amount of respondents versus the prior 2005 survey. The first RECS was conducted in 1978 in order to statistically estimate and represent what the American household demands for energy usage (U.S. Energy Information Administration, n.d.).

The survey data set contains 931 columns of information divided into 14 sections: “Housing Characteristics”, ”Kitchen Appliances”, “Home Appliances and Electronics”, “Space Heating”, “Water Heating”, “Air Conditioning”, “Miscellaneous”, “Fuels Used”, “Housing Unit Measurements”, “Fuel Bills”, “Residential Transportation”, “Household Characteristics”, “Energy Assistance”, and “Scanning of Fuel Bills”, each labeled A through N, respectively. Survey questions involving heating are those that contribute to the decision to use wood and the quantity of wood energy used in this study, such as relevant heating questions, locations, incomes, etc. While the survey contains a lot of information within those sections, specific ones such as water heating fueled by wood, and space heating fueled by wood, are those that are of particular relevance to this study on wood energy demand. Other questions of relevance for this study are those that are potential determinants for the residence choosing wood energy, such as the amount of space heating the home has, household income, and geographical location of the residence.

The resulting survey creates a matrix of 11,249,273 unique data points that you can download at over 41,500 KB in size from the U.S. Department of Energy’s website. Since the survey data is representative of a cross section, not a time series, the values are assumed to be in real values without inflationary, transportation, or geographical adjustments required, and prices are further assumed to represent market clearing equilibrium prices. Unfortunately, no price information on wood energy usage is included, which is noted to not be reflected within variables TOTALBTU and TOTALDOL, which is discussed further in the sections below. The EIA’s weighted estimates approximate 892,671 households using pellets with an estimated 13,107,830 households using all types of fuelwood.

<u>4.2. Choice Model Variables from 2009 RECS</u>

We seek to identify variables for the choice model that can influence a residential household’s end decision in using wood as an energy source, as indicated by the model’s first step dependent variable USEWOOD within the 2009 RECS data set. Only 1,434 out of the 12,083, or 11.87% of total respondents, indicate fuelwood usage during 2009. Table 4.2 summarizes these variables with their descriptions and statistics given by the EIA RECS public codebook. The derived variables for the estimation procedure, as summarized by Table 4.3, are further examined within Table 3.1 of Chapter 3 on the methodology.

Table 4.2. RECS Variables for First Step of Heckman Choice Procedure: Choice.

Variable	Value	Description	Mean (Std. dev.)
USEWOOD	Dummy: 0 = No, 1 = Yes.	Wood is used in home.	0.1187 (0.32)
UR	Dummy: 0 = Urban, 1 = Rural.	Housing unit classified as urban or rural by Census	0.2207 (0.40)
TOTALDOL	Dollars	Total cost, in whole dollars, 2009 for energy expenditures.	$2,036 ($1,174)
TOTALBTU	Thousand BTU	Total usage, in thousand BTU, 2009 for energy consumption.	89,996 (54,469)
YEARMADE	Years	Year housing unit was built.	1971.06 (24.82)
HDD30YR	Heating degree days	Heating degree days, 30-year average 1981-2010, base 65°F.	4,141 (2,318)
KOWNRENT	1 = Owned by someone in household, 2 = Rented, 3 = Occupied without payment of rent.	Housing unit is owned, rented, or occupied without payment of rent.	1.34 (0.50)
TYPEHUQ	1 = Mobile Home, 2 = Single-Family Detached, 3 = Single-Family Attached, 4 = Apartment in Building with 2-4 Units, 5 = Apartment in Building with 5+ Units.	Type of housing unit.	2.66 (1.20)
DIVISION	1 = New England Census Division (CT, MA, ME, NH, RI, VT), base 2 = Middle Atlantic Census Division (NJ, NY, PA) 3 = East North Central Census Division (IL, IN, MI, OH, WI) 4 = West North Central Census Division (IA, KS, MN, MO, ND, NE, SD) 5 = South Atlantic Census Division (DC, DE, FL, GA, MD, NC, SC, VA, WV) 6 = East South Central Census Division (AL, KY, MS, TN) 7 = West South Central Census Division (AR, LA, OK, TX) 8 = Mountain North Sub-Division (CO, ID, MT, UT, WY) 9 = Mountain South Sub-Division (AZ, NM, NV) 10 = Pacific Census Division (AK, CA, HI, OR, WA).	Census Division, 1-10	5.37 (2.86)
HOUSE-HOLDER_RACE	Mutually exclusive dummies; 1 = White Alone, 2 = Black or African American Alone, 3 = American Indian/Alaskan Alone, 4 = Asian Alone, 5 = Hawaiian/Pacific Islander Alone, 6 = Other Alone, 7 = 2 or More Races	Head of household's ethnicity.	1.4422 (1.1415)

(continues)

Table 4.2. RECS Variables for First Step of Heckman Choice Procedure: Choice (continued).

Variable	Value	Description	Mean (Std. dev.)
MONEYPY	1 =Less than $2,500 2 =$2,500 to $4,999 3 =$5,000 to $7,499 4 =$7,500 to $9,999 5 =$10,000 to $14,999 6 =$15,000 to $19,999 7 =$20,000 to $24,999 8 =$25,000 to $29,999 9 =$30,000 to $34,999 10 =$35,000 to $39,999 11 =$40,000 to $44,999 12 =$45,000 to $49,999 13 =$50,000 to $54,999 14 =$55,000 to $59,999 15 =$60,000 to $64,999 16 =$65,000 to $69,999 17 =$70,000 to $74,999 18 =$75,000 to $79,999 19 =$80,000 to $84,999 20 =$85,000 to $89,999 21 =$90,000 to $94,999 22 =$95,000 to $99,999 23 =$100,000 to $119,999 24 =$120,000 or More.	2009 gross household income.	13.03 (6.80)

Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey, public codebook.

Recall equation (3.7) from Chapter 3. This equation represents a vector of decision variables leading a residence to choose wood fuel over alternatives. We look to explain the dependent USEWOOD occurring through Heckman's use of probit estimation. To do so, there exists a set of regressors z_{ji}, which can be integer, continuous, or discrete in nature and describe d_{1i} as the choice outcome for wood use. The final conclusion and estimated coefficients, $\hat{\gamma}_j$, determine the probability of a residence choosing wood as an energy source. These regressors are found within Table 4.2.

To measure residential urbanization, variable UR indicates a dummy value of zero, or one for being rural. Furthermore, this variable is labeled, RURAL, within Tables 4.3 and 4.7. Similarly, geographical location across the U.S. is indicated by categorical dummies, DIVISION. The New England region is our base measure, which is omitted to prevent perfect collinearity. There are a total of ten U.S. Census divisions within this data set with details of each outcome identified within Table 4.2 and used within Table 4.3.

Table 4.3. First Step Selection Model Variables: Descriptive Summary and Statistics.

First step selection model variable descriptive summary and statistics

Variable		Value	Description	Mean (Std. dev.)
USEWOOD		Dummy, base = 0:	Binomial indicator of choice; when WOOD = 1, the user indicates they choose to use wood energy, WOOD = 0 otherwise.	0.1187 (0.3234)
	0 =	No,		
	1 =	Yes.		
NWP		U.S. $ per million BTUs.	Composite non-wood price.	24.94 (9.5667)
INC		Thousands of US dollars.	Obs. Income level.	54.87 (36.4691)
INC^2			Square of Obs. Income level.	4340.66 (4811.03)
DIVISION	1 =	New England Census Division (CT, MA, ME, NH, RI, VT)	Mutually exclusive discrete categorical dummies, base 1.	0.0776 (0.2676)
	2 =	Middle Atlantic Census Division (NJ, NY, PA)		0.1099 (0.3128)
	3 =	East North Central Census Division (IL, IN, MI, OH, WI)		0.0952 (0.2935)
	4 =	West North Central Census Division (IA, KS, MN, MO, ND, NE, SD)		0.1401 (0.3471)
	5 =	South Atlantic Census Division (DC, DE, FL, GA, MD, NC, SC, VA, WV)		0.1859 (0.3890)
	6 =	East South Central Census Division (AL, KY, MS, TN)		0.0508 (0.2196)
	7 =	West South Central Census Division (AR, LA, OK, TX)		0.1018 (0.3024)
	8 =	Mountain North Sub-Division (CO, ID, MT, UT, WY)		0.0368 (0.1883)
	9 =	Mountain South Sub-Division (AZ, NM, NV)		0.0304 (0.1716)
	10 =	Pacific Census Division (AK, CA, HI, OR, WA)		0.1715 (0.3769)
RURAL		Dummy, base 0:	Dummy, indicating Urban or Rural Census classification.	0.2009 (0.4007)
	0 =	Urban,		
	1 =	Rural.		
OWN	0 =	Otherwise,		0.6736 (0.4689)
	1 =	Householder owns home.		
HOMEAGE		Years.	Age of home in years.	37.9376 (24.8179)
HDD30YR			Heating degree days, 30-year average 1981-2010, base 65°F.	4135.15 (2260.54)
SINGLEFAM	0 =	Otherwise,	Type of housing unit.	0.7194 (0.4493)
	1 =	Single-family home, either attached or detached.		
RACE	1 =	White Alone	Estimate whether ethnicity impact the probability of using fuelwood.	1.4422 (1.1415)
	2 =	Black or African American Alone		
	3 =	American Indian/Alaskan Alone		
	4 =	Asian Alone		
	5 =	Hawaiian/Pacific Islander Alone		
	6 =	Other Alone		
	7 =	2 or More Races		

Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey, public codebook.

Certain demographics about the home and those living within the residence can influence a decision to use wood. The gross household income for 2009, MONEYPY, is included. Measured as a categorical variable, it provides a range of estimated respondent income. We modify this measure through the same method of mean value measurement Song et al. (2012b) does for their 2005 RECS study. The RECS survey gives a categorical range of possible income values for each observation (see Table 4.2), which makes it difficult for us to estimate income effects on using wood over other energy sources. Thus, this variable provides interpretation problems. Using the mean of possible monetary values for each of the 24 income categories should prove, on average, a means to calculate income effects. We assume this becomes almost continuous as the sample size becomes large. This is identified as INC in Tables 4.3 and 4.6. In regards to possible homes with joint incomes, no measure is unavailable within the data. The inclusion of this would benefit as a means to determine whether there is someone available frequently to continually supply the fuelwood for heating needs, as working families tend to not have available spare time to maintain such labored work.

Other demographic measures involve home ownership, KOWNRENT, which gives three categories to indicate the home is owned by someone within, rented, or is occupied without payment of rent, and HOUSEHOLDER_RACE, which describes homeowner's ethnicity in seven different categories. To simplify the model, we modify home ownership type into a dummy to indicate those who own the home, or all other possible scenarios as the base noted as dummy OWN within Tables 4.3 and 4.7. Home ownership certainly can influence the residence's ability and willingness to use fuelwood due to risk involvement and home modifications for storage or available appliances. We also include ethnicity to determine whether ethnicity plays a role in originating choice to use fuelwood as some traditions may lead to the use due to specific cooking styles or societal concern for the environment.

The type of home construction, TYPEHUQ in Table 4.2, is also considered a deciding factor for wood use. This is a categorical variable describing homes identified as mobile, single-family detached, single-family attached, or apartments in two ranges (2-4 units, or 5+ units). We hypothesize that homes categorized as a single-family home will show an increasing probability of using fireplaces or heating stoves, meaning the likelihood of home maintenance and renovation decisions are up to the home owner of a single-family type rather than apartments or mobile homes. The other types are possibly influenced by available space to use wood energy, consider responsibility to not occur to the residing occupants, or because of safety and insurance reasons due to increased fire risk associated with using wood energy. Single-family homes are possibly able to control for those factors due to the shift in risk responsibility and available space to use the energy. As such, we simplify the estimation by focusing on observations identified as single-family attached or detached and create a dummy, where $SINGLEFAM_i = 1$ for when $TYPEHUQ_i = 2 \text{ or } 3$, or zero for all other observations, in Table 4.3.

The age of the home is considered a determinant of probable wood usage as well. In Table 4.2, YEARMADE indicates the year of construction the survey respondent believes the home was constructed. The RECS codebook indicates it ranges from 1600-2009, yet the actual range the respondents claim is between 1920 and 2009. We find here the average home is built around 1971, with a standard deviation of 24.8 years. To determine actual age of the home, we calculate this as a reversed de-trended value by using $HOMEAGE_i = 2009 - YEARMADE_i$, which aids in the probability interpretation in Table 4.3. As the age of the home increases, the probability of using wood energy increases. Homes built in the early 1900s likely used self-sustainable energies due to modern fossil fuel networks not being widely implemented across the nation. Likewise, the late 1970s saw an increase in wood energy usage due to the oil embargo from OPEC (Song et al., 2012a). Modern homes built in the late 1990s and 2000s likely use electricity, natural gas, or propane for heating needs.

Historical climatic effects are believed to increase the probability of using wood as well. As the location experiences, on average over the course of 30 years, lower temperatures and greater quantities of days required for heating, the use of localized fuelwood increases the probability of wood use. Table 4.2 describes HDD30YR as this 30 year average of heating degree days, with 65°F the base temperature for calculations. The National Weather Service defines heating degree days as the summation of daily average temperatures below the base over the course of the year (National Weather Servce, n.d.). This temperature is believed to be the point at which a location is to use energy to create heat for comfort. The same variable name is used within Table 4.3 for modeling choice.

Lastly, we look to incorporate price of energy into the model. Song et al. (2012b) and Arabatzis and Malesios (2011) find price of energy impacts the end use decision and quantity demanded. Since the RECS doesn't include wood energy prices, we impose an assumption imposed by Song et al., obtained through economic theory that quantities and prices are simultaneously determined and a reduced form model is applicable (Greene, 2003; Wooldridge, 2013). The RECS includes aggregate energy quantities and expenditures, which conveniently stated by the EIA, doesn't include wood energy. TOTALDOL and TOTALBTU are the total cost of energy and total quantity of energy in thousand BTU, respectively, within Table 4.2. On average, the observations use approximately 90 million BTUs of energy, and expend approximately $2,036 during the year. We further aggregate these values for each observation through equation (4.1) for Table 4.3,

$$NWP_i = TOTALDOL_i/(TOTALBTU_i\ /1{,}000), \tag{4.1}$$

which is the non-wood energy price observed in dollars per million BTUs. This describes the average observed value for the use of other fuel sources. Descriptive statistics on the non-wood fuel prices is found within Table 4.4, detailed as price per million BTUs. If the price for non-wood energy increases, on average, the probability of using wood energy increases. This value is also used within the second-step demand model, as detailed in the next section.

Table 4.4. Descriptive Statistics of Fuels, $ per mBTUs.

	Full Sample					Observations where $/mBTUs is greater than zero				
	NG	LP	FO	KERO	EL	NG	LP	FO	KERO	EL
# Obs.	12,083	12,083	12,083	12,083	12,083	7,494	1,075	904	166	12,081
Avg.	8.1324	2.3380	1.3148	0.3080	37.5607	13.1123	26.2791	17.5742	22.4209	37.5659
Min	0.0000	0.0000	0.0000	0.0000	0.0000	3.8322	10.9379	11.3177	7.4074	0.7879
Max	346.3415	88.2040	35.5989	37.0370	506.5502	346.3415	88.2040	35.5989	37.0370	506.5502
Std. dev.	11.1203	7.9934	4.6593	2.6879	13.2838	11.5797	9.4363	2.0981	5.4987	13.2753
Kurt.	289.2323	17.0373	9.4443	87.1573	164.0440	367.0403	8.5260	10.6096	1.4110	164.4365
Skew.	12.8339	3.8332	3.3315	9.1500	6.5542	17.1715	2.4091	1.9801	0.7730	6.5691

Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey.

4.3. Wood Demand Variables from 2009 RECS

The demand dependent variable, BTUWOOD, is the residential quantity of wood energy consumed in 2009. It measures the thousands of BTUs used for the entire year. For this study, each observation's wood consumption is first divided by 1,000, to adjust for interpretation purposes, for millions of BTUs as indicated within Table 4.7 under MBTUWOOD. The survey respondents indicate the use of wood energy by appliance, main or secondary space heating, type of wood used, and whether other appliances were available and/or not used. A list of variables in Table 4.4, FUELNOHEAT, FUELHEAT, FURNFUEL, RADFUEL, PIPEFUEL, RNGFUEL, DIFFUEL, FUELH2O, and FUELH2O2 aid in understanding how the respondent will use the chosen fuel. It's important to note here that many observations respond within the survey a categorical value of -2, meaning the question/response is not applicable to them. As the summary statistics in Table 4.5 hints, we see there is potential for bias within the dataset from a large response of this category. Even though many observations indicate this response, we convert the -2 responses to a zero for calculation purposes within the modeling. The observation still doesn't use wood even though it is a question of not being applicable. Since observations responding to these variables indicate the type of fuel used, we assume here a direct relationship to BTUWOOD when the respondent indicates category 7, in which wood is the energy source for the space or appliance.

Table 4.5. RECS Variables for Second Step of Heckman OLS Procedure: Demand.

Variable	Value	Description	Mean (Std. dev.)
BTUWOOD	Total wood usage, in thousand BTU, 2009.	Wood consumption is not included in TOTALBTU or TOTALDOL; Thousand BTU.	4,507.57 (20,387.98)
TOTHSQFT	Total heated square footage.	Square feet of space that is heated.	1,675.96 (1,200.82)
HDD65		Heating degree days, base 65°F, for year 2009.	4,141.38 (2,317.76)
HHAGE	Age; 16-95.	Age of head household.	49.74 (16.72)
EDUCATION	0 = No schooling completed, 1 = Kindergarten to grade 12, 2 = High school diploma or GED, 3 = Some college, no degree, 4 = Associate's degree, 5 = Bachelor's degree, 6 = Master's degree, 7 = Professional degree, 8 = Doctorate degree.	Highest education completed by head of household.	3.3916 (1.6837)
WDPELLET	0 = No, 1 = Yes, -2 = Not Applicable.	Wood pellets used.	-1.7539 (0.6768)
WOODLOGS	0 = No, 1 = Yes, -2 = Not Applicable.	Wood logs used.	-1.6592 (0.9359)
FUELNOHEAT,	1 = Natural Gas,	Fuel for unused space heating equipment,	-1.8663 (0.9144)
FUELHEAT,	2 = Propane/LPG,	Main space heating fuel,	2.2612 (2.2654)
FURNFUEL,	3 = Fuel Oil,	Fuel used by warm-air furnace for secondary heating,	-1.9001 (0.7000)
RADFUEL,	4 = Kerosene,	Fuel used by hot water system for secondary space heating,	-1.9759 (0.4283)
PIPEFUEL,	5 = Electricity,	Fuel used by pipeless furnace for secondary space heating,	-1.9920 (0.1956)
RNGFUEL,	7 = Wood,	Fuel used by cooking stove for secondary space heating,	-1.9617 (0.4848)
DIFFUEL,	8 = Solar,	Fuel used by other secondary space heating equipment,	-0.0673 (0.3869)
FUELH2O,	9 = District Steam,	Fuel used by main water heater,	2.7112 (2.0009)
and	21 = Other Fuel,	and	-1.8304 (1.0465)
FUELH2O2	-2 = Not Applicable.	Fuel used by secondary water heater.	
WOODKILN	0 = No, 1 = Yes, -2 = Not Applicable.	Heating stove used for secondary space heating.	-0.0494 (0.4119)
HSFUEL	7 = Wood, 21 = Other Fuel, -2 = Not Applicable.	Fuel used by heating stove for secondary space heating.	-1.7710 (1.4906)
FPFUEL	1 = Natural Gas, 2 = Propane/LPG, 7 = Wood, 21 = Other Fuel, -2 = Not Applicable.	Fuel used by fireplace for secondary space heating.	-1.1943 (2.6289)
DRAFTY	1 = All the time, 2 = Most of the time, 3 = Some of the time, 4 = Never.	Objective view of respondent: Is home too drafty in the winter?	3.3077 (0.9033)
MONEYPY	see Table 4.2	see Table 4.2	
TOTALDOL	see Table 4.2	see Table 4.2	
TOTALBTU	see Table 4.2	see Table 4.2	
UR	see Table 4.2	see Table 4.2	
KOWNRENT	see Table 4.2	see Table 4.2	
YEARMADE	see Table 4.2	see Table 4.2	

Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey, public codebook.
NOTE: Mean and Standard deviation statistics for appliance and wood type terms under each option are within Appendix A.

We further simplify the nine indicators, along with wood types WDPELLET and WOODLOGS, by deriving mutually exclusive dummy/categorical indicators of using wood pellets or wood logs with fireplaces or heating stoves detailed in Table 4.6. It is important here to note that quantities of each are not reported, but the effects each have on overall quantity demanded through the end result of demanding MBTUWOOD can be inferred. While both are likely to have positive signs, indicating a shift in demand from the base scenario, it is the magnitudes that are of particular interest. With the possible growing interest in pellets as a fuel source, we hypothesize WDPELLET to have a larger magnitude than WOODLOGS (U.S. Energy Information Administration, 2014d).

The regressor WDPELLET found in Table 4.5 is of particular interest, as recent production and exportation of wood pellets increased greatly (Peng et al., 2010; Spelter and Toth, 2009). Using the prior stated mutually exclusive dummies to indicate observations, an individual using wood pellets only aids in capturing consumption effects associated with the use of the fuel source. Same scenario exists for wood log usage, the type of appliance, and the frequency of use as indicated within main or secondary space heating as well. Main space heating appliances are hypothesized to have larger magnitude measures relative to secondary space heating appliances. As the summary finds the majority of observations use wood within secondary heating spaces, we hypothesize secondary to have greater magnitude and significance relative to main heating spaces as well, although recent trends find newer and larger homes to not be consuming as much heat energy as they used to (U.S. Energy Information Administration, 2012d, 2013a). These variables are used within the final model, as detailed further in the prior chapter on the methodology in Table 3.3.

Table 4.6. Mutually Exclusive Wood Usage* Options: Data Summary.

Variable	Description	Count	Obs. 12,083 Mean (Std. dev.)		Obs, 1,434 Mean (Std. dev.)	
NOWD	No wood usage by observer; omitted.	10,649	0.8813	(0.3234)	–	–
PELMFP	Main space heating is fireplace using pellets.	3	0.0002	(0.0158)	0.0021	(0.0457)
PELMHS	Main space heating is heating stove using pellets.	36	0.0030	(0.0545)	0.0251	(0.1565)
PELSFP	Secondary space heating is fireplace using pellets.	9	0.0007	(0.0273)	0.0063	(0.0790)
PELSHS	Secondary space heating is heating stove using pellets.	38	0.0031	(0.0560)	0.0265	(0.1607)
LOGMFP	Main space heating is fireplace using logs.	38	0.0031	(0.0560)	0.0265	(0.1607)
LOGMHS	Main space heating is heating stove using logs.	169	0.0140	(0.1174)	0.1179	(0.3225)
LOGSFP	Secondary space heating is fireplace using logs.	623	0.0516	(0.2211)	0.4344	(0.4959)
LOGSHS	Secondary space heating is heating stove using logs	213	0.0176	(0.1316)	0.1485	(0.3558)
OTH	Base: Other uses, wood types, appliances, and/or multiple uses, wood types, and appliances.	305	0.0252	(0.1569)	0.2127	(0.4094)

* See Appendix A for further details regarding usage factors.

Moving onto demographics of the householder, we hypothesize as the homeowner ages (HHAGE) the quantity of wood usage increases as comfort for heat is preferred during colder conditions. We also view the same hypothesized outcome with educational background. As the individual's education increases, the quantity of wood use begins to increase due to having knowledge about specific factors associated with benefits from its energy usage, e.g., climate change, self-sustainability, etc. To simplify estimation and interpretation, we factor in a dummy indicated by PRIMEDUC, which indicates the individual has a K-12 primary education or less. It is important to note that almost all observations indicate education of K-12 or higher (98.34%). Having a dummy variable with this high of occurrence doesn't give enough variability for estimation significance nor meaning. Thus, we include only observations that have a minimal social-norm education, or less, without progressing onto further skill development, which 34.97% of observations have. We look for PRIMEDUC to have a negative shift in demand.

Specific factors of market indicators, non-wood price (NWP) and household income (INC) are included in the second-step estimation as well. Arabatzis and Malesios (2011) and Song et al. (2012b) find household income levels to significantly affect consumption levels, whereas Song et al., (2012b) finds prices of wood-alternative fuels to be significant. What these two variables provide is an opportunity to estimate income and price effects on wood energy consumption. What we expect to see is a parabola type effect as the income class goes up, i.e., lower income ranges will use more as wood can be sourced for little to no cost, and higher income ranges will use wood more as a luxury good, like ambiance and mood settings. Thus, we predict the middle income ranges will likely be located within regions where easy access to heating systems exist, and due to the working middleclass stereotype, we believe those households look to methods of energy where little work is required on a majority basis.

Another factor to consider is an observational objective viewpoint for how drafty the home is. It isn't uncommon to have survey information that involves objective views regarding specific situations, for example Arabatzis and Malesios (2011) use residential views of climate change within their survey estimations. The homeowner's view on how drafty the home is during the winter is hypothesized to increase the quantity of wood energy used during the heating season. A drafty home will require more heat energy to maintain comfort temperatures vs. a home that is drafty of a lesser magnitude. We simplify this as well into a single dummy variable, DRAFT in Table 4.7, where observations indicating it is drafty all the time (1) or most of the time (2) a value of one, and zero for all other instances.

Table 4.7. Second Step Selection Model Variables: Descriptive Summary and Statistics.

Variable	Value	Description	Mean (Std. dev.)
MBTUWOOD	Million BTUs.	Quantity of fuelwood consumed by the residence in 2009.	4.5076 (20.3880)
Usage Variables:			*(**)
NOWD,	0 = Dropped: No wood usage by observer,	Type of appliance and wood used for	0.8813 (0.3234)
PELMFP,	1 = Main space heating is fireplace using pellets,	heating for the indicated location.	0.0003 (0.0158)
PELMHS,	2 = Main space heating is heating stove using pellets,	Main or secondary space heating, wood	0.0030 (0.0545)
PELSFP,	3 = Secondary space heating is fireplace using pellets,	pellets or logs, fireplace or	0.0007 (0.0273)
PELSHS,	4 = Secondary space heating is heating stove using pellets,	heating stove are in question as they are	0.0031 (0.0600)
LOGMFP,	5 = Main space heating is fireplace using logs,	the most observed scenarios	0.0031 (0.0600)
LOGMHS,	6 = Main space heating is heating stove using logs,	(see Appendix A). Observations indicating	0.0140 (0.1174)
LOGSFP,	7 = Secondary space heating is fireplace using logs,	NOWD are	0.0516 (0.2211)
LOGSHS,	8 = Secondary space heating is heating stove using logs,	dropped in the second stage of the	0.0176 (0.1316)
OTH.	9 = Base: Other uses, wood types, appliances, and/or multiple uses, wood types, and appliances.	Heckman (1979) procedure. Other wood types, appliances, spaces, and/or multiple combinations are the base scenario.	0.0252 (0.1569)
NWP	Dollars per million BTU.	Composite price of non-wood energy (see equation 4.1).	***
INC	Thousands of dollars.	Observation income level in thousands of U.S. dollars.	***
TOTHSQFT	Square feet.	Total heated square footage.	1675.96 (1200.82)
HDD65	Sum of 2009 heating degree days.	Heating degree days, base 65°F, for year 2009.	4141.38 (2317.76)
HOMEAGE	Years.	Age of home.	***
DRAFT	Dummy, base 0: 0 = Not drafty, 1 = Drafty.	Observation of how drafty home is, where 1 = Always drafty or Most of the time drafty, 0 = otherwise.	0.1534 (0.3603)
RURAL	Dummy, base 0: 0 = Urban, 1 = Rural.	Dummy, indicating Urban or Rural.	***
OWN	Dummy, base 0: 0 = other, 1 = Householder owns home.	Home ownership indicator.	***
PRIMEDUC	Dummy, base 0: 0 = other, 1 = Householder has a K-12 primary education or less.	The householder has a primary K-12 education or less.	0.3664 (0.4818)
HHAGE	Age; 16-95.	Age of head household.	49.7414 (16.7221)

Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey, public codebook.
*,**: Mean and (Standard deviation) statistics, noted respectively, for data as a whole. Individual response statistics are within Appendix A.
***: See Table 4.3 on descriptive statistics.

The next variable in Table 4.5, UR, is a dummy accounting for whether the residence is located within rural or urban settings. A response from the residence of 1 indicates rural, and 0 indicates urban, which is classified by the Census. If the residence is located within a rural area, they are likely not connected to urban utilities like natural gas, district heating, phone, and internet. Thus, they are likely to use more locally sourced resources, such as wood and wood waste. Here, we expect the model to have a positive value for residence based within rural areas. Further, we label RURAL within Table 4.7 and interact with DIVISION to determine regional effects and different rural geographical locations effects on consumption. DIVISION locations are included to determine base shifts within demand based upon location overall.

Accounting for age of the home is another factor that can determine fuel usage and type of fuel. Within the 2009 RECS survey, respondents indicated the year their home was built under column YEARMADE. Here, the range indicated within the survey states 1600 to 2009 for the year the housing unit was built. What we hypothesize older homes lead to an increase in probability of using wood burning appliances. We include this as HOMEAGE within Table 4.7.

Finally, the 2009 heating degree days (HDD65) and home's heating area (TOTHSQFT) are included to determine variable and seasonal heating requirements, and constant space size heating requirements. Just like the 30-year average heating degree day measure, which is used to approximate average climate effects on heating needs, the yearly variability is used in the demand equation given appropriate seasonality in temperatures. Likewise, a home's heated space doesn't necessarily change every year. The stated total heated space is used to estimate a home's size as a determinant for heating demands.

4.4. Data Summary and Assumptions

There is a wide range of diverse residential home profiles and traits that can influence the quantity and type of energy used to heat the home. The variables chosen within the 2009 RECS will allow an analysis on wood energy consumption when adjusting for selectivity. Types of variables to include involve residential demographic, geographical location, temperature and climate, as well as housing traits such as size and age can influence consumer decisions and demand. To give a better picture of how the data and selectivity can be summarized, Table 4.8 summarizes the RECS data used in one source. We also summarize usage statistics within Appendix A. This gives us a comparison of overall RECS variable statistics chosen.

Table 4.8. 2009 RECS Descriptive Summary Statistics.

Variable		Mean	Standard dev.	Skewness	Kurtosis
BTUWOOD		4507.5730	20387.9800	9.1481	140.0900
USEWOOD		0.1187	0.3200	2.3581	6.5607
TOTHSQFT		1675.9560	1200.8190	1.9152	10.1102
HDD65		4141.3750	2317.7590	0.0790	2.1445
HDD30YR		4135.1470	2260.5440	0.0517	2.1263
YEARMADE		1971.0600	24.8200	0.5394	2.4196
HHAGE		49.7414	16.7221	0.2523	2.2615
EDUCATION		3.3915	1.6837	0.3887	2.4852
	1	0.0855	0.2796	2.9649	9.7905
	2	0.2643	0.4410	1.0693	2.1434
	3	0.2235	0.4166	1.3272	2.7614
	4	0.0987	0.2983	2.6903	8.2378
	5	0.2009	0.4007	1.4926	3.2280
	6	0.0792	0.2701	3.1164	10.7119
	7	0.0183	0.1340	7.1898	52.6928
	8	0.0130	0.1133	8.6009	74.9750
WDPELLET		0.0088	0.0933	10.5356	111.9994
WOODLOGS		0.1035	0.3046	2.6042	7.7818
DRAFTY		3.3077	0.9033	-1.2471	3.6929
	1	0.0741	0.2619	3.2578	11.5806
	2	0.0792	0.2701	3.1164	10.7119
	3	0.3117	0.4632	0.8132	1.6613
	4	0.5350	0.4988	-0.1405	1.0198
MONEYPY		13.0274	6.8041	0.2523	1.9018
TOTALDOL		2036.5540	1174.8030	3.0953	45.2945
TOTALBTU		89995.7000	54468.7600	2.0467	18.1657
UR		0.2009	0.4007	1.4933	3.2299
KOWNRENT		1.3381	0.4973	0.9689	2.5975
	1	0.6737	0.4689	-0.7408	1.5488
	2	0.3146	0.4644	0.7987	1.6378
	3	0.0118	0.1078	9.0611	83.1034
HOUSE-HOLDER_RACE		1.4422	1.1415	3.2536	13.6710
	1	0.7927	0.4054	-1.4440	3.0850
	2	0.1255	0.3313	2.2602	6.1086
	3	0.0091	0.0950	10.3371	107.8546
	4	0.0378	0.1908	4.8455	24.4791
	5	0.0033	0.0574	17.2934	300.0783
	6	0.0175	0.1310	7.3677	55.2832
	7	0.0141	0.1178	8.2517	69.0907
TYPEHUQ		2.6596	1.1926	1.0834	2.7117
	1	0.0448	0.2068	4.4024	20.38.144
	2	0.6458	0.4783	-0.6096	1.3716
	3	0.0737	0.2612	3.2643	11.6559
	4	0.0766	0.2660	3.1830	11.1316
	5	0.1591	0.3658	1.8635	4.4727
DIVISION		5.3700	2.8600	0.3080	1.9819
	1	0.0776	0.2676	3.1569	10.9658
	2	0.1099	0.3128	2.4944	7.2221
	3	0.0952	0.2935	2.7590	8.6121
	4	0.1401	0.3471	2.0736	5.3000
	5	0.1859	0.3890	1.6150	3.6081
	6	0.0508	0.2196	4.0906	17.7327
	7	0.1018	0.3024	2.6338	7.9370
	8	0.0368	0.1883	4.9184	25.1911
	9	0.0304	0.1716	5.4731	30.9550
	10	0.1715	0.3769	1.7431	4.0385

Source: U.S. Energy Information Administration, 2009 Residential Energy Consumption Survey.

While this data is a good fit for this research, we do make one assumption on the data: observations are evenly weighted. There is a clear need to weigh each observation, given it is a survey and proper representation of the population will consider a weighting system. The 2009 RECS includes a weight for each observation, however, the Heckman (1979) procedure doesn't include nor indicates a means to involve observational weights. Furthermore, the estimation procedure within Stata indicates a weight isn't allowed for the two-step procedure. A correlation matrix of the RECS data used within the model is available in Appendix B, which aids in viewing relationships among each other.

Lastly, there are a few limitations of this data. Due to survey design through itemization of fuel usage, it is difficult to determine exactly how fuelwood is used, e.g., if it is used as a backup source. This is a problem as many homes have or looking to utilize multiple heating fuel systems. Being able to test if fuelwood used as a backup source significant would be beneficial as a means to test the insignificance found by Couture et al (2012).

CHAPTER 5: RESULTS

5.1. Summary of the Study

The purpose of this study is to estimate U.S. residential fuelwood demand. Results provide information and insight for national and regional biomass energy firms, including those producing fuelwood-based products for home heating needs, as well as academic institutes, policy makers, and agencies concerned with biomass resources. Given the thin market share fuelwood observes and prior literature on selectivity, a Heckman (1979) two-step correction model for selectivity bias is used. The lambda coefficient estimated within the procedure provides a means to involve a decision model that doesn't necessarily impact quantity of fuelwood demanded directly. However, through residential selectivity it does provide support for estimating demand while correcting to choose to demand fuelwood over other options.

As a review, our research questions focus on the effects of non-wood price and gross household income have on wood energy consumption. These two questions should help firms anticipate estimation and supply issues when market economic conditions become volatile. They should also aid in determining who their targeted customers are by identifying if fuelwood is a normal or luxury good. We also explore the impact of climate and geographical location on demand, and whether we can use the outcome to forecast demand.

To estimate these questions, we view residential demand for fuelwood as an original observation when a consumer chooses to use the heating source. Identified by Heckman (1976, 1979), the decision outcome acts as an omitted variable in demand estimation, thereby creating bias results if not corrected. We hypothesized when correcting for such selection bias this decision is influenced from specific energy economic traits, such as efficiency of the appliance to be used with the fuelwood, availability of fuelwood source, convenience factors for maintenance, and usage over time and periodicity, as well as consumer and home demographics. To further review, as non-wood energy prices increase we hypothesized the probability of using fuelwood increases as well as the quantity of fuelwood demanded; colder climates are believed to influence the probability outcome of using fuelwood in conjunction with the quantity demanded; and fuelwood is hypothesized to be a normal good, i.e., as income increases the probability and quantity decrease as well. We estimate effects of such regressors within the modeling, covering both selectivity deterministic variables and general demand determinants influencing overall quantity to be consumed (see Chapters 3 and 4). These regressors are then a set of factors that influence overall utility to be gained from consumption.

5.2. Heckman Step-one: Selectivity Correction Results

The premise of this section is to focus on the results of the Heckman (1979) two-stage selectivity correction model we've estimated, first exploring the selection model and the estimated lambda coefficient correcting for omitted variable bias. As laid out within Chapter 3 on the methodology, this coefficient is an estimate for omitted variables of selection bias found within the demand procedures, where individuals or researchers create bias through sample selection, i.e., predetermined decisions to use something or making a choice instead of being randomly chosen, respectively. The overall effect the lambda coefficient has on an OLS model should correct for bias found within the model and decision process, thus producing consistent estimates. We then interpret the results of the selectivity bias corrected OLS model the two-stage Heckman (1979) procedure produces. The results are interpreted like a regular OLS estimation with *t* tests. The overall fit of the model and whether the variables chosen represent a form of relationship are tested and reported through the Wald test against a χ^2_{df} distribution.

Table 5.1. Results of Heckman First Stage Probit Correction.

Selectivity correction equation: USEWOOD			
Variable		Coefficient [a]	*t* value
Constant		-2.8786***	-14.72
NWP		0.0156***	8.49
INC		0.0072***	3.74
INC^2		-0.0000***	-2.60
DIVISION			
2		-0.3473***	-4.63
3		-0.2635***	-3.53
4		-0.1729**	-2.53
5		-0.0256	-0.32
6		-0.2973***	-2.90
7		-0.0946	-1.01
8		0.1210	1.31
9		0.1592	1.35
10		0.4295***	5.51
RURAL		0.4704***	12.34
OWN		0.2751***	5.80
HOMEAGE		0.0015**	2.21
HDD30YR		0.0001***	9.42
SINGLEFAM		0.5191***	9.78
RACE			
1		-0.1953	-1.57
2		-0.3166**	-2.35
3		-0.1610	-0.78
4		-0.9156***	-5.23
5		-0.7740**	-2.27
6		-0.2208	-1.23
mills			
lambda		14.7210***	3.05
rho		0.3535	
sigma		41.6419	
	Pseudo R^2 [b] =	0.1199	
	Prob > χ^2 =	0.0000	
	Log likelihood =	-3873.8315	
	Number of obs =	12083	
	LR $\chi^2_{[23]}$ =	1055.65	

Note: [a] (*, **, ***) denote statistical significance at the 10%, 5%, and 1% levels, respectively.
[b] Calculated by estimating a separate Probit model, as Stata doesn't report this within Heckman estimation.

The results of the selectivity model are reported in Table 5.1. Recall these results are estimates of a dichotomous choice outcome with an omitted variable correction term, where the residence uses fuelwood for heating needs through a probit regression as indicated by the RECS variable (USEWOOD). Here we see from the *t* values, in general, the variables chosen each have a fairly good fit. The overall fit, however, proves to be of a poor measure as the Pseudo R^2 measures just 0.1199 (Wooldridge, 2013; McFadden, 1974). This can be from numerous causes including inaccurate data collection, measurement error of the data, incomplete information underlying the variables chosen, or even incorrect chosen variables, theory, or model (Wooldridge, 2013). Most decision outcome variables are of some measureable significance, as discussed below, but the low Pseudo R^2 measure indicates there appears to be omitted variables within the first step (Wooldridge, 2013; Kennedy, 2008). Possible variables that may be missing include opinion based questions on the environment, policy, or future economic conditions like those included within prior research (Couture et al., 2012; Arabatzis and Malesios, 2011). As a reminder, since the first step estimation uses a probit regression the coefficient magnitudes are generally not interpreted due to different models providing different scales, though similar in nature (Wooldridge, 2013). For instance, removing one of the variables, home ownership (OWN), will yield a different model and a different scalable system for interpretation among all other variables. Thus, we interpret the signs of the coefficients rather than magnitudes in the probability of choosing to use fuelwood.

Regarding significance of the chosen variables, several are significant at the 1% level. Those include non-wood energy prices (NWP), household gross income (INC), income squared (INC^2), home ownership (OWN), average heating degree days (HDD30YR), single family home construction (SINGLEFAM), and being located within a rural region (RURAL) or divisions (DIVISION) two, three, six or ten where divisions are the Middle Atlantic, East North Central, East South Central and Pacific, respectively (division one, the base case, is New England). In addition, other variables found to be significant at a lesser extent are the age of home (HOMEAGE) and division four (West North Central), both of which are significant at the 5% level. In regards to the head of household's ethnicity, respondents indicating being of an Asian only decent, RACE-4, are found to be statistically significant at 1%, whereas respondents of Black-African American or Native Hawaiian/Pacific Islander decent, RACE-2 and RACE-5, respectively, are found to be statistically significant at 5% level.

Even though the coefficients are difficult to interpret, we interpret the signs of the coefficients, γ_j in equation (3.2), to determine if a change in base outcome positively or negatively affects the probability of predicting the dependent $d_{1i} = 1$, USEWOOD, indicating the observation selects wood as an energy source (Wooldridge, 2013). Of the variables found to significantly impact the decision, we estimate a positive change in non-wood fuel prices, income, age of the home, and average climate heating requirements will increase the probability of using wood for heating. An increase in income squared will yield a negative probability of using wood, but the magnitude is negligible relative to the other modeled variables.

Dummy variables are interpreted slightly different, as a change in status will yield a positive or negative impact on wood usage. We estimate homes located within rural areas are more likely to use wood than those in urban areas. This is similar to Mansur, Mendelsohn, and Morrison (2008) and their findings for fuel oil selection when natural gas is available to the residence. Home ownership and single-family dwellings are also found to likely use fuelwood over rental units and mobile/apartment complexes, respectively. Unlike the insignificance found within Abaratzis and Malesios (2011), this supports Couture et al. (2012) and their use of a dummy for indication of apartment dwellings and Mansur et al. (2008) across all discrete choice options. Our division dummies, being mutually exclusive of each other, indicate divisions two, three, four, and six to negatively impact the probability of fuelwood usage over the base case, and division ten to positively impact the probability of choosing fuelwood. Lastly, our dummies for ethnicity all negatively impact the probability of choosing fuelwood over our base case of some other race or multiple races being indicated.

Finally, we look at the estimated lambda coefficient. Interpretation of the coefficient isn't necessary, however, the significance is (Heckman, 1979). This value has an estimated *t* value of 3.05 with a probability of 0.002, well beyond our 1% critical value of 1.645 for a two-tail test, thus failing to reject the H_0: $\hat{\lambda}_i = 0$. While the selection regressors have a range of significance, we estimate there is evidence to correct for selectivity bias within this model given the dataset. This differs from

Hassan and Hardie's (1986) estimates of 1980-1981 fuelwood demand, where they found the lambda insignificant, and similar to Mackenzie and Weaver's (1986) lambda at 10% significance. Couture et al. (2012) used a multinomial logistic correction term, with similar findings within the natural gas and fuel oil equations, with fuel oil being the primary and fuelwood as a backup.

5.3. Heckman Step-two: Selectivity-Bias Corrected OLS Results

The interpretation of the overall modeled results are presented here, starting with Table 5.2. The first focus is on the estimation results of the second-step within the Heckman model. We apply the model and determine estimates of fuelwood demand by region and urban/rural locations. Finally, we apply elasticity estimates to the variables of income, non-wood prices, temperature, and heating space. Concluding remarks on the results and study succeed this section of the chapter.

Table 5.2. Results of Heckman Second Stage OLS Estimation vs OLS and Tobit Estimates.

Demand equation for selectivity corrected: MBTUWOOD						
	Heckman		OLS		Tobit	
Variable	Coefficient	t Value	Coefficient	t Value	Coefficient	t Value
Constant	-29.6137**	-2.06	-4.3659***	-5.32	-134.0085***	-21.54
NWP	0.6307***	4.38	0.0861***	5.22	0.4856***	4.62
INC	-0.0970**	-2.68	-0.0166***	-3.55	-0.0446	-1.45
TOTHSQFT	0.0037***	3.93	0.0009***	6.30	0.0053***	6.36
HDD65	0.0024***	3.85	0.0003***	4.04	0.0017***	3.47
HOMEAGE	0.2509***	5.06	0.0312***	5.15	0.1792***	4.34
PELMFP	35.5339	1.55	72.4441***	8.11	160.3632***	5.45
PELMHS	26.8060***	3.77	63.2955***	24.42	147.6086***	16.74
PELSFP	-14.9258	-1.11	13.4023***	2.60	110.4009***	6.45
PELSHS	-19.0483***	-2.76	20.2331***	8.02	103.2228***	12.03
LOGMFP	19.3761***	2.80	58.0866***	23.06	147.3189***	17.15
LOGMHS	38.8207***	9.60	83.5858***	68.29	166.7166***	35.81
LOGSFP	-9.8102***	-3.52	20.9618***	32.57	114.5150***	36.95
LOGSHS	-0.7800	-0.21	36.2255***	33.45	123.1870***	29.24
DRAFT	-3.3936	-1.08	-0.0007	-0.20	-0.0147	-0.12
RURAL	24.5277***	8.24	4.3132***	11.37	20.6749***	9.29
OWN	4.8597	1.31	0.3083	0.86	12.0701***	4.31
PRIMEDUC	8.1851***	3.34	0.7498**	2.38	0.4386	0.20
HHAGE	-0.1792***	-2.21	-0.0184**	-2.04	-0.1543**	-2.32
	Wald $\chi^2_{[18]}$ =	521.18	F(18, 12064) =	496.04	LR $\chi^2_{[18]}$ =	4959.84
	Prob > χ^2 =	0.0000	Prob > F =	0.0000	Prob > χ^2 =	0.0000
	Number of obs =	12,083	Number of obs =	12,083	Number of obs =	12,083
	Censored obs =	10649	R-squared =	0.4253	Pseudo R-squared =	0.2223
	Uncensored obs =	1434	Adj R-squared =	0.4245		
			Root MSE =	15.467		

Note: (*, **, ***) denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Observe the results of our overall choice corrected demand mode in Table 5.2, presented along with a traditional OLS procedure and a Tobit (1958) censored regression, both of which use the same second-step regressors minus the correction term. Variable coefficients and *t* values are reported, with significant variables labeled with asterisks. The Heckman (1979) procedure doesn't report an R^2 value, but does report the Wald statistic for joint significance of included variables. Our Wald χ^2 with 18 degrees of freedom measure is 521.18, well above the 1% critical value of 42.98.

This suggests a general good fit for the entire model. Nearly all variables are significant to some degree. Results provided are the marginal effects at the means, *ceteris paribus*. The three models indicate a difference in estimates, though significance of the coefficients is nearly the same. The OLS estimates, with nearly all variables significant, appear to increase effects on demand for use, type, and appliance use; all other variables tend to be estimated with lower effects on demand. A possible explanation for this is the larger observations demanding 80-600 million BTU of fuelwood can skew results in OLS as they are outliers. The Tobit estimates for USAGE appear over-estimated compared to both OLS and Heckman models' effects on demand. With total pellet users estimated to be nearly 900,000 households within the U.S., this model suggests five to eight cords equivalent of fuelwood burned per household during the heating season (2009)[6]. An explanation for this can result from the omission of non-wood using observations, yielding a generally higher estimate.

One of the primary variables of concern in this study, the composite measure for individual market prices of alternative energy (NWP), is both positive with a coefficient measure of 0.6307 and statistically significant at the 1% level. An increase of $10 in the composite price of non-wood energy alternatives estimates an increase of 6.307 million BTU of fuelwood demanded, or about 32% more cords[7]. The second primary variable of concern, income (INC), is statistically significant

[6] The 2009 RECS dataset approximates a weighted amount for population estimates, with 892,671.71 pellet using households out of the sample weighted 13,107,830.62 households using fuelwood.
[7] The U.S. Energy Information Administration uses a uniform conversion of 20 million BTU per cord (2009).

at the 5% level but negative in sign with a measure of -0.0970. An increase of $10,000 in income estimates a decrease of almost one million BTU of fuelwood demanded at the household level, or about 5% less per cord. Furthermore, estimates have larger magnitudes in comparison to an OLS model.

Other variables in question include heating space in square feet (TOTHSQFT), heating degree days below 65°F (HDD65), and the age of the home (HOMEAGE), all found to be significant at the 1% level. An increase of 1,000 sq ft in heating space estimates an increase of 3.7 million BTU in fuelwood, consistent with suggestions of heating requirements of larger homes (U.S. Energy Information Administration, 2013b). As the total heating degree days increase by a sum of 100 °F, keeping in mind that an increase in HDD reflects a colder winter, an estimated increase of 0.25 million BTU of fuelwood is found. Furthermore, as the age of the home increases by a decade, we estimate fuelwood demand to increase by 2.5 million BTU of those surveyed.

Moving onto the demographics and building characteristics, we find education and head of householder age to be statistically significant at the 1% level, with measured coefficients of 8.18 and -0.18, respectively. This suggests those with a K-12 education will use 8.18 million BTU of fuelwood more than other education levels, and as the head of household ages they will decrease the amount of fuelwood demanded by 0.18 million BTU a year.

We also find home ownership to be of no significance, yet its positive value indicates some effect on demand. A home located within a rural classified area, however, is found significant at the 1% level and having a large magnitude of 24.5. This suggests rural homes consume about 25 million BTU of fuelwood more than urban residences. Lastly, the dummy DRAFT, measuring whether the respondent believes the home is drafty all the time or most of the time during heating months is found to be of no significance, with a negative value as well.

Finally, recall we took into consideration how the fuelwood is used (Table 4.5) in this model by using mutually exclusive dummies to signify appliance, location, wood type, and other options. The base event is those respondents who fall into the other category (OTH), were multiple types, appliances, spaces, or other unforeseen survey situations occur. We find all options are significant at the 1% level except for pellets used for fire places or secondary heating stoves. The former is logical as most pellet appliances lead to using a heating stove, while the latter can be argued that if you're using a heating stove with pellets, you're probably using it for primary space heating.

Of those found to be significant, secondary heating spaces are all negative in sign with primary spaces being positive. Those who use pellets in their main space heating stove use 26.8 million BTU more fuelwood over the base. Those who use logs for main space heating fireplaces and stoves use 19.37 and 38.80 million BTU more fuelwood, respectively, over the base. We find the magnitudes of pellet users to be of lesser degree relative to log users, except for main fireplace users, which isn't found to be significant. Finally, those who use pellets for their secondary space heating stove and those using logs for their secondary space fireplace use 19.0 and 9.8 million BTU less than the base. While the term "secondary" is defined as spaces used less often, this possibly can be indicative of backup usage as well.

Looking into effects of location, Table 5.3, we estimate demand given divisional and urban effects, *ceteris paribus.* Across the OLS and Heckman models, we estimate rural regions to demand more than urban localized residences. Furthermore, it appears the farther north and east you investigate, the greater magnitude of fuelwood demanded. The OLS model significantly estimates lower demand across all regions and metro status.

Table 5.3. Estimates of Geographical and Urban Location Effects on Fuelwood Demand.

Division	OLS		Heckman, conditional	
	Urban	Rural	Urban	Rural
1: New England	4.0986	26.4671	35.9943	54.2755
2: Middle Atlantic	1.8114	16.8462	40.0601	59.5468
3: East North Central	2.3941	19.5146	35.7351	58.2570
4: West North Central	2.4187	18.3907	31.5824	54.2425
5: South Atlantic	1.9363	8.0274	30.3771	49.0733
6: East South Central	1.4679	8.1533	33.6676	53.4333
7: West South Central	1.1190	4.1176	30.5335	47.7124
8: Mountain North Sub-Division	2.0891	12.5581	23.8820	45.0872
9: Mountain South Sub-Division	2.3397	5.0909	22.9714	45.0299
10: Pacific	2.5965	20.4420	13.4395	47.5254

We find a wide variety of significance among the modeled variables, specifically the division and usage variables. As predicted, there is significant impact on demand from income and price effects, as well as climatic and household characteristics of heating needs determinants, which our research questioned. The next section extends the discussion to focus on four variables considered continuous within this study, which aid in determining elasticity measures of response on demand. Those are income, price, heating space, and winter temperature effects on demand.

5.4. Elasticity Results of Income, Non-wood Price, Heating Space, and Winter Temperature

Table 5.4 and Figure 5.1 depict elasticity estimates for wood users. Here we see how responsive to demand at different increments variable, *ceteris paribus*. We find income elasticity of demand is found to be negative in response and is decreasing at an increasing rate, answering the first research question on whether fuelwood is a necessary, inferior, or luxury good. This suggests as incomes increase, residential homeowners view fuelwood as an inferior good and negative income inelastic. Furthermore, we propose this to be attributed from two things: (1) homes necessarily don't require more heating as incomes increase, and (2) the price conscious homeowner is likely not concerned with the overall consumption habits of energy usage. The income demand elasticity response over ranges of different income values is reported in Figure 5.1 (a).

Table 5.4. Estimated Elasticities of Selected Continuous Variables.

INC	%Δfuelwood	NWP	%Δfuelwood	HDD65	%Δfuelwood	TOTHSQFT	%Δfuelwood
5	-0.0219	5	0.9784	500	0.1862	500	0.1834
10	-0.0448	7	0.9845	1070	0.3287	615	0.2164
15	-0.0687	9	0.9879	1640	0.4288	730	0.2469
20	-0.0937	11	0.9901	2210	0.5028	845	0.2751
25	-0.1199	13	0.9916	2780	0.5599	960	0.3012
30	-0.1475	15	0.9927	3350	0.6052	1075	0.3256
35	-0.1764	17	0.9936	3920	0.6421	1190	0.3483
40	-0.2068	19	0.9942	4490	0.6727	1305	0.3695
45	-0.2388	21	0.9948	5060	0.6984	1420	0.3894
50	-0.2726	23	0.9952	5630	0.7204	1535	0.4080
55	-0.3082	25	0.9956	6200	0.7394	1650	0.4256
60	-0.3460	27	0.9959	6770	0.7560	1765	0.4421
65	-0.3860	29	0.9962	7340	0.7706	1880	0.4578
70	-0.4283	31	0.9965	7910	0.7836	1995	0.4725
75	-0.4734	33	0.9967	8480	0.7951	2110	0.4865
80	-0.5214	35	0.9969	9050	0.8055	2225	0.4998
85	-0.5726	37	0.9970	9620	0.8149	2340	0.5124
90	-0.6275	39	0.9972	10190	0.8234	2455	0.5244
95	-0.6862	41	0.9973	10760	0.8312	2570	0.5358
100	-0.7494	43	0.9974	11330	0.8383	2685	0.5466
105	-0.8175	45	0.9976	11900	0.8449	2800	0.5570
110	-0.8912	47	0.9977	12470	0.8509	2915	0.5669
115	-0.9710	49	0.9978	13040	0.8565	3030	0.5764

Next we find the cross-price elasticity estimates using the composite average price of non-wood energy per million BTUs, found in Figure 5.1 (b). We estimate the average price paid for energy alternatives to be almost unit elastic, as when price increases it yields a positive response in fuel consumption across all units of measure. Fuelwood is estimated to be a substitute of energy alternatives, yet as prices increase the consumer's response is not necessarily to increase their consumed quantities of fuelwood at increasing or decreasing rates. They tend to switch at relatively lower price levels and continue to substitute at the same rate. This answers our second research question regarding the effect energy prices have on residential wood energy demand.

Our third research question looked at estimating the impacts of climate and location on demand, two variables hypothesized to impact heating needs as a whole regardless of the fuel used. We also include here the findings on home size on demand response. Both heating degree days below 65 °F and total heating space of the home are positive and increasing at a decreasing rate. A combination of items are thought to influence this including home construction, location, and heat retaining insulation properties, each of which should be addressed and researched for other studies.

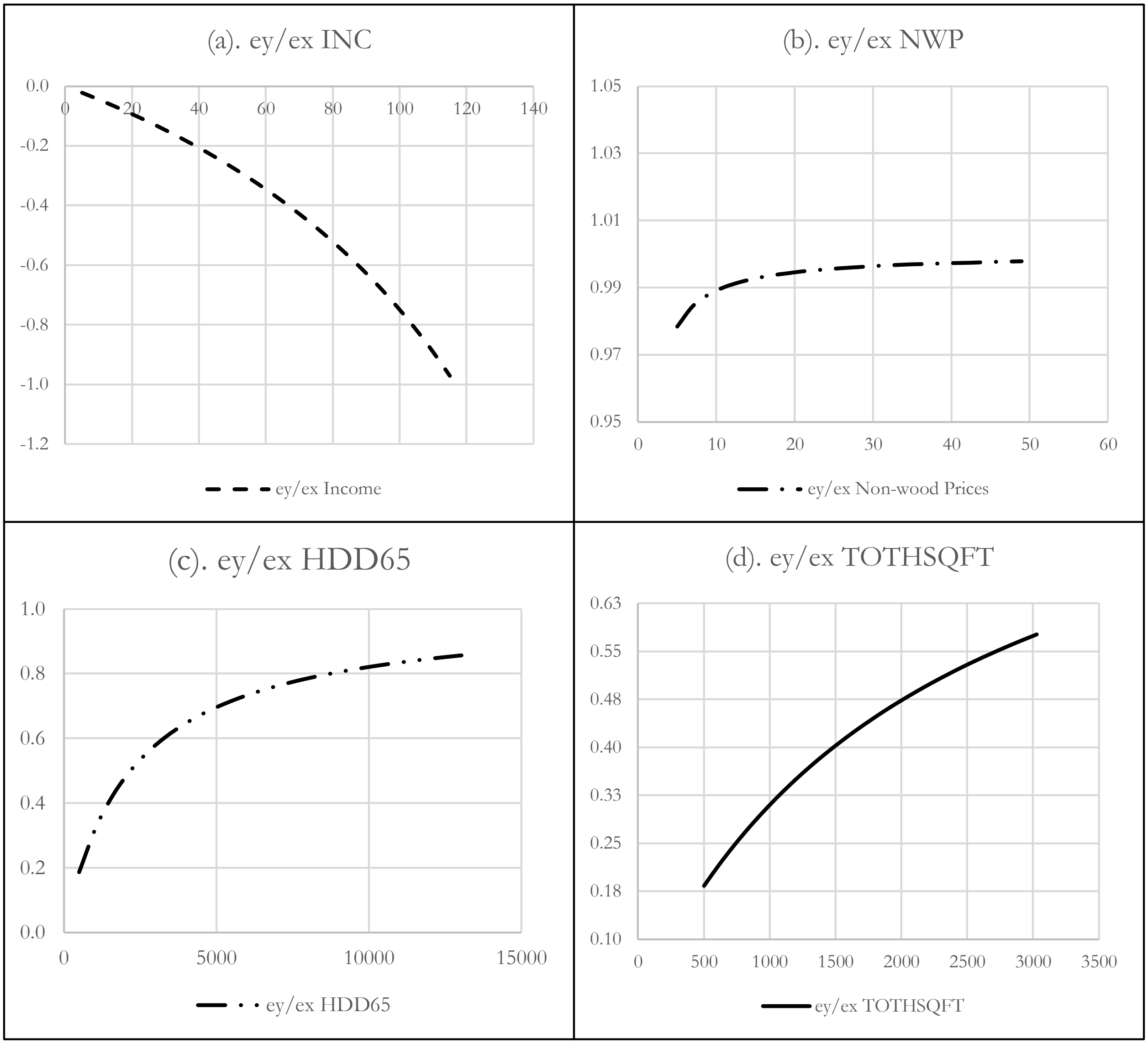

Figure 5.1. Elasticities of Chosen Continuous Variables from Heckman Model.

Finally, our last question asked whether we can forecast possible demand from these findings. Given the historical research by Song et al. (2012a, 2012b) finds demand is on a general decline, the results appear to conclude demand is concurrent with many factors, including location and prices of non-wood fuel sources. If firms are to monitor anything to judge production and pricing of their fuelwood products based upon variables that are economically more likely to change, non-wood energy prices have the greatest impact upon changes within demand for fuelwood. Type of fuelwood, application, and appliance appear to have overall magnitude influence as well, as those using logs for fuel within main-space heating stoves use more fuel than those using pellets for the same application, suggesting a hint at a possible efficiency difference; log use within main fireplaces use more fuel over secondary.

5.5. Summary

This chapter presented findings for applying a simple Heckman two-stage selectivity correction method for fuelwood demand estimation among U.S. residential homes in 2009. Primary research questions explored involve response of fuelwood demand given changes in non-wood alternative prices and household incomes, each of which confirmed to be statistically significant on overall demand. If a fuelwood firm is to question pricing strategy, given this particular market with a market share of approximately 3.2% relative to all available energy sources, it perhaps is ideal to monitor alternative market prices in order to react to pricing schemes or supply levels if demand changes. Income levels tend not to react quickly to market conditions, but we can see there is a clear difference in elasticity responses of lower income households vs. middle and higher income households. An indication of fuelwood being a necessary inferior good is present.

Finally, we looked to estimate and support selectivity playing a role in fuelwood demand. We found the lambda coefficient to be statistically significant at the 1% level, indicating there is evidence of selection among U.S. residential homes. Correcting for selection finds other variables significant in demand for fuelwood, including those attributed to wood type, appliance, and heating space chosen, as well as education, age of homeowner and home, and the large influence rural surroundings have on overall consumption.

CHAPTER 6: CONCLUSIONS AND IMPLICATIONS

This research conducted on U.S. residential fuelwood demand finds several conclusions. Given residential demand for fuelwood is historically declining as fossil fuels provide a modern usage convenience that can't easily be met, we face increasing concerns for sustainable and renewable energy. As such, society may need to reconsider using one of the oldest forms of generating heat. Although we didn't explore the ability of a residence to have multiple heating sources, the use of fuelwood as a backup source can provide a means of reducing exposure to market supply and price volatilities. It can also ease the transition towards a fuel with higher maintenance and lesser convenience due to continued cleanings or continuous self-fueling of heating appliances. This can prove beneficial to both fuel and appliance firms in marketing their products to a residence.

We also find potential for using non-wood fuel prices as a means for marketing such supply and stability issues. The cross-price unit-elasticity estimate suggest fuelwood users face many substitutes. In the event a wood-biomass firm needs to adjust prices, monitoring non-wood energy prices (as well as their supply levels, naturally) will aid in determining how their customers will respond. Furthermore, other identified factors can provide a means to determine how likely a residence will use the product and the quantity they will demand. Items such as household heating space, income level, location, age, ethnicity, and winter temperatures, all have significance to some degree. Fuelwood as a luxury, aesthetic, good is found significant in a prior study (Mackenzie and Weaver; 1986), but 30 years later we find the income elasticity estimate suggesting fuelwood is an inferior good, confirming our hypothesis.

Firms located in or marketing and shipping their products to those states in the Upper North East, Upper Midwest, or Pacific regions of the U.S. are more likely to make a sale leading to fuelwood consumption over those in the South or Rocky Mountains. Similarly, elasticity estimates of colder temperatures and larger heating spaces confirm positive consumption responses, but each face diminishing responses suggesting possible limitations of possible thermal ability fuelwood provides compared to fossil fuel substitutes. Lastly, conclusions on forecasting from these estimates should provide reliable and unbiased results in the short-term (Heckman, 1979; Song et al., 2012a). Given the data available to conduct this analysis is dated, however, an updated survey with renewable-based questions, similar to those found in Arabatzis and Malesios (2011), should aid in continued studies and monitoring of residential heating trends as energy policy continues to expand (The White House, n.d.).

Some implications arise from this study. Policy makers looking to promote fuelwood can focus on specific regions and appliances by developing tax credits, price regulation, or grants. Increasing taxes on non-wood energy prices can be a possible means of convincing residential homeowners to use wood as an alternative. Since the general population tend to dislike taxes, emission-based tax credits can also be designated towards appliances where the primary use will be for main heating spaces where heating stoves are used. There is also risk associated with the use of fuelwood, which the home-owner insurance industry recognizes through higher premiums. Newer homes may also tend not to install wood-based appliances, possibly opting for natural gas fireplaces due to convenience. We also confirm homes in rural locations continue to use more fuelwood over urban areas. While this can possibly be due to sourcing cheap, if not free, locally produced forestry, it can also be due to the lack of available fossil fuel alternatives leaving the residence with fuelwood as an only choice. As demand is higher within these locations, rural economic support through policy programs can aid in expanding firm startups or business expansion to promote in these areas in possible need. Similarly, policymakers concerned with poverty and ethnicity within specified locations can identify and target their objectives to those individuals' concerns.

This study also provides updated demand estimates given variable means in identified Census divisions. As selectivity is estimated to provide significance in the outcome of using fuelwood, focusing on these attributes will aid in increasing the probability of identified targeted markets in using fuelwood.

A few shortcomings of this study surround the use of data that is heavily designed to estimate fossil fuel consumption. Since it is the best available source for data as of this writing, developing a survey designated for renewable resource demand could prove useful in this case. Own-price estimates can provide firms a means to respond to observed fuelwood sales, but isn't possible due to the exclusion of fuelwood prices within the survey design. Lastly, use of the Dubin-McFadden polychotomous choice model with an analysis as fuelwood as a backup source, as in the Couture et al (2012) study, would provide additional angle in residential decision making.

In summary, to estimate demand of a product with such relatively little market impact, extra measures must be accounted for to prevent incorrect statistical findings, i.e., biased results of estimated coefficients. Fuelwood hasn't been monitored into such great detail as fossil fuel energy has. Because of this, accurate demand estimation is not possible, nor is it obtainable without modernizing survey data. The best data available to estimate from is from 2009, data that is already five years old. We merely find evidence the Heckman procedure corrects potential selection bias. Heckman (1979) himself states, "Given its simplicity and flexibility, the procedure outlined in this paper is recommended for exploratory empirical work" (p. 160).

We find evidence the Heckman procedure corrects potential selection bias. Heckman (1979) states, "given its simplicity and flexibility, the procedure outlined in this paper is recommended for exploratory empirical work" (p. 160). Using this study in conjunction with the Song et al (2012a, 2012b) findings, identifying market structures is possible. Had those papers included appliance and fuelwood types within the Tobit modeling, we'd have findings of latent demand and overall demand together in a unified summary. Perhaps this encourages other researchers to explore newer, more modern, methods of estimation or it may provide means to enforce a more robust dataset that contains price data for all energy sources. As the need for alternative renewable energy increases, for many reasons beyond the scope of this research, the need for consistent and accurate data will increase as well. We will need better measures for both supply and demand estimations so firms and governments alike can make due diligent decisions. The European market for wood based fuel pellets is growing (Peng et al., 2010). As such, we need to understand the U.S. energy market to its fullest in order to prepare for possible similar domestic demand increases as well.

REFERENCES

Anderson, Kent P. (1973). *Residential Energy Use: An econometric analysis.* RAND Corporation R-1297-NSF. Retrieved from the RAND.org website: http://www.rand.org/content/dam/rand/pubs/reports/2007/R1297.pdf on August 7, 2014.

Arabatzis, G., and Ch. Malesios. (2011). An econometric analysis of residential consumption of fuelwood in a mountainous prefecture of Northern Greece. *Energy Policy,* 39: 8088-8097. Retrieved on January 18, 2014 from Science Direct website: http://www.sciencedirect.com/science/article/pii/S0301421511007877.

Bourguignon, Fraçnois, Fournier, Martin and Gurgand, Marc. (2007). Selection Bias Correction Based on the Multinomial Logit Model: Monte Carlo Comparisons. *Journal of Economic Surveys,* 21(1): 174-205. Web, Retrieved on July 29, 2014.

Couture, Stéphane, Garcia, Serge, and Reynaud, Arnaud. (2012). "Household energy choices and fuelwood consumption: An econometric approach using French data." *Energy Economics*, 34: 1972-1981. DOI: 10.1016/j.eneco.2012.08.022.

Dahl, Gordon B. (2002). Mobility and the Return to Education: Testing a Roy Model with Multiple Markets. *Econometrica,* 70(6): 2367-2420.

Dubin, Jeffrey A., and McFadden, Daniel L. (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica,* 52(2): 345-362. Retrieved from the web on July 20, 2014.

Favero, Alice, and Mendelsohn, Robert. (2014). Using markets for woody biomass energy to sequester carbon in forests. *Journal of Association Environmental Resource Economics* 1: 75-95.

Gately, Dermot, H.G. Huntington. (2001). The asymmetric effects of changes in price and income on energy and oil demand. *The Energy Journal*, 23(1): 19-37. ISSN: 0195-6574.

Greene, William H. (2003). *Econometric Analysis* (5th ed.). Prentice Hall, Upper Saddle River New Jersey, USA.

Hardie, Ian W., and Aziz Abu Hassan. (1986). "An Econometric Analysis of Residential Demand for Fuelwood in the United States, 1980-1981. *Forest Science,* 32 (4): 1001-1015. Retrieved form the web on July 20, 2014.

Heckman, James J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement,* 5 (4): 475-492. Retrieved August 11, 2014 from the National Bureau of Economic Research website: http://www.nber.org/chapters/c10491.pdf.

Heckman, James J. (1979). Sample Selection Bias as a Specification Error. *Econometrica,* 47(1): 153-161.

Jebaraj, S. S. Iniyan. (2006). A review of energy models. *Renewable & Sustainable Energy Reviews*, 10: 281-311. doi: 10.1016/j.rser.2004.09.004

Johansson, Linda S., Bo Leckner, Lennart Gustavsson, David Cooper, Claes Tullin, and Annika Potter. (2004). Emission characteristics of modern and old-type residential boilers fired with wood logs and wood pellets. *Atmospheric Environment,* 38: 4183-4195. doi: 10.1016/j.atmosenv.2004.04.020.

Kaza, Nikhil. (2010). Understanding the spectrum of residential energy consumption: A quantile regression approach. *Energy Policy*, 38: 6574-6585. doi: 10.1016/j.enpol.2010.06.028.

Kebede, Ellene, Gbenga Ojumu, and Edinam Addozssi. (2013). Economic impact of wood pellet co-firing in South and West Alabama. *Energy for Sustainable Development,* 17: 252-256. doi: 10.1016/j.esd.2013.01.004.

Kennedy, Peter. (2008). *A guide to econometrics.* (6th ed.). Malden, MA: Blackwell Publishing.

Kialashaki, Arash, J.R. Reisel. (2013). Modeling of the energy demand of the residential sector in the United States using regression models and artificial neural networks. *Applied Energy*, 108: 271-280. doi: 10.1016/j.apenergy.2013.03.034.

Lee, Jung-Fei. (1982). Some Approaches to the Correction of Selectivity Bias. *Review of Economic Studies* 49, 355-372. Web, Retrieved July 29, 2014 from JSTOR website.

Lee, Jung-Fei. (1983). Notes and Comments Generalized Econometric Models with Selectivity. *Econometrica* 51, 507-512. Web, Retrieved July 29, 2014 from JSTOR website.

Mackenzie, John, and Weaver, Thomas F. (1986). A household production analysis of fuelwood demand in Rhode Island. *Northeastern Journal of Agricultural and Resource Economics,* 15 (1): 53-60. Retrieved on July 8, 2014 from AgEcon Search: http://ageconsearch.umn.edu/handle/28884.

Mansur, Erin T., Mendelsohn, Robert, and Morrison,Wendy. (2008). Climate change adaptation: A Study of fuel choice and consumption in the U.S. energy sector. *Journal of Environmental Economics and Management,* 55: 175-193. Retrieved on August 24, 2014 from http://www.sciencedirect.com/science/article/pii/S0095069607001040.

National Renewable Energy Laboratory. (2013). Energy Analysis: International Trade of Wood Pellets. Web, Retrieved May 21, 2014 from the National Renewable Energy Laboratory website: http://www.nrel.gov/docs/fy13osti/56791.pdf.

National Weather Service (NWS). (n.d.). Explanation of the weekly and monthly degree day data summaries. Climate Prediction Center, National Weather Service. Web, Retrieved June 17, 2014 from the National Weather Service website: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/ddayexp.shtml.

Office of Management and Budget. (n.d.). *Creating the Clean Energy of Tomorrow and Protecting the Environment.* Retrieved Sept 13, 2014 from the Whitehouse website: http://www.whitehouse.gov/omb/factsheet/creating-the-clean-energy-of-tomorrow-and-protecting-the-environment.

Pa, Ann, Bi, X.T., Sokhansanj, Shahab. (2013). Evaluation of wood pellet application for residential heating in British Columbia based on a streamlined life cycle analysis. *Biomass and Bioenergy*, 49: 109-122. doi: 10.1016/j.biombioe.2012.11.009.

Peng, J.H., Bi, H.T., Sokhansanj, S., Lim, J.C., and Melin, S. (2010). An economical and market analysis of Canadian wood pellets. *International Journal of Green Energy,* 7: 128-142. Retrieved January 16, 2014 from http://www.tandfonline.com/doi/pdf/10.1080/15435071003673518.

Pindyck, Robert S., and Rubinfeld, Daniel L. (1998). *Econometric models and economic forecasts* (4th ed.). Boston: Irwin/McGraw-Hill.

Quigley, John M., and Rubinfeld, Daniel L. (1989). Unobservables in consumer choice: Residential energy and the demand for comfort. *The Review of Economics and Statistics,* 71(3): 416-425. Retrieved from the web: http://www.jstor.org/stable/1926898 on August 11, 2014.

Schmertmann, Carl P. (1994). Selectivity bias correction methods in polychotomous sample selection models. *Journal of Econometrics,* 60: 101-132. Retrieved July 30, 2014.

Song, N, F.X. Aguilar, S.R. Shifley, and M.E. Goerndt. (2012a). Analysis of U.S. residential wood energy consumption: 1967-2009. *Energy Economics*, 34: 2116-2124. doi: 10.1016/j.eneco.2012.03.004.

Song, Nianfu, F.X. Aguilar, S.R. Shifley, M.E. Goerndt. (2012b). Factors affecting wood energy consumption by U.S. households. *Energy Economics*, 34: 389-397. doi: 10.1016/j.eneco.2011.12.009

Spelter, Henry, and Toth, Daniel. (2009). North America's Wood Pellet Sector. U.S. Department of Agriculture, Forest Service. FPL-RP-656. Retrieved May 21, 2014 from Forest Products Laboratory website: http://www.fpl.fs.fed.us/documnts/fplrp/fpl_rp656.pdf?origin=publication_detail.

Standards of Performance for New Residential Wood Heaters, 40 C.F.R. § 60.531 (1988).

Standards of Performance for New Residential Wood Heaters, 40 C.F.R. § 60.531 (1988).

Standards of Performance for New Residential Wood Heaters; New Residential Hydronic Heaters and Forced-Air Furnaces, and New Residential Masonry Heaters, 79 Fed. Reg. 2014-00409 (Feb. 3, 2014) (to be codified at 40 C.F.R. pt. 60). [(Standards of Performance for New Residential Wood Heaters, 2014)]

Stern, David I. (2011). The Role of Energy in Economic Growth. *Ecological Economics Reviews* 1219: 26-51. Retrieved August 25, 2014 from the web.

Sustainable. (2014). In *Merriam-Webster.com* Retrieved May 21, 2014, from http://www.merriam-webster.com/dictionary/sustainable.

Tobin, James. (1958). Estimation of relationships for limited dependent variables. *Econometrica,* 26: 24-36. Retrieved Dec. 29, 2013 from JSTOR website: http://www.jstor.org/stable/1907382.

U.S. Census Bureau. (n.d.[a]). Census Regions and Divisions of the United States. Retrieved June 14, 2014 from the U.S. Census Bureau website: http://www.census.gov/geo/maps-data/maps/pdfs/reference/us_regdiv.pdf.

U.S. Census Bureau. (n.d.[b]). Census Bureau Regions and Divisions with State FIPS Codes. Retrieved June 14, 2014 from the U.S. Census Bureau website: http://www.census.gov/geo/maps-data/maps/pdfs/reference/us_regdiv.pdf.

U.S. Coast Guard. (2011). *On Scene Coordinator Report Deepwater Horizon Oil Spill.* pg. 21. Retrieved January 10, 2014 from the United States Coast Guard website: http://www.uscg.mil/foia/docs/dwh/fosc_dwh_report.pdf.

U.S. Energy Information Administration. (n.d.). *About the RECS*. Retrieved July 4, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/consumption/residential/about.cfm.

U.S. Energy Information Administration. (1994). *Estimates of U.S. Biomass Energy Consumption 1992.* Retrieved July 4, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/renewable/renewables/054892.pdf.

U.S. Energy Information Administration. (2008). *Short-Term Energy and Winter Fuels Outlook: Highlights*. Retrieved August 8, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/forecasts/steo/archives/oct08.pdf.

U.S. Energy Information Administration. (2009a). *Residential Energy Consumption Survey*: Household Questionnaire, Form A, pg. 2. Retrieved June 6, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/survey/form/eia_457/form.pdf.

U.S. Energy Information Administration. (2009b). *Residential Energy Consumption Survey*, Survey Forms. Retrieved June 6, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/survey/#eia-457.

U.S. Energy Information Administration. (2010a). *Annual Energy Review 2009. Release Date: August 2010.* Figure 1.0 Energy Flow, 2009 (Quadrillion BTU). Retrieved January 9, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/archive/038409.pdf.

U.S. Energy Information Administration. (2010b). *Annual Energy Review 2009. Release Date: August 2010.* Table 10.1 Renewable Energy Production and Consumption by Primary Energy Source, Selected Years, 1949-2009 (Trillion BTU), pg. 283. Retrieved January 10, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/archive/038409.pdf.

U.S. Energy Information Administration. (2010c). *Annual Energy Review 2009. Release Date: August 2010.* Figure 2.1a Energy consumption by Sector Overview (End-Use Sector Shares of Total Consumption, 2009), pg 38. Retrieved January 9, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/archive/038409.pdf.

U.S. Energy Information Administration. (2010d). *Annual Energy Review 2009. Release Date: August 2010.* Table 2.1b Residential Sector Energy Consumption, Selected Years, 1949-2009 (Trillion BTU), pg. 41. Retrieved January 10, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/archive/038409.pdf.

U.S. Energy Information Administration. (2010f). *Short-Term Energy and Winter Fuels Outlook. Release Date: October 6, 2009.* Highlights, pg. 1; Projected Winter Fuel Expenditures by Fuel and Region, pg. 2-11. Retrieved January 11, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/forecasts/steo/archives/oct09.pdf.

U.S. Energy Information Administration. (2011). *Flattening of per-capita residential energy consumption reflects multiple factors.* Today in Energy, Chart, Retrieved June 20, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/todayinenergy/detail.cfm?id=3590.

U.S. Energy Information Administration. (2012a). Annual Energy Review 2011. Release Date: September 27, 2012. Table 2.4 Household Energy consumption by Census Region, Selected Years, 1978-2009 (Quadrillion BTU). Retrieved August 8, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/showtext.cfm?t=ptb0204.

U.S. Energy Information Administration. (2012b). *Annual Energy Review 2011. Release Date: September 27, 2012.* Table 1.3 Primary Energy Consumption Estimates by Source, 1949-2011 (Quadrillion BTU). Retrieved January 9, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/showtext.cfm?t=ptb0103.

U.S. Energy Information Administration. (2012c). Residential Energy Consumption Survey (RECS): 2009 RECS Survey Data, Table CE1.1 Summary totals and intensities. U.S. homes, and CE5.2 Household wood consumption. Release date December 14, 2012 and February 21, 2014, respectively, from the U.S. Energy Information Administration website: http://www.eia.gov/consumption/residential/data/2009/index.cfm?view=consumption#summary.

U.S. Energy Information Administration. (2012d). *The impact of increasing home size on energy demand.* Analysis & Projections, Residential Energy Consumption Survey. Retrieved June 21, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/consumption/residential/reports/2009/square-footage.cfm.

U.S. Energy Information Administration. (2012e). Annual Energy Review 2011. Release Date: September 27, 2012. Table 2.6 Household End Uses: Fuel Types and Appliances and Electronics, Selected Years, 1978-2009 (Quadrillion BTU). Retrieved August 8, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/annual/showtext.cfm?t=ptb0206.

U.S. Energy Information Administration. (2013a). *New U.S. homes are 30% larger but consume about as much energy as older homes.* Today in Energy, February 12, 2013. Retrieved June 20, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/todayinenergy/detail.cfm?id=9951&src=%E2%80%B9%20Consumption%20%20%20%20%20%20Residential%20Energy%20Consumption%20Survey%20(RECS)-f2.

U.S. Energy Information Administration. (2013b). *Heating and cooling no longer majority of U.S. home energy use.* Today in Energy, March 7, 2013. Retrieved June 26, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/todayinenergy/detail.cfm?id=10271&src=%E2%80%B9%20Consumption%20%20%20%20%20%20Residential%20Energy%20Consumption%20Survey%20(RECS)-b1.

U.S. Energy Information Administration. (2013c). Short-Term Energy and Winter Fuels Outlook, October 8, 2013. Retrieved January 30, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/forecasts/steo/special/winter/2013_winter_fuels.pdf.

U.S. Energy Information Administration. (2013d). Monthly Energy Review, August 2014. Table 2.1 Energy Consumption by Sector (Trillion BTU). Retrieved August 30, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/monthly/pdf/sec2_3.pdf.

U.S. Energy Information Administration. (2014a). Today in Energy: Northeast and Mid-Atlantic power prices react to freeze and natural gas constraints. Retrieved May 21, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/todayinenergy/detail.cfm?id=14671.

U.S. Energy Information Administration. (2014b). Monthly Energy Review July 2014. Release Date: July 28, 2014. Table 2.1 Energy Consumption by Sector, 1949-2013. Retrieved August 8, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/browser/xls.cfm?tbl=T02.01&freq=m.

U.S. Energy Information Administration. (2014c). Monthly Energy Review July 2014. Release Date: July 28, 2014. Table 10.2a Renewable Energy Consumption: Residential and Commercial Sectors, 1949-2013. Retrieved August 8, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/browser/xls.cfm?tbl=T10.02A&freq=m.

U.S. Energy Information Administration. (2014d). *U.S. wood pellet exports double in 2013 in response to growing European demand*. Today in Energy, May 22, 2014. Retrieved July 27, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/todayinenergy/detail.cfm?id=16391.

U.S. Energy Information Administration. (2014e). Monthly Energy Review August 2014. Release Date: August 27, 2014. Table 2.2 Residential Sector Energy Consumption (Trillion BTU). Retrieved August 31, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/totalenergy/data/monthly/pdf/sec2_5.pdf.

U.S. Energy Information Administration. (2014f). Today in Energy: Propane use for crop drying depends on weather and corn markets as well as crop size. Retrieved October 28, 2014 from the U.S. Energy Information Administration website: http://www.eia.gov/todayinenergy/detail.cfm?id=18231.

The White House, Office of Management and Budget. (n.d.). Creating the Clean Energy of Tomorrow and Protecting the Environment. [Press release]. Retrieved on October 4, 2014 from the White House website: http://www.whitehouse.gov/omb/factsheet/creating-the-clean-energy-of-tomorrow-and-protecting-the-environment.

Van der Klaauw, Bas, and Koning, Ruud H. (2003). Testing the normality assumption in the sample selection model with an application to travel demand. *Journal of Business and Economic Statistics,* 21 (1): 31-42. Retrieved August 15, 2014 from http://personal.vu.nl/b.vander.klaauw/snpmlejbes.pdf.

Vella, Francis. (1998). Estimating Models with Sample Selection Bias: A Survey. *The Journal of Human Resources* 33 (1): 127-169. Retrieved August 11, 2014 from http://www.jstor.org/stable/146317.

Wooldridge, Jeffery. (2013). *Introductory econometrics: A modern approach* (5th ed.). Mason, OH: Cengage Learning.

Yergin, Daniel. (2012). The Quest. New York, New York. The Penguin Group.

APPENDIX A. USAGE DESCRIPTIVE STATISTICS

Variable	Value	Mean (Std. dev.)	Variable	Value	Mean (Std. dev.)
FUELNOHEAT	1 = Natural Gas	0.0100 (0.1000)	RNGFUEL	1 = Natural Gas	0.0027 (0.0522)
	2 = Propane/LPG	0.0002 (0.0129)		2 = Propane/LPG	0.0004 (0.0203)
	3 = Fuel Oil	– –		3 = Fuel Oil	– –
	4 = Kerosene	– –		4 = Kerosene	– –
	5 = Electricity	0.0137 (0.1164)		5 = Electricity	0.0036 (0.0602)
	7 = Wood	0.0004 (0.0234)		7 = Wood	0.0003 (0.0182)
	8 = Solar	– –		8 = Solar	– –
	9 = District Steam	0.0000 (0.0091)		9 = District Steam	– –
	21 = Other Fuel	0.0000 (0.0091)		21 = Other Fuel	– –
	-2 = Not Applicable	0.9754 (0.1548)		-2 = Not Applicable	0.9929 (0.0841)
FURELHEAT	1 = Natural Gas	0.4885 (0.4999)	DIFFUEL	1 = Natural Gas	0.0007 (0.0273)
	2 = Propane/LPG	0.0403 (0.1967)		2 = Propane/LPG	0.0024 (0.0489)
	3 = Fuel Oil	0.0676 (0.2511)		3 = Fuel Oil	0.0000 (0.0091)
	4 = Kerosene	0.0043 (0.0655)		4 = Kerosene	– –
	5 = Electricity	0.3342 (0.4717)		5 = Electricity	0.0026 (0.0514)
	7 = Wood	0.0244 (0.1543)		7 = Wood	0.0000 (0.0091)
	8 = Solar	0.0000 (0.0091)		8 = Solar	0.0002 (0.0158)
	9 = District Steam	0.0019 (0.0436)		9 = District Steam	– –
	21 = Other Fuel	0.0017 (0.0417)		21 = Other Fuel	0.0003 (0.0182)
	-2 = Not Applicable	0.0369 (0.1886)		-2 = Not Applicable	0.9935 (0.0806)
FURNFUEL	1 = Natural Gas	0.0104 (0.1016)	FUELH2O	1 = Natural Gas	0.5280 (0.4992)
	2 = Propane/LPG	0.0040 (0.0629)		2 = Propane/LPG	0.0330 (0.1787)
	3 = Fuel Oil	0.0022 (0.0472)		3 = Fuel Oil	0.3782 (0.1908)
	4 = Kerosene	– –		4 = Kerosene	0.0002 (0.0158)
	5 = Electricity	0.0057 (0.0754)		5 = Electricity	0.3950 (0.4889)
	7 = Wood	0.0000 (0.0091)		7 = Wood	0.0009 (0.0302)
	8 = Solar	– –		8 = Solar	0.0014 (0.0375)
	9 = District Steam	– –		9 = District Steam	– –
	21 = Other Fuel	– –		21 = Other Fuel	0.0007 (0.0273)
	-2 = Not Applicable	0.0001 (0.0091)		-2 = Not Applicable	0.0028 (0.0530)
RADFUEL	1 = Natural Gas	0.0021 (0.0454)	FUELH2O2	1 = Natural Gas	0.0134 (0.1150)
	2 = Propane/LPG	0.0003 (0.0182)		2 = Propane/LPG	0.0024 (0.0489)
	3 = Fuel Oil	0.0018 (0.0463)		3 = Fuel Oil	0.0013 (0.0364)
	4 = Kerosene	– –		4 = Kerosene	– –
	5 = Electricity	0.0004 (0.0203)		5 = Electricity	0.0142 (0.1181)
	7 = Wood	0.0000 (0.0091)		7 = Wood	0.0002 (0.0129)
	8 = Solar	– –		8 = Solar	0.0005 (0.0223)
	9 = District Steam	– –		9 = District Steam	– –
	21 = Other Fuel	0.0002 (0.0129)		21 = Other Fuel	0.0003 (0.0182)
	-2 = Not Applicable	0.9951 (0.0697)		-2 = Not Applicable	0.9677 (0.1767)
PIPEFUEL	1 = Natural Gas	0.0014 (0.0375)	HSFUEL	7 = Wood	0.0238 (0.1523)
	2 = Propane/LPG	0.0000 (0.0091)		21 = Other Fuel	0.0007 (0.0257)
	3 = Fuel Oil	– –		-2 = Not Applicable	0.9756 (0.1543)
	4 = Kerosene	– –			
	5 = Electricity	0.0005 (0.0223)	FPFUEL	1 = Natural Gas	0.0457 (0.2088)
	7 = Wood	– –		2 = Propane/LPG	0.0114 (0.1063)
	8 = Solar	– –		7 = Wood	0.0584 (0.2346)
	9 = District Steam	– –		21 = Other Fuel	0.0042 (0.0648)
	21 = Other Fuel	– –		-2 = Not Applicable	0.8802 (0.3247)
	-2 = Not Applicable	0.9980 (0.0445)			

Source: U.S. Energy Information Administration, Residential Energy Consumption Survey, public codebook (2009).

APPENDIX B. CORRELATION MATRIX OF RECS VARIABLES USED

	btuwood	usewood	tothsqft	hdd65	hdd30yr	yearbuilt	hhage	educ-1	educ-2	educ-3
btuwood	1									
usewood	0.6025	1								
tothsqft	0.0925	0.1485	1							
hdd65	0.1009	0.0962	0.1844	1						
hdd30yr	0.1029	0.0993	0.1849	0.9909	1					
yearbuilt	-0.0252	0.0051	0.1689	-0.1797	-0.1839	1				
hhage	0.0235	0.0234	0.1028	0.0303	0.0301	-0.0638	1			
educ-1	-0.0052	-0.0472	-0.1446	-0.035	-0.0357	-0.0665	0.0783	1		
educ-2	0.0353	-0.011	-0.1015	0.0188	0.0185	-0.0495	0.0509	-0.1832	1	
educ-3	-0.0006	0.007	-0.0311	-0.002	-0.0046	0.0264	-0.0578	-0.1641	-0.3216	1
educ-4	0.0161	0.0287	0.0121	0.0118	0.011	0.0132	-0.0385	-0.1012	-0.1984	-0.1776
educ-5	-0.0301	0.0082	0.1363	-0.0017	0.0005	0.0449	-0.0605	-0.1533	-0.3005	-0.2691
educ-6	-0.0155	0.0127	0.117	0.0105	0.0125	0.0514	0.0116	-0.0897	-0.1758	-0.1574
educ-7	-0.0102	0.0187	0.0791	-0.0023	-0.0027	-0.0186	0.0269	-0.0417	-0.0818	-0.0732
educ-8	-0.0114	0.0076	0.0557	0.0044	0.0051	-0.002	0.0298	-0.0351	-0.0688	-0.0616
wdpellet	0.1714	0.2564	0.0403	0.0547	0.0561	-0.0018	-0.0036	-0.0161	0.002	-0.0079
woodlogs	0.565	0.9257	0.1445	0.0937	0.0972	0.0014	0.0266	-0.0397	-0.0131	0.0134
drafty-1	-0.0131	-0.0129	-0.0771	0.0054	0.0039	-0.1136	-0.0879	0.0288	0.0211	0.0318
drafty-2	-0.0157	-0.0081	-0.0735	0.0031	0.0032	-0.0751	-0.1	0.0298	0.0161	0.0037
drafty-3	0.0163	0.0376	-0.026	0.0779	0.0787	-0.1085	-0.0645	-0.0108	0.0015	0.0108
drafty-4	0.0002	-0.0237	0.1044	-0.0769	-0.0768	0.2011	0.1602	-0.0212	-0.0212	-0.0287
moneypy	0.0234	0.1152	0.4285	0.0439	0.0481	0.1263	-0.0699	-0.2433	-0.2009	-0.0515
totaldol	0.0449	0.1019	0.5114	0.1189	0.1265	-0.0198	0.0478	-0.0876	-0.0442	-0.0189
totalbtu	-0.0003	0.0599	0.5516	0.3705	0.376	-0.1139	0.077	-0.0797	-0.0489	-0.0109
ur	0.2212	0.1724	0.1694	0.103	0.0969	0.1752	0.0701	-0.0077	0.0457	0.0062
own-1	0.0978	0.1512	0.4189	0.0754	0.0762	0.0924	0.2894	-0.1211	-0.0408	-0.0341
own-2	-0.1008	-0.1538	-0.4164	-0.0775	-0.0783	-0.0858	-0.2898	0.1129	0.0406	0.0378
own-3	0.009	0.0051	-0.0284	0.0061	0.0059	-0.0321	-0.0104	0.0408	0.0026	-0.0143
race-1	0.0555	0.0797	0.1054	0.1205	0.12	-0.0038	0.1112	-0.0181	0.0028	-0.0074
race-2	-0.0447	-0.0664	-0.0787	-0.0989	-0.0994	0.0026	-0.0555	0.0181	0.0363	0.0257
race-3	0.0015	0.0025	-0.0295	0.0048	0.0046	-0.0048	-0.0287	0.0112	0.0038	0.0009
race-4	-0.0385	-0.054	-0.0294	-0.0755	-0.0732	0.0228	-0.0563	-0.0172	-0.0677	-0.0481
race-5	-0.005	-0.0078	-0.0296	-0.0384	-0.0414	0.0087	-0.005	0.003	0.0112	-0.0102
race-6	-0.014	-0.0138	-0.0339	-0.0112	-0.0102	-0.0174	-0.0615	0.027	0.0032	-0.0063
race-7	0.0139	0.017	-0.0179	0.0132	0.0132	-0.0122	-0.0414	-0.0013	-0.0142	0.0422
typehuq-1	0.0289	-0.0126	-0.1248	-0.0361	-0.039	0.1153	0.0147	0.0941	0.0527	0.002
typehuq-2	0.1104	0.1819	0.4439	0.0437	0.0452	-0.0027	0.1607	-0.0997	-0.0212	-0.0113
typehuq-3	-0.0482	-0.0614	-0.0589	0.0136	0.0156	0.0212	-0.0219	-0.0058	-0.0181	0.0114
typehuq-4	-0.0545	-0.0846	-0.1848	0.0106	0.0107	-0.1366	-0.1053	0.0577	0.0186	0.0075
typehuq-5	-0.0866	-0.1254	-0.3334	-0.0542	-0.056	0.0225	-0.1262	0.0393	-0.0026	0.0001
d-1	0.0509	0.0628	0.0098	0.2933	0.3017	-0.1622	0.0162	-0.0102	-0.009	-0.0205
d-2	-0.0211	-0.0389	-0.0061	0.2025	0.2221	-0.1832	0.0142	0.009	0.0204	-0.0526
d-3	0.0152	-0.0065	0.0707	0.3256	0.3221	-0.0995	0.0146	-0.0003	0.0173	0.0155
d-4	0.0509	0.0222	0.1159	0.3519	0.3459	-0.0113	0.0228	-0.0185	0.0149	0.0244
d-5	-0.0232	-0.03	0.0143	-0.3631	-0.3766	0.1738	-0.0034	-0.0206	0.0046	-0.0174
d-6	0.001	-0.0232	0.007	-0.0717	-0.0707	0.0693	0.0134	0.033	0.0459	-0.0138
d-7	-0.0464	-0.0482	-0.0337	-0.3041	-0.2991	0.0818	-0.0502	0.0282	0.0056	0.0224
d-8	-0.0038	0.0356	0.0425	0.1835	0.179	0.056	-0.0208	-0.0315	-0.0275	0.0417
d-9	-0.0152	-0.0068	-0.0193	-0.157	-0.1569	0.1032	0.0113	-0.0075	-0.0208	0.0173
d-10	-0.0079	0.0408	-0.168	-0.3136	-0.3159	-0.0053	-0.0143	0.0156	-0.0506	0.0031

Note: All variables found within the 2009 RECS survey responses, data, and codebook.
educ-n: found as education,
own-n: found as kownrent,
race-n: found as householder_race,
d-n: found as division.

	educ-4	educ-5	educ-6	educ-7	educ-8	wdpellet	woodlogs	drafty-1	drafty-2	drafty-3
btuwood										
usewood										
tothsqft										
hdd65										
hdd30yr										
yearbuilt										
hhage										
educ-1										
educ-2										
educ-3										
educ-4	1									
educ-5	-0.166	1								
educ-6	-0.0971	-0.1471	1							
educ-7	-0.0452	-0.0684	-0.04	1						
educ-8	-0.038	-0.0575	-0.0337	-0.0157	1					
wdpellet	0.0135	0.006	-0.0046	0.0137	0.0049	1				
woodlogs	0.0206	0.0053	0.0151	0.0165	0.0066	0.0147	1			
drafty-1	-0.0078	-0.0369	-0.0478	-0.015	-0.0157	-0.0097	-0.0172	1		
drafty-2	-0.0056	-0.0094	-0.0225	-0.024	-0.0147	0.0118	-0.0101	-0.083	1	
drafty-3	0.0067	-0.0061	-0.0061	0.0028	-0.0046	0.0133	0.0366	-0.1903	-0.1974	1
drafty-4	0.0009	0.0302	0.043	0.0183	0.0205	-0.0137	-0.0195	-0.3034	-0.3146	-0.7219
moneypy	0.0388	0.2365	0.2093	0.1118	0.1077	0.0411	0.1043	-0.1068	-0.0593	-0.0056
totaldol	0.0015	0.068	0.0561	0.0718	0.0369	0.0479	0.0925	-0.0141	-0.0241	0.0198
totalbtu	-0.0057	0.0642	0.0632	0.0701	0.0292	0.0136	0.0585	-0.0005	-0.0105	0.0499
ur	0.0106	-0.0354	-0.0246	-0.0191	-0.0156	0.0636	0.1641	-0.0456	-0.0285	-0.0096
own-1	0.0233	0.0741	0.091	0.0476	0.0315	0.0465	0.1408	-0.1449	-0.0933	-0.0103
own-2	-0.0235	-0.0737	-0.0891	-0.0446	-0.0321	-0.0484	-0.1429	0.1405	0.0924	0.0082
own-3	-0.0001	-0.0049	-0.0121	-0.0149	0.0011	0.0062	0.0033	0.0249	0.0078	0.0095
race-1	0.0023	0.0119	0.0086	-0.0048	0.0046	0.024	0.0752	-0.0658	-0.0375	-0.0155
race-2	0.0069	-0.0498	-0.0297	-0.0219	-0.0192	-0.0276	-0.0598	0.0607	0.0433	0.0271
race-3	0.0092	-0.0198	0.0009	-0.0001	-0.011	0.0097	0.0018	0.0295	-0.0055	-0.0081
race-4	-0.0133	0.0868	0.0495	0.0571	0.0271	-0.0093	-0.0531	-0.0213	-0.0051	-0.007
race-5	0.0051	-0.0001	-0.0062	-0.0079	-0.0066	-0.0054	-0.0054	-0.0053	0.0098	-0.0077
race-6	-0.0081	-0.0148	-0.0134	0.0007	0.0014	0.001	-0.0183	0.0395	-0.0017	0.0031
race-7	-0.0066	-0.009	-0.009	-0.0111	0.0049	0.0038	0.0171	0.0252	0.017	-0.0045
typehuq-1	-0.018	-0.0766	-0.0591	-0.0206	-0.0213	-0.0075	-0.0157	0.0427	0.0136	0.0055
typehuq-2	0.0288	0.054	0.048	0.0197	0.0147	0.0548	0.1726	-0.0746	-0.0493	-0.0015
typehuq-3	-0.0073	0.0072	0.017	-0.003	0.0096	-0.0231	-0.0594	0.0279	0.01	-0.0003
typehuq-4	-0.0077	-0.035	-0.0396	-0.0184	-0.0083	-0.0204	-0.0805	0.0658	0.048	0.0124
typehuq-5	-0.0166	-0.007	-0.0128	0.0014	-0.008	-0.0361	-0.1158	0.0057	0.0148	-0.0099
d-1	0.0087	0.0058	0.0272	0.0227	0.0213	0.0623	0.0528	0.0053	-0.0015	0.0218
d-2	-0.0134	0.0173	0.0233	0.0093	-0.0076	0.001	-0.0368	0.0037	0.0047	0.0435
d-3	-0.009	-0.0184	-0.0022	-0.0106	-0.0173	-0.0033	-0.0037	0.0181	0.0051	0.0235
d-4	0.0127	-0.0156	-0.023	0.0018	0.0063	-0.0124	0.0296	-0.0086	-0.0204	0.0197
d-5	0.0144	-0.0028	0.0158	-0.0081	0.0034	-0.0221	-0.0289	-0.0174	-0.0054	-0.034
d-6	-0.0134	-0.0295	-0.0176	-0.0175	-0.0099	-0.0137	-0.0143	-0.0093	0.0061	-0.019
d-7	-0.0068	-0.0151	-0.0258	-0.0174	-0.012	-0.0258	-0.047	0.025	-0.0014	-0.0179
d-8	0.003	0.0083	0.0029	-0.007	0.0008	0.0193	0.0302	-0.015	0.0159	0.0354
d-9	0.0077	0.0063	-0.0019	-0.0062	0.001	-0.0115	-0.0079	-0.0169	-0.0323	-0.0504
d-10	-0.0063	0.0354	-0.0017	0.0215	0.0098	0.0161	0.0329	0.0055	0.0219	-0.0222

Note: All variables found within the 2009 RECS survey responses, data, and codebook.
educ-n: found as education,
own-n: found as kownrent,
race-n: found as householder_race,
d-n: found as division.

	draft-4	moneypy	totaldol	totalbtu	ur	own-1	own-2	own-3	race-1	race-2
btuwood										
usewood										
tothsqft										
hdd65										
hdd30yr										
yearbuilt										
hhage										
educ-1										
educ-2										
educ-3										
educ-4										
educ-5										
educ-6										
educ-7										
educ-8										
wdpellet										
woodlogs										
drafty-1										
drafty-2										
drafty-3										
drafty-4	1									
moneypy	0.0934	1								
totaldol	0.002	0.3569	1							
totalbtu	-0.0404	0.3269	0.7986	1						
ur	0.0482	0.023	0.1439	0.0548	1					
own-1	0.1361	0.3545	0.3261	0.3287	0.193	1				
own-2	-0.1314	-0.3425	-0.3304	-0.3305	-0.2013	-0.9734	1			
own-3	-0.0261	-0.0666	0.0051	-0.006	0.0278	-0.1567	-0.0739	1		
race-1	0.0693	0.1006	0.06	0.0594	0.1056	0.1809	-0.1807	-0.0086	1	
race-2	-0.0805	-0.1352	-0.0255	-0.0249	-0.0647	-0.1593	0.1564	0.0189	-0.7409	1
race-3	-0.005	-0.0293	-0.0272	-0.0216	0.0128	-0.0318	0.0345	-0.0105	-0.1874	-0.0363
race-4	0.0204	0.0539	-0.0488	-0.0531	-0.0842	-0.0304	0.0301	0.0025	-0.3877	-0.0751
race-5	0.0046	0.0078	-0.0097	-0.0301	-0.0073	-0.0152	0.0168	-0.0063	-0.1127	-0.0218
race-6	-0.0227	-0.0225	-0.0153	-0.0179	-0.04	-0.0581	0.0607	-0.0087	-0.2607	-0.0505
race-7	-0.0183	-0.0084	-0.0121	0.0037	-0.0073	-0.0278	0.0295	-0.0065	-0.2336	-0.0453
typehuq-1	-0.0349	-0.1384	-0.0418	-0.0826	0.1402	0.0568	-0.0613	0.0172	0.0584	-0.0494
typehuq-2	0.0673	0.311	0.3639	0.3846	0.1942	0.5825	-0.5872	-0.0043	0.1412	-0.1084
typehuq-3	-0.0198	-0.0028	-0.0606	-0.0532	-0.0947	-0.0632	0.0635	0.0016	-0.0379	0.0222
typehuq-4	-0.072	-0.1515	-0.1115	-0.0943	-0.1118	-0.3184	0.3201	0.0061	-0.0829	0.0645
typehuq-5	-0.0018	-0.2162	-0.3278	-0.3497	-0.1842	-0.517	0.5243	-0.0097	-0.1302	0.1069
d-1	-0.0222	0.0504	0.1789	0.1288	-0.0165	-0.0184	0.0186	-0.0001	0.0248	-0.0437
d-2	-0.0449	0.0296	0.1682	0.117	-0.0903	-0.0438	0.0434	0.0034	-0.005	0.001
d-3	-0.0341	-0.0236	0.0011	0.1563	-0.0176	0.0447	-0.043	-0.0092	0.0072	0.0201
d-4	-0.0028	-0.0056	-0.0483	0.1311	0.0916	0.0567	-0.0558	-0.0064	0.0929	-0.0817
d-5	0.0436	-0.0403	0.0203	-0.152	0.07	0.0222	-0.0241	0.0071	-0.091	0.149
d-6	0.0193	-0.0478	-0.0165	-0.0372	0.1539	0.0284	-0.031	0.0097	0.0179	0.0113
d-7	0.0043	-0.0403	0.0178	-0.0718	-0.0062	-0.0208	0.0183	0.0115	-0.027	0.0624
d-8	-0.0336	0.0191	-0.0759	0.0565	-0.0037	0.0227	-0.0236	0.0031	0.0664	-0.0608
d-9	0.0731	0.0033	-0.0141	-0.0485	-0.0225	0.0059	-0.0057	-0.0014	0.0334	-0.0409
d-10	0.0059	0.0538	-0.2043	-0.2009	-0.1289	-0.0744	0.0781	-0.0129	-0.0479	-0.0717

Note: All variables found within the 2009 RECS survey responses, data, and codebook.
educ-n: found as education,
own-n: found as kownrent,
race-n: found as householder_race,
d-n: found as division.

	race-3	race-4	race-5	race-6	race-7	typehuq-1	typehuq-2	typehuq-3	typehuq-4	typehuq-5
btuwood										
usewood										
tothsqft										
hdd65										
hdd30yr										
yearbuilt										
hhage										
educ-1										
educ-2										
educ-3										
educ-4										
educ-5										
educ-6										
educ-7										
educ-8										
wdpellet										
woodlogs										
drafty-1										
drafty-2										
drafty-3										
drafty-4										
moneypy										
totaldol										
totalbtu										
ur										
own-1										
own-2										
own-3										
race-1										
race-2										
race-3	1									
race-4	-0.019	1								
race-5	-0.0055	-0.0114	1							
race-6	-0.0128	-0.0264	-0.0077	1						
race-7	-0.0115	-0.0237	-0.0069	-0.0159	1					
typehuq-1	0.0172	-0.0408	-0.0055	-0.0044	-0.0021	1				
typehuq-2	-0.0347	-0.05	-0.0055	-0.0466	-0.0173	-0.2923	1			
typehuq-3	0.013	0.0238	-0.0052	0.0108	0.0094	-0.061	-0.3807	1		
typehuq-4	0.015	0.0293	-0.0004	0.0305	0.0105	-0.0624	-0.389	-0.0812	1	
typehuq-5	0.0155	0.0501	0.0143	0.0335	0.0095	-0.0942	-0.5874	-0.1227	-0.1253	1
d-1	-0.0018	-0.0105	-0.0113	0.0251	0.0336	-0.0374	-0.0457	-0.0025	0.1024	0.0082
d-2	-0.0253	0.0122	-0.0064	0.04	-0.0263	-0.0351	-0.1126	0.0569	0.0699	0.0757
d-3	-0.0103	-0.0244	-0.0138	-0.0152	-0.01	-0.0334	0.058	-0.0278	-0.0107	-0.0293
d-4	0.0115	-0.0525	-0.015	-0.0192	0.0145	-0.0101	0.0851	-0.0171	-0.0491	-0.0576
d-5	-0.0055	-0.0423	-0.0201	-0.0036	-0.0191	0.0642	0.0011	-0.0036	-0.0417	-0.0049
d-6	-0.0222	-0.034	-0.0068	-0.0021	-0.0148	0.0228	0.05	-0.0277	-0.0242	-0.0409
d-7	0.0167	-0.028	-0.0194	-0.0177	-0.0216	0.0012	0.0147	-0.0184	-0.0157	0.0047
d-8	-0.0095	-0.0319	-0.0036	-0.0093	0.014	-0.002	0.0144	0.0374	-0.0233	-0.0274
d-9	0.0033	-0.0048	0.015	-0.0089	0.0075	0.0526	0.006	-0.013	-0.0274	-0.0084
d-10	0.0281	0.1688	0.0732	0.0064	0.0277	-0.0125	-0.0496	0.0138	0.0167	0.05

Note: All variables found within the 2009 RECS survey responses, data, and codebook.
educ-n: found as education,
own-n: found as kownrent,
race-n: found as householder_race,
d-n: found as division.

	d-1	d-2	d-3	d-4	d-5	d-6	d-7	d-8	d-9	d-10
btuwood										
usewood										
tothsqft										
hdd65										
hdd30yr										
yearbuilt										
hhage										
educ-1										
educ-2										
educ-3										
educ-4										
educ-5										
educ-6										
educ-7										
educ-8										
wdpellet										
woodlogs										
drafty-1										
drafty-2										
drafty-3										
drafty-4										
moneypy										
totaldol										
totalbtu										
ur										
own-1										
own-2										
own-3										
race-1										
race-2										
race-3										
race-4										
race-5										
race-6										
race-7										
typehuq-1										
typehuq-2										
typehuq-3										
typehuq-4										
typehuq-5										
d-1	1									
d-2	-0.1019	1								
d-3	-0.0941	-0.114	1							
d-4	-0.1171	-0.1418	-0.1309	1						
d-5	-0.1386	-0.1679	-0.155	-0.1929	1					
d-6	-0.0671	-0.0813	-0.075	-0.0934	-0.1106	1				
d-7	-0.0977	-0.1183	-0.1092	-0.1359	-0.1609	-0.0779	1			
d-8	-0.0567	-0.0687	-0.0634	-0.0789	-0.0934	-0.0452	-0.0658	1		
d-9	-0.0513	-0.0622	-0.0574	-0.0714	-0.0846	-0.041	-0.0596	-0.0346	1	
d-10	-0.132	-0.1599	-0.1475	-0.1836	-0.2174	-0.1053	-0.1532	-0.089	-0.0805	1

Note: All variables found within the 2009 RECS survey responses, data, and codebook.
educ-n: found as education,
own-n: found as kownrent,
race-n: found as householder_race,
d-n: found as division.

APPENDIX C. STATA PROGRAMMING

```
. heckman mbtuwood pelmfp pelmhs pelsfp pelshs logmfp logmhs logsfp logshs nwp inc tothsqft
    hdd65 homeage draft i.rural own primeduc hhage, twostep select(wood = nwp inc c.inc#c.inc
    i.division i.rural own homeage hdd30yr singlefam b7.race)

. predict heckmbtuwood

. probit wood nwp inc c.inc#c.inc i.division i.rural own homeage hdd30yr singlefam b7.race

. predict phat, xb

. gen mills = exp(-.5*phat^2)/(sqrt(2*_pi)*normprob(phat))

. histogram mills

. reg mbtuwood pelmfp pelmhs pelsfp pelshs logmfp logmhs logsfp logshs nwp inc tothsqft hdd65
    homeage draft i.rural own primeduc hhage

. predict regmbtuwood

. tobit mbtuwood pelmfp pelmhs pelsfp pelshs logmfp logmhs logsfp logshs nwp inc tothsqft hdd65
    homeage draft i.rural own primeduc hhage, ll(0)

. summarize heckmbtuwood regmbtuwood

. test

. predict hcndmbtuwood, ycond

. by division: summarize mbtuwood hcndmbtuwood if rural ==0

. by division: summarize mbtuwood hcndmbtuwood if rural ==1

. margins, eyex(inc) atmeans at(inc=(5(5)115)) subpop(if wood==1)

. marginsplot, noci

. margins, eyex(tothsqft) atmeans at(tothsqft=(500(115)3030)) subpop(if wood==1)

. marginsplot, noci

. margins, eyex(nwp) atmeans at(nwp=(5(2)49)) subpop(if wood==1)

. marginsplot, noci

. margins, eyex(hdd65) atmeans at(hdd65=(500(570)13040)) subpop(if wood==1)

. marginsplot, noci
```