THREE ESSAYS ON THE EFFICIENCY AND EQUITY OF ENERGY PRODUCTION AND CONSUMPTION IN THE UNITED STATES

by

AMANDA JEAN HARKER STEELE

(Under the Direction of John C. Bergstrom)

ABSTRACT

Since the turn of the century, the way energy is produced and consumed across the United States has changed substantially. The objective of this dissertation is to provide an overview of the economic efficiency and distributional equity implications of changing the way energy is both produced and consumed. This work is outlined in three separate essays. The first essay entitled, "Gone with The Wind: Understanding the Impact of Intermittent Renewable Resources on Power System Reliability Across the United States," investigates if increasing the capacity of electricity generated by intermittent renewable resources affects the ability of electric utility companies across the United States to provide a reliable supply of power to their customers. Results suggest as the capacity of electricity generated by intermittent renewable resources increases, customers can expect to experience longer power system outages. The second essay, entitled "Estimating and Comparing Empirical Measures of Household Energy Insecurity," examines the extent to which different classification procedures used for identifying energy insecure households provide an accurate representation of what it means to be energy insecure in the United States. In this essay we compare and contrast five different approaches used for measuring household energy insecurity and propose one of these as a unique, conceptually and empirically strong, and preferred measure.

Across the different measures, we find between 9 to 22% of households living in the U.S. identify as energy insecure. The third essay entitled, "Examining the Theoretical and Empirical Relationships Between Household Energy Efficiency and Security," develops a theoretical model and empirical procedure for examining how investing in energy efficiency affects household energy insecurity. We use our preferred measure of energy insecurity developed in the second essay to test our hypothesis that investing in energy efficiency has a significant negative effect on a household self-identifying as being energy insecure. The results of this essay indicate investing in energy efficiency decreases the likelihood that a household will become more energy insecure. This dissertation concludes in the last chapter with a brief summary and general conclusions from this research.

INDEX WORDS: Efficiency, Equity, Electricity, Power System Outages, Energy Insecurity, Energy Efficiency, Household Production, Stochastic Production Frontier, State Contingent Production

THREE ESSAYS ON THE EFFICIENCY AND EQUITY OF ENERGY PRODUCTION AND CONSUMPTION IN THE UNITED STATES

by

AMANDA JEAN HARKER STEELE

BS, West Virginia University, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

THREE ESSAYS ON THE EFFICIENCY AND EQUITY OF ENERGY PRODUCTION AND CONSUMPTION IN THE UNITED STATES

by

AMANDA JEAN HARKER STEELE

Major Professor:	John C. Bergstrom
Committee:	J. Wesley Burnett
	Gregory Colson
	Craig E. Landry

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

DEDICATION

To Bert and Toonces, for their endless love and support. To my grandparents KJ and POB, for whom my entire life would have been different without. Missing you always.

ACKNOWLEDGEMENTS

Funding support for this research was provided by the National Institute of Food and Agriculture/United States Department of Agriculture (USDA) through grant award 2013-38420-20521 entitled, "Preparing Students for Leadership in Natural Resources and Economics of Alternative Energy: A Cooperative Approach.” All opinions, findings, conclusions, and recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the National Institute of Food and Agriculture/USDA.

First, I would like thank my advisory committee for their support and advice throughout my entire graduate study. Each of you have encouraged me to think critically about the science of economics and helped me to become a better researcher. Thank you for answering my countless emails, writing numerous recommendation letters for me, and helping me navigate my future as an applied economist. You truly are the dream team.

Dr. Bergstrom, it has been such an honor to have worked with you over the past five years. Your sound advice and support have made all the difference. You have continued to push me to be the best version of myself and supported me in evert facet of my professional career. From helping fund my travel, reading each and every draft of my papers, to working up to the deadline with me on many of my numerous tasks, you have been nothing but patient and supportive. Throughout every trial and triumph, I have had during graduate school, you have been there with words of support and encouragement, while continuing to push me to be the best version of myself. You are truly an outstanding advisor, colleague, and someone I consider to be a friend.

Dr. Burnett, without you none of this would have ever happened. If you had not graciously agreed to be my research advisor at WVU, I never would have known about the program at UGA. In fact, I probably would have never even made it to graduate school. Thank you for continuing to support my research ideas, helping me see that I am both knowledgeable and capable, and having faith in me both as your student and as a professional colleague. You have continued to support me throughout graduate school and I will be forever grateful that you decided to be on my committee. I feel honored to have crossed paths with you at WVU and hope that we can continue to work together in the future.

Dr. Colson, you are someone I have always and will continue look up to. You have been honest with me about my research ideas and provided realistic, honest advice. Thank you for supporting my desire to teach and encouraging me to pursue the teaching certificate, while at the same time being honest with me about the amount of work it would require. Thank you for helping with the graduate student association over the years and for making me feel supported through my efforts. Thank you for your help with my many applications, reading over all of my slides, and supporting my research endeavors. You are always there to lend a helping hand and I sincerely appreciate it.

Dr. Landry, I feel very fortunate to have started at UGA when you did. Over the past five years you have helped me become a better researcher, a more critical thinker, and overall a better economist. You have encouraged me and always made me feel more like a junior faculty than a student. You are incredibly insightful and have helped to make my research better overall. I appreciate your endless support, your guidance, and help with answering the tough questions. Thanks for always being honest with me, while still managing to make me feel as if I could achieve my goals.

I would also like to acknowledge the other faculty, staff and administration of the Department of Agricultural and Applied Economics in the College of Agricultural and Environmental Sciences, without whom this would not have been possible. I feel extremely fortunate and lucky to have been in a department that supports graduate students and helps them succeed. In particular, I would like to acknowledge Dr. Shonkwiler and Dr. Dorfman for their assistance in helping to answer my many econometrics questions. Your knowledge on the subjects I presented helped me ensure my approaches were accurate. Your patience with me did not go unnoticed.

Lastly, I would like to thank my husband, Bert for his endless love and overall patience during my graduate school journey. From wiping away my tears to pushing me to be better, you have been my rock through it all. You are truly an amazing human and I feel lucky each and every day to be your wife. I’ve told you this before, but without you I never would have even gone to college or I probably would have quit in the first semester. You helped me to see my potential early on and have been there supporting me ever since. Because of you I have been able to laugh at something every single day. You make me a better version of myself. Thank you for believing in me (and for catching most of my typos).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 RESEARCH MOTIVATION

Since the OPEC oil embargo of the 1970s, the energy policy agenda of the United States has been concentrated on achieving three main goals: (1) ensuring an adequate and secure supply of energy remains available for both current and future consumption; (2) guaranteeing energy prices remain affordable to consumers, especially for low-income and other economically marginalized households; and (3) reducing the environmental impacts associated with producing and consuming energy resources (Yacobucci 2016). In pursuit of these goals, countless public policies and programs have been implemented to "improve" the way energy is produced and consumed across the United States (Yacobucci 2016).

Recently, many of these policies and programs have called for a modification of the energy resource mix to include a greater share of intermittent renewable resources (e.g., wind energy and solar radiation), as well as a reduction in overall energy consumption through increased energy efficiency (Yacobucci 2016). The energy changes described above have implications related to both the efficiency and equity of energy production and consumption in the United States. For example, modifying the energy resource mix to include more intermittent renewable resources and fewer traditional fossil fuels can help to reduce carbon dioxide emissions, but perhaps with tradeoffs in the efficiency of energy production and/or equity of energy consumption. In addition, if traditional fossil fuel resources are still economically feasible to extract and utilize, then phasing them out too quickly for intermittent renewables could potentially lead to lost profits (Hotelling 1931).

Other energy policies including home energy assistance programs have focused on reducing energy insecurity in the United States. Energy insecurity refers to situations where households do not have adequate access to energy services such as adequate home heating and cooling. Household energy insecurity is similar to household food insecurity where households do not have adequate access to food services such as adequate calorie and nutrition intake. As with the transition to the use of more renewable resources in the power grid, household access to adequate energy services has both efficiency and equity considerations and implications related to energy production and consumption.

1.2 ECONOMIC EFFICIENCY AND DISTRIBUTIONAL EQUITY

From a positive ("what is") economics perspective, economists tend to focus on achieving economic efficiency as the leading objective when evaluating any proposed policy change. In general, policies that achieve economic efficiency are those that allocate resources in such a way that it would be impossible to reallocate the same resources to make someone better off without simultaneously making someone else worse off (Bergstrom and Randall 2016). Economic efficiency occurs when desired outputs are produced at minimum cost, while taking individual consumer preferences into account.

As a policy assessment criterion, a presumed advantage of economic efficiency is it does not necessitate equal distribution of society's scarce resources across all individuals and groups (Bhattacharrya 1995). Another major advantage of economic efficiency as a policy assessment criterion is that it prevents wasteful use and allocation of society's scarce resources to economic production and consumption. A disadvantage of this criterion, however, is that it allows economic injury to occur and ignores who gains and who loses (Bergstrom and Randall 2016). Thus, economic efficiency alone may be insufficient to evaluate proposed policy changes in the U.S. energy industry. Instead, other criteria which take into account the distribution of benefits and costs across households, income groups, geographical regions, and states should also be considered (Goulder and Perry 2008).

Examining the equity or "fairness" of policy changes involves considering normative ("what should be") types of questions and issues. Distributional equity is often a concern when a proposed policy change could potentially harm certain individuals or groups. For example, if a policy is pursued that leads to higher energy prices, then low-income households may be unable to afford the expenses necessary to provide adequate household energy services (e.g., maintenance of comfortable indoor air temperatures). [1] From a distributional equity perspective, these households are made "worse off" when policies are adopted that lead to higher energy prices.

1.3 ARE U.S. ENERGY POLICIES EFFICIENT AND EQUITABLE?

Currently, around 63 percent of the energy resources used to generate electricity in the United States are traditional fossil fuels (e.g., coal and natural gas) (United States Energy Information Administration [EIA] 2018). While these types of energy resources outperform other types of energy resources used to produce electricity (e.g., intermittent renewables) in terms of their predictability and reliability, the potential hazardous risks burning fossil fuel resources imposes on the environment are hard to ignore. From an economics perspective, one way to motivate a shift in production away from fossil fuels would be to assign a price to emissions, either through taxes or the implementation of a cap-and-trade program.

[1] We assume economically marginalized households include households whose characteristics and resources render them potentially vulnerable to economic hardship. Examples might include households with differently abled persons, households with low incomes, households aged 60 and over, or households where an adult is unable to provide a sufficient source of income due to economic hardship (e.g., unemployment).

Previous studies have examined the economic implications of fossil fuel emission control policies and conclude that in most cases the benefits outweigh the costs and implementation of such programs results in positive impacts for the environment (Palmer and Burtraw 2005; Palmer et al. 2011; Rocha et al. 2015; Schmalensee and Stavins 2017). However, in isolation, these policies could lead to higher electricity prices for consumers. For example, when a cap-and-trade program goes into effect, the "cap" on emissions lowers from year to year. As a result, electricity prices go up to reflect the increasing scarcity of the "right" to emit pollutants (Morris 2009; Bird et al. 2011; Palmer et al. 2011; Linn and Richardson 2013).

When electricity prices increase, low-income households may struggle to pay their home energy bills. As a result, these households could end up being disconnected from their electric utility service provider and forced to go without electricity for an extended period of time. Therefore, while a cap-and-trade program represents a cost-effective policy solution for reducing emissions, the resulting increase in home energy prices can have adverse, inequitable effects on low-income households (Stone 2015). As a result of these and other equity concerns, U.S. public policy and decision makers have yet to fully embrace a national cap-and-trade program as a first-best means to mitigating emissions (Murray 2015).

Instead, a number of policy initiatives have been proposed to reduce emissions by directly shifting production toward cleaner, more sustainable, intermittent renewable energy resources. Principal among these initiatives are renewable portfolio standards (RPS), which are state-level regulatory mandates that require a minimum amount of electricity to be generated from renewable energy resources (National Renewable Energy Laboratory [NREL] 2015).[2] Currently 29 states

[2] Several cities across the United States have also committed to having a greater share of their electricity generated by renewable energy resources. Examples include the city of Atlanta, GA who has committed to have 100% of its electricity supplied by renewables by 2035, the city of Blacksburg, VA who has committed to 100% renewables by 2050, and Denver, CO who plans to be 100% renewable by 2030 (Sierra Club 2019).

across the U.S. have a standing RPS in place (EIA 2012). To meet the requirements of their state's RPS, electric utility companies can choose to either generate electricity from renewable energy resources themselves, or purchase renewable energy credits (RECs) from other utilities.

Palmer and Burtraw (2005) examined the economic impacts of state-level RPS at a time when only six states had active policies in place. [3] Their results suggest that because fossil fuel generation declines over time as the level of the state RPS increases, emissions reductions targets are achieved. However, they find that implementation is often accompanied by an increase in the price of electricity for consumers, which as discussed earlier can result in inequitable outcomes across different household groups.

A more recent report by Wiser et al. (2016) however, suggests the opposite. Analyzing RPS in place up to the year 2013, Wiser et al. (2016) finds compliance obligations are not met with electricity price increases. Instead, Wiser et al. (2016) suggests that as more renewable energy resources are being brought online, the supply curve for electric power is likely shifting outward. This outward shift in the supply curve results in lower wholesale electricity prices overall. If wholesale prices decrease, then so too will retail electricity prices overtime.

As Wiser at al. (2016) also points out while decreasing electricity prices are good for consumers, it is important to understand that these cost savings likely come at the expense of the producer (e.g. owners and other shareholders who experience reduced revenues when electricity prices go down). Thus, any reduction in the wholesale price of electricity does not necessarily represent an overall gain in economic welfare, but rather a transfer of wealth from the producer to

[3] In 2004, only Connecticut, Maine, Nevada, Massachusetts, New Jersey, and California had RPS policies in place (Palmer and Burtraw 2005).

the consumer (Felder 2011). This wealth redistribution again raises a concern over the equity implications of such policies.

In addition to motivating a shift in production toward renewable energy resources, many public policies and programs have called for increased investment in energy efficiency to reduce harmful pollutants and emissions. Because energy efficiency investments enable households to provide energy services while consuming fewer units of energy, they decrease demand and lower household fuel expenditures. Thus, from a theoretical perspective, households should invest in energy efficiency as long as the marginal benefits of the investment outweigh the marginal costs. Investments in energy efficiency are also advocated for because they can help to displace generation from traditional fossil fuels and help to lower operation and maintenance costs for utilities by preventing the need to invest in additional operating capacity (Environmental Protection Agency [EPA] 2018).

As stated earlier, when a customer invests in energy efficiency such as an *Energy Star*® appliance, they consume fewer units of energy to produce household energy services. If consumers require fewer units of energy to produce the same level of household energy services, then utility companies, especially those who face consumer price regulation, can expect profits to decline. Especially, if their lower operating and maintenance costs are more than offset by their decreased revenue from customers consuming fewer units of energy. If profits are declining, then utility companies are becoming worse off as a result of the household's decision to invest in energy efficiency. These distributional effects again raise concerns about the equity implications of investing in energy efficiency as a means to achieving the three main goals of the U.S. energy policy agenda listed in Section 1.1.

1.4 OVERALL AND SPECIFIC RESEARCH OBJECTIVES

The overall objective of this dissertation is to better understand, both theoretically and empirically, how firms and households are affected by changes in energy production and consumption from economic efficiency and distributional equity perspectives. The specific research objectives of this dissertation corresponding to the three essays contained therein, are as follows:

1. Examine how modifying the energy resource mix to include a greater share of intermittent renewable resources (i.e., wind energy and solar radiation) affects the efficiency and reliability of energy production and delivery to consumers.
2. Develop and compare alternative measures of household energy insecurity as indicators of the distributional equity of energy consumption.
3. Examine theoretically and empirically how household investments in energy efficiency influences a household's level of energy insecurity.

1.5 STRUCTURE OF THE DISSERTATION

The remainder of this dissertation is structured as follows. In Chapter 2 (Essay 1) entitled, "Gone with the Wind: The Unintended Consequences of Increasing the Capacity of Intermittent Renewable Resources Used for Electricity Generation," we investigate how increasing the capacity of intermittent renewable resources (e.g., wind energy and solar radiation) impacts electric system reliability. We model disruptions in reliability of service using a state-contingent production function approach. Using data from the U.S. Energy Information Administration (EIA), we examine whether or not the average frequency and/or duration of disturbances has increased over time as more intermittent renewable resources have been brought online.

In Chapter 3 (Essay 2) entitled, "Estimating and Comparing Empirical Measures of Household Energy Insecurity," we compare and contrast five different approaches for measuring the extent and severity of energy insecurity. These five measures all measure household energy insecurity on discrete scales that are designed to represent the full extent and severity of energy insecurity being experienced by households living in the United States. Our overall objective in this chapter (essay) is to identify an energy insecurity measure that can be universally applied and produces consistent and accurate measurements of the experience of being energy insecure. We use household responses to the 2015 Residential Energy Consumption Survey (RECS) to construct the different index measures, and then compare the validity of the different index measure results using three separate validity tests.

In Chapter 4 (Essay 3) entitled, “Examining the Theoretical and Empirical Relationships Between Household Energy Efficiency and Security,” we develop a theoretical model and empirical procedure for examining how improvements in household energy efficiency affect the presence of household energy security/insecurity. Our theoretical approach relies on the theory of household production, which allows us to capture a household’s demand for and production of energy services. We then utilize a stochastic production frontier approach to explain why households who are inefficient in their production of household energy services might choose to invest in energy efficiency. We explain how the return to such an investment could lead to higher felt-levels of household energy security. Using the index measure identified in Chapter 3 (Essay 2) we empirically examine how making energy efficiency improvements in the home affects the presence of household energy insecurity. The key findings, general conclusions, and policy implications from all three essays are discussed in the concluding chapter (Chapter 5) of the dissertation.

CHAPTER 2

ESSAY 1: GONE WITH THE WIND: THE IMPACT OF INTERMITTENT RENEWABLE RESOURCES ON POWER SYSTEM RELIABILITY ACROSS THE UNITED STATES*

* Harker Steele, Amanda. To be submitted to *Energy Policy*.
This chapter has benefited from comments received at the 2017 AERE Annual Meetings and 2017 Columbia University Sustainable Development Workshop.

ABSTRACT

Using intermittent renewable resources (e.g., wind and solar) to generate electricity has long been cited as one potential solution to reduce carbon dioxide and other harmful greenhouse gas emissions. However, given their sporadic nature, electricity generation using intermittent renewable resources raises concerns about reliability of power production and delivery. This chapter provides a theoretical and empirical examination of how increasing the capacity of intermittent renewable resources impacts power system reliability across the United States. Our theoretical motivation relies on a state-contingent production function approach, which allows us to account for the uncertainty presented by the use of intermittent renewable resources. Our empirical analysis focuses on end-user interruptions as measured by the System Average Interruption Duration Index and the System Average Interruption Frequency Index. Overall, our results suggest increasing the capacity of intermittent renewable resources used to generate electricity, on average leads to longer power disturbances for end-consumers. However, at the margin, these effects are relatively small, indicating a significant amount of intermittent renewable resources would need to be added to the grid before large-scale power system interruptions are realized. Nevertheless, given the current policy landscape, the capacity of electricity generated by intermittent renewable resources is likely and expected to increase. Therefore, to provide some insight on the policy implications of our results, we use one set of our estimation results to forecast interruptions in power system reliability and the associated costs under different renewable energy policy scenarios.

Keywords: Electricity Reliability, Intermittent Renewables, Reliability Metrics, Power System Interruptions, Costs

2.1 INTRODUCTION

Moving away from a carbon-intensive electric utility industry continues to remain a key policy objective for many public and private decision makers across the United States. Since the early 1900s, electricity in the United States (U.S.) has been predominantly supplied by conventional fossil fuels including coal, oil, and natural gas. Although these fossil fuels still dominate the market, there has been much public concern about the negative environmental impacts associated with utilizing these resources. As a result of this concern, interest has been expressed in modifying the energy resource mix used to generate electricity across the U.S. to include a greater share of intermittent renewable resources (e.g., wind and solar energy) (Energy Information Administration [EIA] 2017).

Transforming the electric utility sector to include more intermittent renewables has raised some key, fundamental questions within the electric utility industry. Perhaps the most important of which is: "How will increasing the capacity of intermittent renewable resources impact power system reliability?" The overall objective of this chapter (essay) is to help answer this question and contribute to helping to fill an important information gap in the literature. More specifically, drawing from economics and engineering literature, we provide a theoretical and empirical analysis of the effects of expanding the capacity of intermittent renewable resources used to generate electricity on the ability of electric utility companies across the United States to provide a consistent and reliable supply of electricity to end-consumers.

Currently, there are many different definitions of electrical system reliability available in the literature. [4] For the purposes of this study, we assume electrical system reliability can be

[4] According to North American Electric Reliability Corporation (NERC), a "reliable" power system is one that ensures "adequacy of supply" and "operational reliability." Ensuring adequacy of supply involves being able to provide the electricity required to meet the energy requirements of all electricity consumers at all times, considering both scheduled and unscheduled power outages (2013). A power system is operationally reliable if it is able to withstand sudden interruptions from an unanticipated loss of system

defined as the ability of the electrical grid generating system and its components to provide a consistent, steady, uninterrupted supply of power to end-consumers. Following this definition, our empirical analysis on power system reliability focuses on end-user interruptions as measured by two indices, namely the System Average Interruption Duration Index (*SAIDI*) and the System Average Interruption Frequency Index (*SAIFI*). [5]

Many prior published examinations of U.S. electric system reliability have been motivated by attempts to uncover time-trends in bulk power system interruptions, which have implications for public policy and investment decisions surrounding the revitalization of the U.S. electrical grid (Eto and LaCommare 2008; Hines et al. 2009; Larsen et al. 2015, 2016). Past research findings have suggested, even as technology advances, the frequency of adverse system interruptions (e.g., blackouts and large-scale brownouts) has not decreased over time. Moreover, the majority of these interruptions, when and if they do occur, are happening within the utility's distribution system; which indicates it may be worthwhile to focus on outages experienced by end-consumers when addressing grid reliability concerns (Eto et al. 2012).

In its current state, electricity along the electrical power grid of the United States only flows in one direction, from central generation to the end-consumer (U.S. Department of Energy 2015). [6] As a result, any disruption that occurs at any point along the grid (e.g., during the generation, transmission, or distribution process) can impact the ability of customers to receive an uninterrupted supply of electricity. And while there is an agreement among public and private decision makers alike that the electrical grid generating system needs to be revitalized, investments

---

components. Burtraw et al. (2013) define reliability more generally, as the potential to account for disruptions in the supply of electricity, including shortages in generation capacity and available reserve margins, or an inability of an electrical utility to meet the reliability standards currently in place.

[5] Both SAIDI and SAIFI are indices constructed from outage information reported by customers of electric utilities. These indices are discussed in more detail later on in the paper under Section 3. Measuring Reliability of Service.

[6] For further clarification see Figure 2. A1. Three Phases of the Electricity Production Process in the list of figures at the end of the paper.

thus far to improve and modernize the grid have been insufficient (Campbell 2018). Moreover, what is needed are not just a few additional distribution lines, but rather a complete modification of the grid, including the addition of advanced infrastructure that is able to accommodate new technologies such as rooftop solar and electric vehicles (O'Connor 2018).

A specific issue that makes increasing the capacity of intermittent renewable resources used to generate electricity difficult is the lack of economically viable, commercial storage along the grid. Unlike traditional fossil fuels, the availability of intermittent renewable resources depends on the daily weather. Thus, intermittent renewable resources are inherently sporadic and without available storage capacity, their forecasted output must be continuously updated and significant errors can occur (Delaure and Morris 2015). Lastly, because the bulk of intermittent renewable resources are located in rural regions in the central United States, away from major population centers where demand is normally the highest, new transmission infrastructure is needed to ensure supply from intermittent renewables can instantaneously meet demand (Crabtree et al. 2011).

Nevertheless, calls from both public and private decision makers for increased capacity in the power grid from intermittent renewable resources continue to persist – and utility companies all across the United States are responding. For example, in 2017 wind energy accounted for close to 6% of total U.S. electricity generation and approximately 37% of the total electricity generated from renewable energy resources overall (EIA 2018). Additionally, according to the Solar Energy Industries Association (SEIA), solar energy was the number one resource of additional capacity added to the grid in the first quarter (January to March) of 2018, surpassing all other types of energy resources (both renewable and non-renewable).

However, concerns over how grid reliability will be affected from increasing the capacity of electricity generated by intermittent renewable resources across the United States has yet to be adequately addressed in the economics literature. In fact, the majority of studies that have examined the reliability implications of transitioning to an electrical power system that relies more heavily on intermittent renewable resources to generate capacity have been focused on areas outside of the United States. Examples include studies of grid reliability in Germany (Schmid, Pahle, and Knopf 2013; Abrell and Kunz 2015; Abrell and Rausch 2016), Canada (Zaidi 2007), China (Chang and Wu 2011), and France and Great Britain (Pean, Pirouti, and Qadrdan 2016).

Hibbard, Tierney, and Franklin (2017) examined how the reliability of the U.S. electrical power grid has been affected by policy and market changes that have resulted in decreasing financial viability of conventional power plant technologies (e.g., power plants fueled by coal). They find no supporting evidence that reliability has been negatively affected by the replacement of coal-fired power plants. Instead, they suggest that as conventional, but no longer economically profitable coal-fired power plants/generating units are being phased out and replaced by natural gas turbines and other intermittent energy resources, services to ensure power system reliability are likely becoming more prevalent.[7]

Some previous research has been done to investigate issues beyond electricity sector operations that can negatively influence power grid reliability across the United States. More specifically, recent work by Pless and Fell (2017) examines how bribes (i.e., informal payments made by firms to utilities to obtain kilowatt-hours of electricity) might negatively influence electrical system reliability. The idea behind their analysis is that a bribe to secure a connection to the electrical grid often goes unseen by the capacity resource planner. This unseen or stealth bribe can result in an inadequate supply of electricity on the grid to meet demand, which would result in

[7] Examples of reliability services include frequency and voltage management, increasing ramping and load following capabilities, ensuring the provision of contingency and replacement reserves, and having black start capability (Hibbard, Tierney, and Franklin 2017).

more power system interruptions being experienced by end-consumers. Using five years of bribery and electricity reliability data, Pless and Fell (2017) find stealth bribes for electricity connections lead to significant increases in monthly power outages for all consumers.

There has also been some work done to address reliability concerns brought about by U.S. public policy initiatives designed to reduce the amount of greenhouse gas (GHG) emissions from electricity power plants. For example, Burtraw et al. (2013) examine the impacts of the Cross-State Air Pollution Rule (CSAPR) and the Mercury and Air Toxins Standards (MATS) on electrical system reliability in the United States. Using the Haiku electricity market simulation model developed by Resources for the Future (RFF), Burtraw et al. (2013) find that instead of leading to changes in the energy resource mix used to generate electricity, both regulations lead to investments in pollution control technology, which provide the opportunity for utilities to reduce emissions without necessarily changing inputs.

The findings of Burtaw et al. (2013) raise the question, "What happens when the energy resource inputs used to generate electricity are changed to include a greater share of intermittent renewable resources?" This fundamental question has been considered in the engineering literature for some time now. A recent example includes Wangdee (2014) who uses a systems well-being-analysis framework to investigate the effect of adding wind capacity to a generating system that has historically relied on traditional fossil fuels. Wangdee (2014) finds that grid generating systems that utilize wind resources are more likely to encounter loss of load situations as compared to grid generating systems that do not utilize wind resources and instead depend on traditional fossil fuels.

Wangdee's (2014) main analysis is theoretical in nature, which illustrates a gap in the literature of empirical studies that examine relationships between grid system reliability and increasing the capacity of electricity generated by intermittent renewable resources. In order to help to fill this gap in the literature, this paper reports on a theoretical and empirical analysis which shows that increasing the capacity of intermittent renewable resources has a positive and significant effect on the duration of power outages experienced by end-consumers.

In the next section of this paper (Section 2.2) we provide some background context for our analysis by briefly outlining the different types of energy resources used to generate electricity across the United States, distinguishing between the characteristics of traditional fossil fuels and intermittent renewable resources. In Section 2.3 we discuss the electricity generation (i.e., production) process in detail, focusing on how power system operators make decisions about which generating units to bring online and when. Building on Sections 2.2 and 2.3, in Section 2.4 we employ a state-contingent production function approach to examine how the uncertainty surrounding the use of intermittent renewable resources might influence a power system operator's decision to bring those resources online, and how the decision to bring them online could lead to disruptions in power system reliability.

The theoretical model in Section 2.4 provides motivation for our empirical analysis starting in Section 2.5 where we provide a more detailed explanation of how we empirically measure reliability of service focusing on end-user interruptions. In Section 2.6 we describe the data set used for our empirical analysis and provide an outline and explanation of our empirical model specifications. In Section 2.7 we discuss the results of our analysis in detail along with policy implications, focusing on projected service interruptions and associated costs resulting from implementation of different renewable energy resource standards across the United States. In Section 2.8 we provide a brief summary of our study and offer some conclusions including needed future research.

2.2 ENERGY RESOURCES USED TO PRODUCE ELECTRICITY

Currently, there are inconsistencies in the terminology used to distinguish between the types of energy resources used to generate electricity. From an economics perspective, energy resources used to produce electricity can be divided into two main categories: fund resources and flow resources. Fund resources include energy resources that exist as a given fixed stock, both in terms of quality and quantity (Bergstrom and Randall 2016). Fund resources can be further categorized as being either exhaustible, non-renewable resources or exhaustible, renewable resources.

Exhaustible, non-renewable resources are resources whose supply is depletable, or in other words, is not able to be renewed within a practical planning horizon (Bergstrom and Randall 2016). In terms of the energy resources used to generate electricity in the United States, exhaustible, non-renewable fund resources include traditional fossil fuels such as coal, oil, and natural gas. Conversely, an exhaustible, renewable fund resource is one in which the supply is able to be depleted but also renewed within a practical, designated planning horizon (Bergstrom and Randall 2016).

An example of an energy resource used to generate electricity that is exhaustible and renewable would be woody biomass or bio-fuels. The natural commodities (e.g., timber and maze) used to produce these energy resources can be extracted. However, due to natural ecosystem processes and functions and in some cases human intervention (e.g., farming, planting, cultivating) they can also be replenished (Bergstrom and Randall 2016). In contrast, to exhaustible, renewable and non-renewable fund resources, flow resources are resources that exist as a continuous stream, also with given quantity and quality dimensions (Bergstrom and Randall 2016).

Flow resources can be further categorized as being either storable or non-storable. In terms of electricity generation, because there is currently no economic solution to commercial energy storage, the term "storable" is often used synonymously with the term "dispatchable." [8] The dispatchability of an energy resource refers to the energy resource's ability be brought online at any time or stored for future use. [9] Dispatchable (storable), flow resources include energy resources that can be stored for future use and therefore, are always available for electricity production when and if they are needed. In terms of electricity generation, dispatchable (storable), flow resources include renewable energy resources such as hydropower and geothermal energy (Nikoletatos and Tselepis 2015).

Non-dispatchable (non-storable), flow resources are resources that are unable to be stored for future use. Instead, their availability primarily depends on real-time meteorological conditions. Energy resources that are considered to be non-dispatchable (non-storable) flow resources include wind energy and solar radiation. These types of energy resources are commonly referred to as "variable" or "intermittent" renewable resources (Nikoletatos and Tselepis 2015). [10]

The difficulty associated with integrating intermittent renewable energy resources into the energy resource mix used to generate electricity stems from the fact that the electrical power grid of the United States was originally designed around the concept of large, controllable electrical power system generators (i.e., power plants). These types of power plants can be brought online at any time as needed because the availability of the energy resources inputs used to power them (e.g., coal, natural gas, and nuclear energy) are readily available (Federal Energy Regulatory Commission [FERC] 2015).

---

[8] The term energy storage is typically used to describe technology or devices that store energy for future use. The devices are generally large-scale lithium-ion based batteries. Because storage is not used on a large scale yet, when classifying energy resources used to produce electricity, system operators and utilities refer to resources as being either dispatchable or non-dispatchable

[9] By their nature, all exhaustible, fund resources (both renewable and non-renewable) are dispatchable.

[10] See Figure 2. A2. Energy Resources Used to Produce Electricity for a graphic that explains the difference between the types of energy resources used to generate electricity.

Compared to traditional fossil fuels, the availability of intermittent renewable resources is largely sporadic, making it difficult to accurately predict the contribution these type of resources can make to overall power generation capacity (Skea et al. 2008). Additionally, in its current state, the U.S. electrical grid operates as a one-way street where electrons are passed through the three phases of production sequentially (Gerrity and Lantero 2014). [11] Therefore, any disturbance in the generation and/or transmission phases of the electricity production process can affect the ability of distribution centers to ensure customers are able to receive an uninterrupted supply of electricity.

Furthermore, while the development of a smart grid [12] has been discussed and considered by public policy and decision makers for some time now, investments thus far have been marginal and mostly only consist of the addition of advanced metering infrastructure used by end-customers (Campbell 2018).[13] Another option available to help control the contribution made by intermittent renewable resources includes the use of energy storage devices (e.g., large and small-scale commercial batteries). Energy storage devices have the potential to revolutionize the way electric grid systems operators meet projected demand because access to storage would allow operators to smooth out the availability of intermittent renewable resources over time. In other words, if storage

---

[11] The electricity production process begins when the energy produced from burning fossil fuels, or harnessing power from the wind, the sun, or a nuclear reaction is converted into kinetic energy to begin the power generation process. The electricity generated by these power plants is then transported through a system of high-voltage transmission lines to a transformer. The transformer operates to lower the voltage so the electricity can be passed off to consumers through the system of distribution power lines (Gerrity and Lantero 2014). For further information on the traditional one-way flow of power along the grid see Figure 2. A1. Three Phases of the Electricity Production Process.

[12] Currently, there is no universally accepted definition of the smart grid. However, in a general sense, the smart grid refers to an interconnected electrical grid system that permits electricity to flow in both directions (e.g., from producer to consumer and vice versa), permits flexible generation and can provide real-time feedback to consumers and producers tracking energy consumption. It is anticipated that when power outages occur on the smart grid, the technology in place will be able to detect and isolate the outage, containing it before it becomes a large-scale brownout or blackout. The smart grid technologies will also make it easier for customers to become net suppliers to the grid (Department of Energy 2015).

[13] Advanced metering infrastructure (AMI) is an integrated system of smart meters, communications networks, and data management systems that enable two-way communication between utilities and customers. These are used by households who supply electricity to the grid through their own solar panels or wind turbines, or whose utility charges time of use pricing and therefore need to be able to remotely measure electricity use. Additionally, AMI provides several important functions, not previously available for utilities and their customers. Examples include the ability to automatically connect and disconnect service, identify and isolate outages, and monitor voltage (Department of Energy 2016).

was an option, then the energy produced by intermittent renewable resources (for example wind energy at night) could be stored in batteries and brought online to be dispatched as needed to meet system demand. However, the batteries currently available are all lithium-ion based batteries, and therefore can only store energy in small quantities for a certain amount of time (Department of Energy 2016; Rathi 2017).

While the complexities of the evolving U.S. power system have implications for this research, a more detailed explanation and assessment of its intricacies is beyond the scope of this chapter (essay). Instead, for the purposes of this chapter (essay), we concentrate on how current operations can contribute to the frequency and duration of power system outages being experienced by end-consumers, assuming the electrical grid generating system still only operates in one direction. In this one-directional operation, power first flows from generation to transmission, and then to distribution and without access to adequate storage technology along the way. In this system, the ability of end-customers to receive an uninterrupted and reliable supply of electricity can be impacted by disturbances in the power generation phase and/or the electricity transmission phase.

2.3 ELECTRICITY PRODUCTION

The electricity production process begins when one of the energy resources discussed previously (i.e., non-renewable and renewable fund resources or storable and non-storable flow resources) is used to spin a turbine that converts the kinetic energy from the energy resource into mechanical energy available for work. The mechanical energy is then converted into electricity via electromagnetic induction. The electricity produced from different generating units moves through a complex system, called the electrical power grid, which consists of electricity substations, transformers, and power lines that are used to connect electricity generators (i.e., producers) to end-consumers (EIA 2018). Electricity must be generated before it can be transported along the grid and distributed to end-consumers.

The first-stage of the electricity production process (i.e., generation) takes place inside a power system generator, also known as generating unit or power plant. Some electric utility companies generate all the electricity they sell to consumers using only the generating units they own and operate. Other utilities purchase their electricity directly from other utilities, different power producers, or sometimes even from the wholesale market (EIA 2018). [14] Generating units (i.e., power plants) are categorized by both the energy resource inputs they consume to produce electricity and their specific operational technology. The costs to operate the unit, which depend on the unit's individual characteristics, determine when, where, and how each generating unit will be brought online to produce electricity and contribute to overall capacity (FERC Primer 2017).

The electricity provided to end-consumers by their electric utility can be produced from multiple different types of generating units. The combination of generating units used to serve customers can also change depending on the type and location of the customer to be served. For example, the electricity provided to one group of customers could be being supplied by a natural gas-fired generating unit while the electricity supplied to a different group of customers, served by the same utility, could be being supplied from a coal-fired generating unit. Utilities often serve customers in multiple states, which can dictate what type and combination of generating units are assigned to serve different customer groups.

For example, First Energy Corp. is an electric utility company based in Akron, Ohio. It services 6 million customers across six different states: Ohio, West Virginia, Pennsylvania,

[14] The electricity production process is different for regulated and unregulated electric utility companies. Regulated electric utilities are normally vertically integrated, meaning the own and operate each phase of their production process (Froger et al. 2016). Unregulated electric utility companies but power from the market to supply to their customers (Froger et al. 2016).

Virginia, Maryland, and New Jersey (First Energy 2018). First Energy Corp. owns and operates its own coal-fired power plants, as well as numerous other hydroelectric power facilities in the states of Virginia, New Jersey, and Pennsylvania (First Energy 2018). The type of power plants used by First Energy Corp. to generate the electricity it provides to its end customers, changes depending on which state the customer resides in. For example, from 2001 to 2017, 93% of the electricity provided to customers who live in West Virginia and are served by First Energy, was supplied by coal-fired generating units (Popovich 2018). In New Jersey however, only 16% of the electricity supplied to end consumers was produced by coal (Popovich 2018).

As stated early, large-scale batteries are currently not economically viable, and as a result commercial storage for electricity does not presently exist along the electrical power grid generating system of the United States. Electricity must therefore be produced in real-time, the instant that it is needed to insure that supply (generation) equals demand (load) at all times. To insure that demand and supply are able to match exactly at every moment throughout the day, in every day of the year, and in every location, an electrical power system operator must determine in advance which generating units will be brought online and when (FERC Primer 2015).

The combination of power plants scheduled to be online by the electrical power system operator provides the solution to the utility's "unit commitment problem." [15] When solving the unit commitment problem, the objective of the electrical power system operator is to determine the least-cost combination or supply order of generating units that should be brought online to meet forecasted demand over a discrete period of time (Yang et al. 2017; Foger et al. 2016; Saravanan et al. 2013).

[15] The solution to the unit commitment problem is considered to be the least cost schedule of generating units that can supply enough power to meet forecasted demand over a pre-designated period of time (Saravanan et al. 2013; Ozturk, Mazumdar, and Norman 2004).

Generating units scheduled to be online can be units that are owned and operated by an individual utility or units that the utility has access to (Villumsen, Clausen, and Pisinger 2011; Blumsack 2018). The electrical power system operator uses mathematical dynamic programing models (e.g., mixed-integer linear programming models) to solve the unit commitment problem (Yang et al. 2017; Foger et al. 2016). The solution to the unit commitment problem considers both the cost to operate each individual unit, the specific operational capabilities of the units chosen, and operational constraints that are present on the system (Saravanan et al. 2013; Ozturk, Mazumdar, and Norman 2004).

Of specific interest to the power system operator is how long it takes to start up and shut-down each generating unit, which is also known as each generating unit's ramp rate (i.e., how quickly its total output can be increased or decreased) and its minimum run time. Some generating units by design usually take longer to start up and shut down. As a result, they are normally scheduled to be online for a longer period of time, as compared to other generating units that do not take as long to start up or shut down (EIA 2012).[16]

For example, a generating unit fueled by coal or nuclear energy takes a significant amount of time to start up and shut down. As a result, coal and nuclear generating units are normally scheduled to stay online for an entire day, or in some cases for multiple days, unless periodic maintenance has been scheduled (EIA 2012). Generating units fueled by coal and nuclear energy are also difficult to ramp up (startup) and ramp down (shut down). Instead, they produce a steady amount of electricity at a near constant rate. Therefore, they are often used as baseload generating

[16] Intermediate generating units fill the gap between baseload demand and peak demand. They are typically fueled by natural gas. Their total output is easier to control than baseload unites and as a result their power levels can be adjusted as needed by power system operators. Peaking units, on the other hand, run very little, sometimes only for a few hours per year. These plants, however, can be brought online within minutes and are often used during summer heat waves or on other occasions when demand surges (Penwell Corp 2019).

units, which are designed to run continuously to meet base load demand (i.e., the minimum level of demand that exists on an electrical power grid at all times) (EIA 2012).

Hydropower and geothermal generating units can also be used as baseload generating units, but only if those resources are available to the power system operator for use. For example, electricity produced using hydropower is less common in Florida and Kansas because these states have low topography, which does not allow for enough force to be produced from the falling water resources which is required to spin turbines in a hydropower facility being used to produce electricity (United States Geological Survey [USGS] 2018). In Idaho, Washington, and Oregon however, hydropower contributes to a greater share of the overall electricity produced because water resources are readily available in these states and their natural topography allows for both enough force to be produced from falling water resources and an opportunity to store water for future use (USGS 2018).

In these states where hydropower can be directly controlled, it is often used for baseload generation. Unlike generating units fueled by coal and nuclear power, generating units fueled by natural gas take a shorter amount of time to start up and shut down. They can also more easily be ramped up or ramped down to meet demand. As a result, they are typically brought online by power system operators as intermediate or peaking units, which are used during times when the power produced by baseload generating units is insufficient to meet customer demand (EIA 2012).

As stated earlier, in terms of electricity production, wind energy and solar radiation are considered to be non-dispatchable (non-storable) flow resources. Therefore, they cannot be completely controlled nor dispatched whenever needed by power system operators. Instead, their availability depends primarily on real-time meteorological conditions. As a result, generating units fueled by intermittent renewable resources (i.e., wind and solar) tend to work best as intermediate units because the resources they depend on are sporadic in nature. Thus, they are unable to be completely relied upon to meet the constant electricity supply needs present during baseload demand or the immediate supply needs present during peak load demand.

If a generating unit (baseload, intermediate, or peaking) is scheduled to be online during a given time period, then that generating unit is said to have been "committed" to produce electricity during the time period under consideration. Once the power system operator has decided which generating units to bring online, it must then decide when exactly each unit will come online during the predesignated time period. Determining when each unit will come online is more commonly known as determining the “order of dispatch,” and involves examining a variety of different economic factors including the cost to operate each individual unit. Typically, a power system operator will choose to bring generating units with lower operating costs online first, and then generating units with higher operating costs online later, unless regulatory constraints exist that prevent the preferred order of dispatch.

The unit commitment problem is solved and the order of dispatch is determined by the power system operator at least one day (24 hours) in advance of the need to meet real-time electricity demand (Blumsack 2018). A troubling consequence is that the dispatch schedule is determined in advance of actual operations, when the future state-of-the-world is not known to the power system operator (Yang et al. 2017). Therefore, the power system operator must take into consideration the uncertainty that exists about the future state-of-the-world when making decisions about which generating units to bring online and when each unit will be dispatched.

The uncertainty involving the future state-of-the-world used to only be a function of potential variations in consumer demand, which due to technological advances can now be forecasted quite effectively by power system operators (Hahn 2009). However, in more recent years, increased grid capacity from intermittent renewable resources has significantly increased, resulting in increased uncertainty over resource availability. Compared to coal, nuclear energy, natural gas, geothermal energy and hydroelectricity, the ability of intermittent renewable resources to supply power to the grid through a generating unit depends largely on real-time (daily) weather conditions. Therefore, the key issue faced by power system operators who decide to bring intermittent renewable resources online, is being able to accurately predict the contribution the resources will be able to make to meet system demand in order to maintain adequate system reliability at all times (Skea et al. 2008).

For example, if wind and/or solar are “committed” to be online but the wind is not blowing or the sun is not shining as expected, then the power scheduled to be provided by these resources does not exist or is diminished. In this case, there is not enough power available along the grid to meet system demand. The lack of power availability along the grid will cause the power system’s voltage to drop below its minimum requirement, which will cause circuit breakers to trip and in some cases shut off in order to prevent equipment damage. When circuit breakers trip and/or shut off, power system outages are experienced by end-consumers (Apogee Interactive 2013).

While not having enough power available to meet demand represents one problem to be solved from increasing the capacity of electricity supplied by intermittent renewables, another problem for power system operators arises when there is too much wind and/or solar energy being produced. One country in particular has experienced first-hand just how increasing the grid capacity of intermittent renewable resources by too much too soon can impact its power generating system. In 2013, in order to address global climate change concerns, Germany had plans in place that by the year 2050; at least 80% of its energy supplied would be from intermittent renewable resources. As a result, Germany began phasing out its nuclear power plants in favor of wind and solar.

However, this sudden shift in the inputs needed for a more diversified, climate change conscious generation resource mix resulted in more than one third of all wind turbines being developed in the country's eastern region (Institute for Energy Research [IER] 2013). During certain time periods with sustained high winds, this large concentration of turbines in one location led to the region producing three to four times the total amount of electricity actually being consumed. This drastic increase in the amount of electricity being supplied by intermittent renewables resulted in massive overloads along the grid which led to massive power outages being experienced by end-consumers.

To prevent power outages (like those in Germany) from being experienced by end-consumers as more intermittent renewable resources are brought online, power system operators must consider how the uncertainty of the availability of the resources could influence its production process (i.e., unit commitment and scheduled order of dispatch). A number of mathematical models have been suggested in the engineering literature including dynamic programming models, integer and mixed integer linear programming models, and decomposition models (Tahanan et al. 2015). For the purposes of this chapter (essay) however, we focus only on the production decision of the utility and its corresponding power system operator using a state-contingent production function approach.

2.4 STATE-CONTINGENT PRODUCTION THEORY

In order to build the theoretical foundation for this chapter, we draw on previous economic theory-based literature on the efficiency of electricity production to model electricity market operations and develop a unique theoretical approach which motivates our empirical analysis. Following the work of Fu, Li, and Shahidephpor (2004) and Moghaddam et al. (2014) we begin by assuming the objective of each electric utility company is to maximize its expected profit. However, each utility is still regulated by the price it can charge to provide electricity to end-consumers (Averch and Johnson 1962).

We model the objective of the electric utility company as follows:

(1) $\max E[\pi_{it}] = E(\sum_{t=1}^{T} \sum_{i=1}^{I} [E_{it} * \bar{p} - C_{it}]),$

where $E[\pi_{it}]$ is the expected profit for each utility i in time period t; $\bar{p}$ represents the fixed, regulated price per kilo-watt hour (kWh) of electricity the utility company can charge end-consumers; C_{it} is the cost to produce electricity the electric utility company; and E_{it} is the quantity of electricity actually produced (e.g., total capacity generated) by each utility i in time period t. [17]We assume electric utility companies produce electricity using a set of inputs, including capital technology inputs, labeled here as K_{it}, labor inputs, labeled here as L_{it}, and energy resource inputs, labeled here as R_{it} according to the following production function:

(2) $E_{it} = E(K_{it}, L_{it}, R_{it}).$

We assume the production function in (2) is well-behaved such that electricity production is continuous; strictly increasing in all inputs and all three inputs are necessary for production to occur.

The energy resource inputs R_{it} used by the electric utility company and its corresponding power-system operator are used to represent a variety of different energy resource inputs including

[17] In a regulated electricity market, electric utility companies own and operate all aspects of their electricity supply chain, including power plants, transmission and distribution systems (EIA 2018). To prevent these utilities from having monopoly power in the market, the price they can charge end-consumers is set in advance by an acting third party and remains fixed. In an unregulated electricity market structure, the price per kWh of electricity supplied is not fixed, but rather determined by interactions between competing parties in the market (EIA 2018). Customers can compare pricing structures of different utilities and decide who to purchase electricity from. Because deregulated electric utility companies do control the generation aspect of their supply, they have little incentive to ensure reliability of service in the generation phase of production.

both exhaustible, non-renewable (e.g., coal, oil, natural gas, and nuclear energy) and exhaustible, renewable fund resources (e.g., woody biomass), as well as dispatchable (storable) flow resources (e.g., geothermal energy and falling water for hydroelectricity) and non-dispatchable (non-storable) flow resources (e.g., solar radiation and wind energy). Therefore, we assume the energy resources used to produce electricity R_{it} can be written as a function of the four different types of energy resources as follows,

(3) $R_{it} = R[\, r_1, r_2, r_3, r_4]$.

We label exhaustible, non-renewable fund resources as r_1, exhaustible renewable fund resources as r_2, dispatchable (storable) flow resources as r_3, and non-dispatchable (non-storable) flow resources as r_4.

Under the proceeding assumption, we can modify the production function in (2) by replacing for R_{it} as follows,

(4) $E_{it} = E(K_{it}, L_{it}, R[\, r_1, r_2, r_3, r_4])$.

To examine how increasing the capacity of non-dispatchable (non-storable) flow resources, r_4 (i.e., intermittent renewable resources) will impact the ability of electric utility companies across the United States to provide a consistent and reliable supply of electricity to end-consumers, we modify the production function in (2) using a state-contingent production function approach, outlined in the literature by Chambers and Quiggin (2000).

Under the state contingent approach, we assume the electric utility company and corresponding power system operator intend to produce only one output, electricity labeled here still as E_{it} using the same inputs as before: $K_{it}, L_{it}, r_1, r_2, r_3, r_4$. However, when solving its unit commitment problem and deciding on its preferred order of dispatch, the power system operator must also consider the probability α_s that a random state of nature s will occur in the future, after the operating decision (solution to the unit commitment problem) has been made.

As stated earlier, the availability of intermittent renewable resources primarily depends on real-time weather conditions. Therefore, when deciding on whether or not to include intermittent renewable resources in the unit order of dispatch (i.e., energy resource mix used to generate electricity), power system operators must also consider the probability that the weather conditions for intermittent renewable resources will be favorable in the future. More specifically, utility company operators have to consider the probability that the intermittent renewable resources included in the unit order of dispatch (i.e., generating units they have scheduled to be online) will be available as expected.

For ease of exposition and without loss of generality, we consider only two states of nature $s \in \{1,2\}$. More specifically, we assume state of nature $s = 1$ represents a state-of-the-world where meteorological conditions for intermittent renewable resources are as expected, while state of nature $s = 2$ represents a state-of-the-world where meteorological conditions for intermittent renewables are not as expected. Thus, depending on which state of nature prevails, electricity produced using intermittent renewable resources will either be as the operator expected for meeting customer demand, or will be more or less than expected which can result in grid problems and power delivery interruptions and outages.

Under the state-contingent production approach, if we assume outputs are independent from one another, in the sense that the electricity produced under state-of-the-world $s = 1$ does not influence and is not influenced by the lack of electricity produced under state-of-the-world $s = 2$, then we can assume the production of electricity is non-joint in inputs (Chambers 1998). Under this scenario, we can model the production of electricity E_{it} as follows,

(5) $E_{it} = E(K_{it}, L_{it}, R[r_1, r_2, r_3, r_4]) \quad s \in \Omega = \{1,2\}$.

Under this specification, at the time when the state of nature reveals itself, the production decision regarding the use of inputs has already been made and therefore the inputs including capital and labor (K_{it}, L_{it}) and energy resource inputs (r_1, r_2, r_3, r_4) have already been committed to and/or scheduled to be online for the production of electricity by each utility i in time period t. That is, the inputs are committed based on forecasts, *ex ante* (e.g., prior to the state of nature being realized) while the total number of kWh of electricity produced is realized *ex post* (e.g., after the state of nature has revealed itself).

Using the state-contingent production approach we can modify the electric utility's objective to maximize expected profit in (1) in the following way:

(6) $\max E[\pi_{its}] = \alpha \cdot [\sum_{t=1}^{T} \sum_{i=1}^{I} (E_{its} * \bar{p} - C_{it})\ ] + (1-\alpha) \cdot [\sum_{t=1}^{T} \sum_{i=1}^{I} (E_{its} * \bar{p} - C_{it})\ ]$,

subject to the following constraint

(7) $E_{its} = E(K_{it}, L_{it}\ r_1, r_2, r_3, r_4) \quad s \in \Omega = \{1,2\}$.

Because we assume there are only two possible states of nature, we remove the s subscript from the probability and assume α represents the probability that state-of-the-world $s = 1$ occurs, and $(1-\alpha)$ represents the probability that state-of-the-world $s = 2$ occurs. Under this framework we can see the cost to produce electricity C_{it} is constant and does not depend on which state of nature s prevails.

However, the total revenue received by each utility i in time period t, which is displayed in the expected profit function (6) as follows, $E_{its} * \bar{p}$, does depend on which state of nature s reveals itself. In the case where the state of nature $s = 1$ reveals itself, the utility can expect to receive a profit of $[\sum_{t=1}^{T} \sum_{i=1}^{I} (E_{its} * \bar{p} - C_{it})\ ]$. However, in the case where state of nature $s = 2$ reveals itself, and the electric utility company has scheduled intermittent renewable resources to be online but conditions for these resources do not match operator expectations, the utility can expect electricity E_{its} will not delivered to end consumers in a reliable manner – for example, due to power interruptions and outages some end-consumers will experience reduced electricity delivery for a certain amount of time.

If a power system interruption or outage occurs and electricity is not delivered to end-consumers, the utility will not receive any revenue from the customers who are affected by the power system outage. Because, in the case of a power system outage, those end-consumers who are affected by the outage do not consume any kWh of electricity or pay price $\bar{p}$ and as a result, $E_{its} * \bar{p} = 0$ for the duration of the outage. However, the electric utility company still faces costs C_{it} even if electricity is not delivered to all end-consumers.

In the case when the electric utility company utilizes wind and/or solar as one of its energy resource inputs and has scheduled those resources to be online but state of nature $s = 2$ reveals itself, a power system outage is expected to occur. In the case of a power system outage, the electricity production process can be modeled as follows:

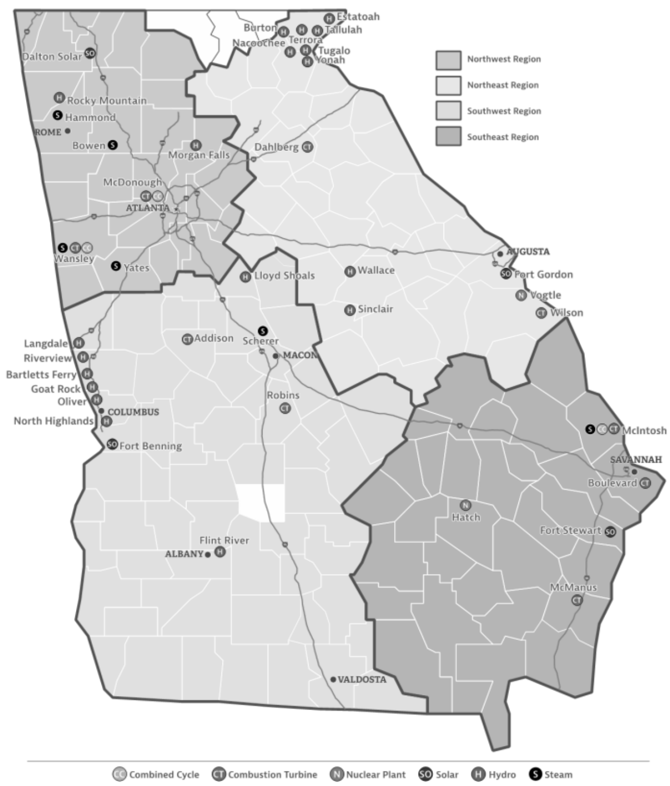
(8) $$E_{it2} = E(K_{it}, L_{it}, R[r_1, r_2, r_3, r_4]) \equiv OUT_{it},$$

where the variable OUT_{it} is used to represent a power system outage. As stated earlier, all scheduled inputs are strictly necessary for production to occur. Therefore, if wind and/or solar are used by the electric utility company and scheduled to be online but the conditions for those resources are not as expected (i.e., the wind is not blowing or the sun is not shining), then the utility can expect to experience an outage in power system reliability (i.e., a power system outage).

In addition to permitting total electrical output (e.g., the number of kWh of electricity produced) to vary under two separate states of the world, the state-contingent production approach outlined above also allows us to consider how the amount electricity delivered to end-consumers who are all serviced by the same electric utility company might be different, even during the same time of day. For example, it could be the case that only some of the customers serviced by an electrical utility company are experiencing a disturbance in power system reliability as a result of the utility's decision to utilize intermittent renewable resources, while others are not. Using the state-contingent production allows us to consider the case where customers serviced by the same utility operate receive two different outputs under the same state-of-the-world.

For example, Georgia Power is the primary electricity provider for the State of Georgia. It provides service to approximately 2.4 million customers in 155 of the 159 counties (Georgia Power 2018). To generate enough power for all of its customers, Georgia Power utilizes a variety of different energy resources, including coal, nuclear energy, natural gas, woody biomass, wind energy, and solar radiation as inputs for the many generating units (i.e., power plants) it operates (Georgia Power 2018). When determining which generating units to bring online to supply electricity, it is plausible that power system operators working at Georgia Power could assign different combinations of generating units (i.e., power plants) to supply power for customers who live in different regions of the state.

Figure 2.1 below provides a map of Georgia Power's network of generating plants.



Source: Georgia Power Generating Plants (www.georgiapower.com)

Figure 2.1 Network of Power Plants Owned and Operated by Georgia Power

For example, it could be the case that only hydroelectric power plants are scheduled to be online to provide electricity to customers who live in the Northeast Region of the state, while a combination of hydro, solar, and nuclear power plants have been scheduled to be online to provide electricity to customers living in the Southeastern part of the state.

If such is the case, then assuming state-of-the-world $s = 2$ reveals itself, the subset of customers who receive electricity from wind or solar energy resources as part of the power supply generating order will likely experience a power system interruption. However, the other group of customers, who only receive electricity from hydroelectric energy sources as part of the power supply generating order will likely not experience a disturbance in power system reliability. As a result, in the case when state-of-the-world $s = 2$ reveals itself, only a portion (i.e., fraction) of the total number of customers serviced by the electric utility will experience a disturbance in power system reliability.

Under this framework, the utility's expected profit function can still be modeled as it is in equation (6). The amount of electricity produced, or lack thereof, will depend on the number of customers serviced by the utility. The framework outlined above is justified by the structure of our data, given that it measures power outages as function of the specific subset of customers served by an individual utility who experienced an outage as a function of the total number of customer served overall by an electric utility company. [18]

Equation (8) represents a theoretical model for a disruption in power system reliability (i.e., a power system outage) being experienced by a subset of end-consumers serviced by an electricity utility when state of nature $s = 2$ reveals itself and the electric utility company has decided *ex ante* to utilize wind and/or solar as one of its primary energy resource inputs. The equation can be empirically estimated using regression analysis techniques, assuming a measure for power system outages as the dependent variable as follows,

(9) $OUT_{it} \equiv E(K_{it}, L_{it}, R[r_1, r_2, r_3, r_4]).$

Here we have replaced the variable E_{it2} in equation (8) with the variable OUT_{it} which as stated earlier is assumed to measure interruptions in power system reliability.

The equation shows that disturbances in power system reliability (i.e., power system outages) OUT_{it} depends on many factors including capital, labor, and energy resources inputs, which are committed *ex ante* by the electrical utility company with the intention of producing electricity before the power system outage occurs. Table 2.1 below shows the theoretical expectation for the estimated sign of the regression coefficient for each of the theoretical variables

[18] The data is described in more detail in the section on Empirical Analysis.

included in equation (9), assuming the electrical utility company is operating in either the first or second stage of production under state-of-the-world $s = 2$ and all other variable inputs are held fixed.

Table 2.1 Theoretical Variables

Variable	Expected Sign of Estimated Regression Coefficient
K_{it}	Positive
L_{it}	Positive
r_1	Negative
r_2	Negative
r_3	Negative
r_4	Positive

Because capital and labor inputs are assigned *ex ante* and outages are defined to be equal to electrical output, theoretically the marginal products of capital and labor inputs are expected to be positive in the first two stages of production, assuming all other inputs remain fixed. Energy resource inputs, r_1, r_2, r_3, r_4 are also assigned *ex ante* via the unit commitment problem. However, the availability of r_4 dictates whether or not a power system outage occurs or does not under state-of-the-world $s = 2$. If scheduled to be online but not available as needed, as r_4 decreases it should have a direct positive effect on power system outages. Conversely, it is expected that other energy resource inputs, r_1, r_2, r_3 will have a negative effect on power system outages, leading to a more reliable power system under state-of-the-world $s = 2$.

2.5 MEASURING RELIABILITY OF SERVICE

For our empirical analysis of electrical power system reliability, the state-contingent theoretical approach discussed above suggests the need for a measure of power system interruptions and outages experienced by end-customers. Focusing on end-customers is also consistent with past research findings suggesting that the majority of power system outages, when and if they do occur, take place within the utility's distribution system, while only a small percentage occur within the bulk power generating system (Eto et al. 2012). Under the state-contingent approach, the electric utility’s (firm’s) expected profit function (equation 6) depends on the amount of electricity produced by a utility that reaches end-consumers as expected. Recall, we assume electrical system reliability can be defined as the ability of the electrical grid generating system and its components to provide a consistent, steady, and uninterrupted supply of power to end-consumers. An empirical measure of electrical power reliability should also account for the number of end-customers who are affected and not affected by power system outages.

Therefore, our empirical analysis focuses on end-user interruptions as measured by two metrics reported by the U.S. Energy Information Agency (EIA). The two metrics we consider include the system average interruption frequency index (*SAIFI*) and the system average interruption duration index (*SAIDI*), both of which provide measures of disruptions in power system reliability that are consistent with the state-contingent theoretical model and criteria. *SAIFI* and *SAIDI* measure end-customer power outages within a utility’s distribution system consistent with power grid and electricity delivery problems caused when wind and solar energy generation does not meet operator expectations (e.g., state-of-the-world $s = 2$ occurs). In addition, these metrics account for the number of end-customers affected by power interruptions and outages as called for the state-contingent profit function. In particular, these metrics measure the specific subset of customers served by an individual utility who experienced an outage as a function of the total number of customer served overall by an electric utility company (firm).

Since 2013, respondents to EIA survey Form EIA-861 have been mandated to report values for *SAIDI* and *SAIFI*, which measure the duration and frequency of power outages experienced by end-consumers respectively. More specifically, *SAIDI* is an index value used to measure the duration of a power interruption for the average utility customer and is calculated as follows:

(10) $SAIDI = \frac{\sum_{i=1}^{n}(d_i * N_i)}{N_t}$.

Here d_i is used to represent the restoration time in minutes (i.e., the duration of the power outage or the amount of time it takes for power to be restored for the customer); N_i is used to denote the number of customers who experienced the power disruptions; and N_t is equal to the total number of customers an individual utility serves during a given time t.

The other measure, *SAIFI* is an index used to measure the frequency of power outages experienced by a utility over a given period (i.e., the number of times a customer goes without power during a year). *SAIFI* is calculated as follows:

(11) $SAIFI = \frac{\sum_{i=1}^{n} N_i}{N_t}$,

where, as before, N_i is equal to the number of customers who experienced a disruption in power, and N_t is equal to the total number of customers an individual utility serves. Larger values of *SAIDI* and *SAIFI* indicate less reliable electricity service (i.e., customers on average, experience longer or more frequent power interruptions), while smaller values of *SAIDI* and *SAIFI* represent a more reliable supply of electricity service is being supplied to end-consumers.

While *SAIDI* and *SAIFI* are both widely recognized metrics used to measure power system reliability, there are differences among the ways utilities define and measure interruptions using these two indices (Eto at al. 2012; Malla 2013). For example, while some follow the IEEE-1366 Standard to measure values for *SAIDI* and *SAIFI*, others use their own set of criteria to measure disturbances using the indices. To control for the differences in criteria used to record and measure outages we include an indicator variable equal to one if the utility used the IEEE 1366 standard to measure values for *SAIDI* and *SAIFI* and zero otherwise in our empirical model specification, which is described in detail below in Section 2.6.

According to Eto et al. (2012), when examining reliability data, outages that occur during major event days (MEDs) should be analyzed separately from outages that do not occur during MEDs.[19] Similarly, Malla (2013) argues that failure to separate out major weather events could degrade the comparability of the indices across different utilities because depending on which region the electric utility operates in, they could naturally be more prone to experiencing major weather events. For example, utilities that operate along the coast are inherently more prone to hurricanes than those who operate inland.

When completing EIA Form-861 utilities are instructed to separate outages based on whether or not they occurred during a MED. Because we are only interested in understanding the impact of using intermittent renewable resources on reliability, we exclude the values for *SAIDI* and *SAIFI* that occur during MEDs. While it is true that extreme weather events will likely contribute to shortages in supply from intermittent renewables, we are more interested in non-severe weather-related events that impact the intermittency of intermittent renewable resources, such as unexpected cloud cover. Separating power outages that occur during MEDs with those that do not occur during allows us to control for weather.

2.6 DATA AND EMPIRICAL METHODS

This section provides an overview of the data, statistical analysis techniques, and model specifications we use to test whether or not increasing the capacity of electricity generated by

[19] A major event day is defined as a day where a power system interruption is likely the result of from some kind of faulty weather-related event (e.g., a lightning strike, snowstorm, hurricane, flood, or other major weather occurrences). During major event days, power outages tend to last longer because multiple people are affected at once.

intermittent renewable resources (e.g., wind and solar) affects disturbances in electrical power system reliability, as measured by *SAIDI* and *SAIFI*.

DATA

To measure power system outages, we collected data from two annual surveys administered by the U.S. Energy Information Administration (EIA). More specifically we obtained utility-level disturbance data from survey Form EIA-861 and operational, power-plant level data from survey Form EIA-923. Because the reliability indices (*SAIDI* and *SAIFI*) we use for our study have only been reported on survey Form-861 since 2013, our dataset spans only five years (2013-2017). The additional data collected from survey Form EIA-861 included: information on whether or not the utility transmitted, distributed, and/or generated its own electricity in each year; the total summer/winter peak load demand measured in megawatts (MW); the total retail sales and total revenue received as measured in US dollars ($); and the number of customers served by each utility in each year.

The data obtained for our study from survey Form EIA-923 included: the net electricity generated, as measured in megawatt-hours (MWh), the primary fuel source used to generate electricity, and the net capacity generated from intermittent resources (e.g., wind energy and solar radiation) measured in MWh by the individual power plants operating under each specific utility. Additionally, from Form EIA-923 we collected information on the North American Electric Reliability Council (NERC) region and census division in which each utility operates a power plant (see Figure 2. A3). Data from survey Form EIA-923 and survey Form EIA-861 were matched by operator *id* and utility *id*. Because different electric utilities collected and retained data for varying numbers of years, our data exists as an unbalanced panel with 276 observations from individual utilities across the five years. The summary statistics are presented below in Table 2.2.

Table 2.2 Summary Statistics EIA Form-861 and EIA Form-923 (2013-2017)

Variable	**Description**	**Mean**	**Std. Dev.**	**Min**	**Max**
Utility	Individual Utilities	12,000.89	7,818.14	84	56,146
SAIDI	System Average Interruption Duration Index	67.67	76.86	0	700.20
SAIFI	System Average Interruption Frequency Index	0.82	1.77	0	43.20
Customers	Number of Customers Served	382,958.20	785,668.90	0	5,517,212
Voltage	Voltage (kV)	24.38	15.98	0	69
Year	Year	2014.86	1.28	2013	2017
Auto	=1 if Utility has an Automated OMS	0.46	0.50	0	1
IEEE	=1 if Utility uses IEEE Standard	0.57	0.50	0	1
Circuits	Number of Distribution Circuits	427.73	712.71	1	4,552
Meters	Number of AMI Meters	178,478.20	678,715.90	0	5,262,080
Summer Peak	Summer Peak Demand (MWh)	2,137.39	3,929.07	1.80	23,858
Winter Peak	Winter Peak Demand (MWh)	1,784.89	3,293.21	1.40	20,541
Retail Sales	Total Electric Retail Sales ($)	9,091,348	16,800,000	0	110,000,000
Revenue	Total Operating Revenue ($1,000)	1,122,687	2,101,950	0	14,500,000
Transmit	=1 Utility Transmits Electricity	0.65	0.48	0	1
Buy Transmit	=1 Utility Buys Transmitted Electricity	0.71	0.45	0	1
Distribute	=1 if Utility Distributes Electricity	1.00	0.06	0	1
Buy Distributed	=1 if Utility Buys Distributed Electricity	0.11	0.31	0	1
Generation	=1 if Utility Generates Electricity	0.91	0.30	0	1
Total Generation	Total Capacity Generated (MWh)	69,600,000	147,000,000	0	984,000,000
Net Generation	Net Capacity Generated (MWh)	7,237,104	15,600,000	-113,055	116,000,000
Renewable	Net Renewable Capacity Generated	98,929.21	646,377.20	0	11,700,000
Renew w/o Hydro	=1 if Utility Uses Wind or Solar as Prime Mover	0.23	0.42	0	1
Renew w/ Hydro	=1 if Utility Uses Hydro as Prime Mover	0.35	0.48	0	1
Observations	937				

EMPIRICAL METHODS

Our empirical analysis followed four main steps. First, consistent with Eto et al. (2012), we transformed both reliability metrics using a log-transformation. We also employed an inverse hyperbolic sine (IHS) transformation of the reliability metrics. Second, we conducted F-tests on both sets of the transformed reliability metrics to determine if accounting for utility-level specific effects was warranted. Third, we conducted a Hausman (1978) specification test to determine rather a fixed effects or random effects specification was more appropriate for our analysis. Fourth, we estimated two sets of models using the transformed values of *SAIDI* and *SAIFI* as the dependent variable. The first set of models included the log-transformed values of *SAIDI* and *SAIFI* as the dependent variable, while the second set of models included the IHS transformed values of *SAIDI* and *SAIFI* as the dependent variable. The raw (untransformed) data for values of *SAIDI* and *SAIFI* are presented below in Figure 2.2 and 2.3.

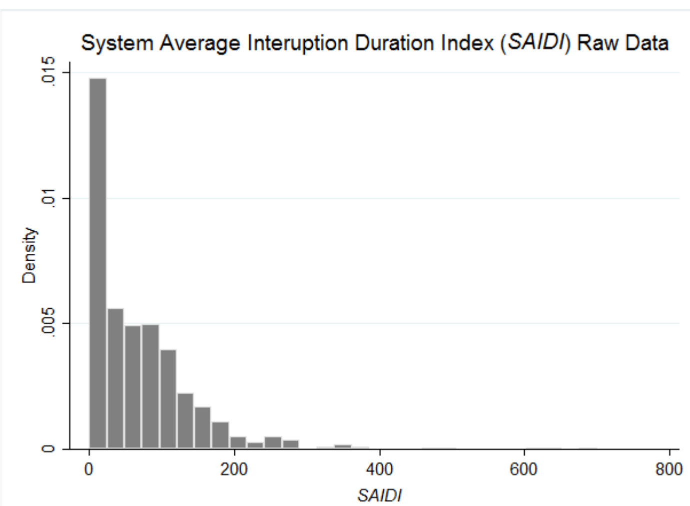


Figure 2.2 Raw Data *SAIDI* (System Average Interruption Duration Index)

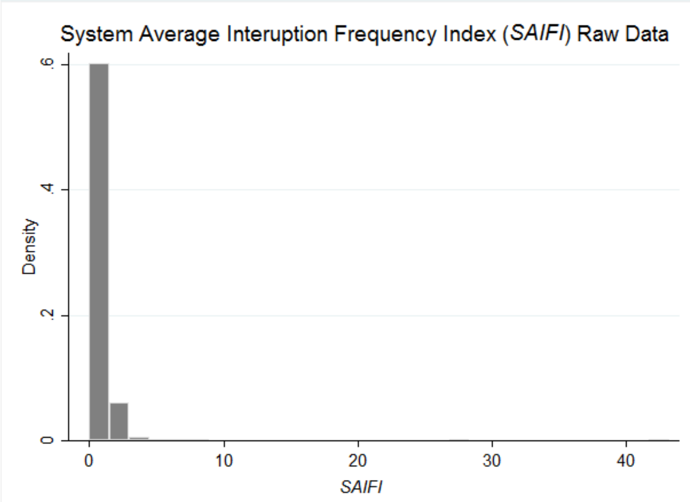


Figure 2.3 Raw Data *SAIFI* (System Average Interruption Frequency Index)

There are two main reasons for applying a log-transformation to the reliability metrics *SAIDI* and *SAIFI*. First, because the metrics themselves tend to follow a log-normal distribution, transforming them into natural logs can result in a relatively normal distribution (Eto et al. 2012; Larsen et al. 2017). Second, using the log-transformed values of the reliability metrics as the dependent variable allows us to interpret the coefficients on our parameters of interest as semi-elasticities. In other words, after multiplying the parameter estimates of interest from the regression equations by 100, we can interpret them as percentage changes in reliability given a one unit change in the variable of interest (Eto et al. 2012; Wooldridge 2015)

However, while the log-transformation is convenient, applying such a transformation does not always work well with data that contain a large number of zeros, as the log transformation of zero is undefined. As suggested by Eto et al. (2012), one way to alleviate this issue is to code all the zeros as ones, and then apply the log-transformation to the reliability metrics. We follow Eto et al. (2012)'s suggestion first and log-transform the reliability metrics after coding the zeros as ones. However, another issue with the log-transformation approach is that it leads to negative values for our reliability metrics, which is problematic as the negative of a power outage is nonsensical. In addition, the estimated skewness and kurtosis values from applying the log-transformation indicated a non-normal distribution of the reliability metrics. The log-transformed values for *SAIDI* and *SAIFI* are presented below is Figures 2.4 and 2.5.

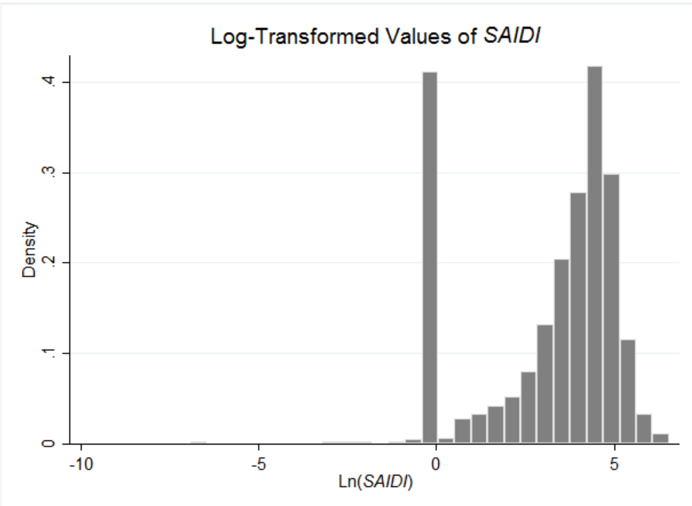


Figure 2.4 Log-Transformed Values *SAIDI* (System Average Interruption Duration Index)

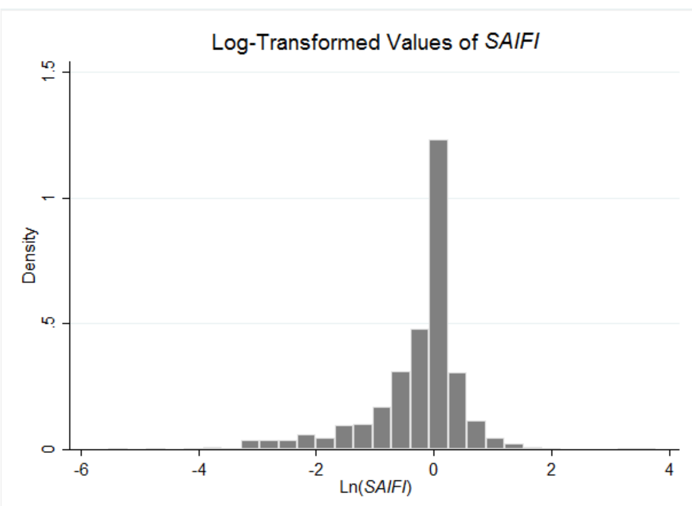


Figure 2.5 Log-Transformed Values *SAIFI* (System Average Interruption Frequency Index)

Therefore, in addition to applying the log-transformation, we also transformed both the values of *SAIDI* and *SAIFI* using an IHS transformation, originally proposed in the literature by Johnson (1949). To apply the inverse hyperbolic sign transformation, we take the reliability metrics *SAIDI* and *SAIFI*, defined here for simplicity as y_{it} and apply the following formula: $\log\left(y_{it} + (y_{it}^2 + 1)^{\frac{1}{2}}\right)$. The HIS transformed values of SAIDI and SAIFI are presented below in Figures 2.6 and 2.7.

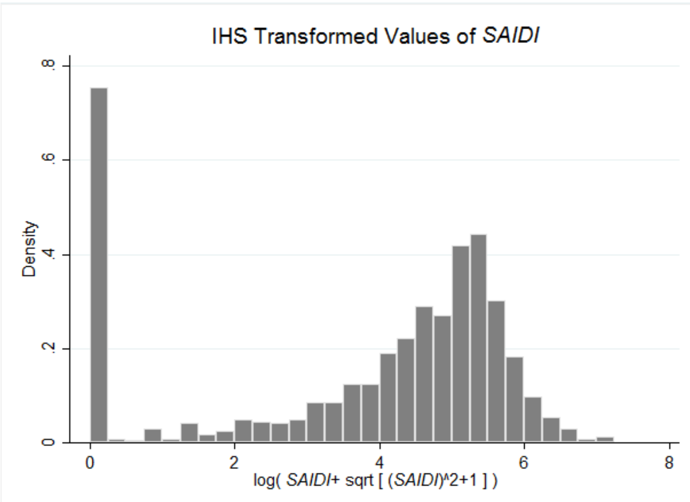


Figure 2.6 IHS Transformed Values SAIDI (System Average Interruption Duration Index)

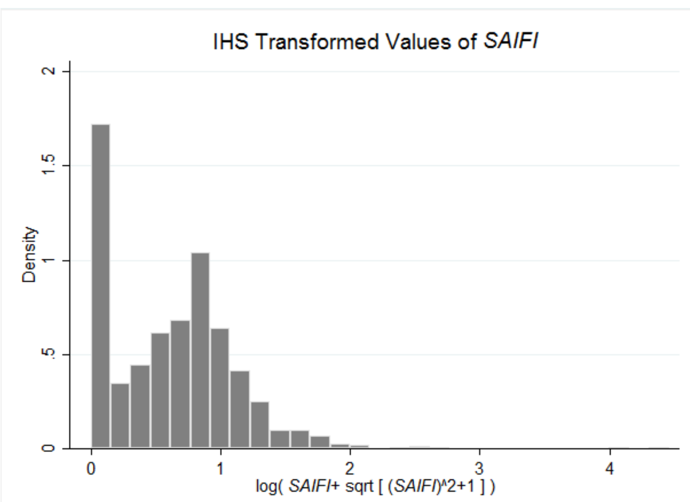


Figure 2.7 IHS Transformed Values SAIFI (System Average Interruption Frequency Index)

Except for very small values of the dependent variable, the IHS transformation is approximately equal to, $\log(2y_{it})$ or $\log(2) + \log(y_{it})$. Therefore, when using the IHS transformed values of the reliability metrics as the dependent variable we can interpret the coefficients on our parameters of interest in exactly the same way as a standard logarithmic dependent variable, that is, the percentage change in the reliability metric given a one unit change in the variable of interest.

Moreover, unlike the log-transformation, the IHS transformation is defined at zero and therefore we do not need to recode the zeros as ones, and do not lose any information. More specifically, given the large presence of zeros in our dataset, applying the IHS transformation allows us to examine what factors might lead no power disturbances. Moreover, after applying the IHS transformation, we find skewness and kurtosis values are within range to assume that *SAIDI* and *SAIFI* are both normally distributed $\left(N(\mu, \sigma^2)\right)$. Therefore, we can estimate the regression equations assuming left censoring at zero.

After applying the log-transformation and IHS transformation to the reliability metrics, we conduct *F*-tests on both sets of the transformed values to confirm that accounting for utility- level specific effects is warranted in our analysis. The *F*-test is a standard statistical analysis technique used to determine if unobserved individual effects have the potential to influence the outcome variable of interest. In the case of the electric utility industry, these unobserved effects might include things such as location, climate region, structure, size, or ownership. If the null hypothesis that no unobserved effects are present can be rejected with a degree of statistical certainty, then estimating fixed-effects and/or random-effects is warranted. The results our *F*-test indicate utility-level specific effects are present and therefore fixed or random effects are appropriate for our analysis.

To determine whether the fixed or random effects approach was more appropriate for our analysis we applied the Hausman (1978) specification test. The Hausman test examines whether, under the null hypothesis, the utility-specific and time-specific effects are uncorrelated with the other regressors of interest included in the model (Wooldridge 2010; Malla 2013). Failure to reject the null hypothesis indicates both specifications are consistent, but that the random effects specification is more efficient than the fixed effects specification. The results of the Hausman test for both sets of transformed *SAIDI* and *SAIFI* values indicate that we fail to reject the null hypothesis in favor of the alternative. Therefore, we estimated the model using a random-effects regression model. [20] Additionally, we estimate the model using a random effects specification because there are time-invariant variables of interest, believed to influence reliability such as ownership type and NERC region operated in.

EMPIRICAL MODEL SPECIFICATION

For our empirical analysis, we use the frequency and duration of power system outages, as measured by the transformed values of *SAIDI* and *SAIFI*, as measures of interruptions in electric utility service reliability. The basic model used to test our hypothesis takes the following form:

(12) $$Outage_{it} = \beta_0 + \beta_1 Intermittent_{it} + \beta X_{it} + \delta Year_{t-1} + c_i + \mu_{it}.$$

In this model, interruptions in electric utility service reliability for each utility i in time period t is measured by the dependent variable, $Outage_{it}$ which is equal to either the natural log or the IHS transformed values of the reliability metrics used (e.g., *SAIDI* and *SAIFI*). We control for year fixed effects by including the variable $Year_{t-1}$, which is used to represent a set of year dummy variables included for all but one of the five years of observation. The unobserved individual-utility-level characteristics, which are believed to influence power system reliability, are denoted by the term c_i, and the idiosyncratic error term is represented here by μ_{it}.

The term X_{it} is used to represent a vector of observable variables that capture operational characteristics of each electric utility i believed to influence power system reliability. More

[20] The random effects specification also allows us to examine the effects of time-invariant variables of interest (Wooldridge 2010).

specifically, the operational characteristics in X_{it} include: $Auto_{it}$, an indicator variable equal to one if the electric utility has an automated outage management system that can be used to automatically detect disturbances; TR_{it} an indicator variable equal to one if the utility transmits electricity; D_{it} is an indicator variable equal to one if the utility operates the distribution lines; G_{it} an indicator variable equal to one if the electric utility identifies as a generator; and $Circuits_{it}$, which is a continuous variable equal to the number of distributional circuits operated by each utility i in time period t.

Consistent with the previous literature, we also control for the size and operations of each electric utility by including additional variables in X_{it}, including $Sales_{it}$, which is equal to the total retail sales of each electric utility (measured in megawatt-hours [MWh]); and $IEEE_{it}$, which is an indicator dummy variable equal to one if the utility uses the IEEE Standard 1366 to record and measure values for *SAIDI* and *SAIFI*; the total number of customers serviced by each utility i in a given year t, labeled as $Customers_{it}$ (Fenrick and Getachew 2012; Malla 2013).[21] To account for differences in the way reliability standards are monitored and enforced, we include a set of indicator variable for all but one of the ten North American Electrical Reliability (NERC) regions individual utilities can operate in, labeled here as $NERC_{it}$. Additionally, because different types of utilities have different procedures for measuring and ensuring the reliability of service using the indices *SAIDI* and *SAIDI*, we include an indicator variable for each ownership type excluding one, labeled here as $ownership_{it}$.

Lastly, because the primary focus of this study is to investigate the relationship between using intermittent renewables resources and power system reliability, we include two key variables

[21] According to Fenrick and Getachew (2012) and Malla (2013) utilities serving fewer customers can be at a disadvantage when measuring the reliability of service using SAIDI and SAIFI values because they have fewer customers overall.

to account the use of intermittent renewable resources by electric utility companies in our sample. First, we include the variable $Renewable_{it}$, an indicator variable equal to one if the utility identifies either wind or solar (photovoltaics) as the prime mover for at least one of the power plants they operate and zero otherwise. [22] Second, we include the variable $Intermittent_{it}$, a continuous variable equal to the net capacity of electricity generated (as measured in MWh) by intermittent renewable resources by each utility i in time period t.

Table 2.3 below lists all of the variables used for our empirical analysis. To provide a clear connection to our theoretical framework, we also include a label for each variable's theoretical counterpart and the hypothesized sign of their respective regression coefficients. For variables not included in our theoretical framework, we include information from the literature to include them. Due to limitations within our data, we do not have an empirical measure for labor.

Table 2.3 Empirical Explanatory Variables used to Model Power System Outages

Empirical Variable	**Label**	**Theoretical Counterparts**	**Expected Sign of Estimated Coefficient**
=1 if Utility has an Automated Outage Management System	$Auto_{it}$	K_{it}	Positive
=1 if Utility Transmits Electricity	TR_{it}	K_{it}	Positive
=1 if Utility Distributed Electricity	D_{it}	K_{it}	Positive
=1 if Utility Generates Electricity	G_{it}	K_{it}	Positive
Number of Distributional Circuits	$Circuits_{it}$	K_{it}	Positive
=1 if Utility Identifies Wind/Solar as Prime Mover and 0 otherwise	$Renewable_{it}$	R_{it}	Positive
Net capacity of Electricity Generated by Intermittent Renewables (MWh)	$Intermittent_{it}$	r_4	Positive

[22] According to the EIA, a prime mover is the engine, turbine, water wheel, or another similar machine responsible for driving the electric generator in a power plant; or, for reporting purposes, the device that converts energy to electricity directly.

Empirical Variable	Label	Literature	Expected Sign of Estimated Coefficient
Total Retail Sales (MWh)	$Sales_{it}$	Eto et al. (2012)	Positive
=1 if Utility uses the IEEE Standard 1366	$IEEE_{it}$	Eto et al. (2012)	Positive
Number of Customers	$Customers_{it}$	Malla (2013)	Positive
North American Electrical Reliability (NERC) Regions	$NERC_{it}$	Eto et al. (2012)	Indeterminant

When analyzing short (e.g., many entities i but few time periods T), unbalanced ($T_i \neq T$ for some i) panel datasets, the conventional technique used in economics is to rely on multivariate regression models that specifically account for unobserved heterogeneity (Cameron and Trivedi 2009; Wooldridge 2010). Similar to Eto et al. (2012), our panel dataset is unbalanced because it does not contain reliability metrics for each electric utility across all five years of observation (Wooldridge 2010). The most common multivariate regression models used when analyzing panel data of this nature are fixed effects and random effects models. The choice of which depends on assumptions surrounding the correlation between the individual unobserved effects (heterogeneity) and included explanatory variables of interest.

Because fixed effects models allow for correlation between the unobserved effect and any included explanatory variables of interest, they are often preferred. However, random-effects models are still used, especially in cases when key explanatory variables of interest do not vary over time. In our specific case, the time-invariant variable of interest is the regional reliability organization of each individual utility, also known as its NERC region. Regional reliability organizations are charged with monitoring and enforcing reliability standards. Being able to estimate the impact of the operations of these specific organizations on reliability is the first reason we prefer the random effects model specification over the fixed effects.

The second reason we preferred the random effects model to the fixed effects model is based on the results of our Hausman test. For both sets of transformed values of *SAIDI* and *SAIFI*, we find a statistically significant difference in the parameter estimates for the time-varying explanatory variables of interest produced by both models, which indicates the random effects specification is preferred to the fixed effects. Therefore, we begin the empirical analysis by estimating equation (1) using both sets of transformed reliability metrics (log-transformed and IHS transformed values of *SAID*I and *SAIFI*) assuming a random effects model specification.

It is important to note, however, that a common problem when analyzing strictly positive microeconomic data, such as the frequency and duration of power outages, is censoring of the dependent variable. A dependent variable is censored when all values within a certain range (e.g., values less than zero) are all reported as a single value (e.g., zero). As a result, conventional regression analysis techniques including ordinary least squares (OLS), are not sufficient because they allow for the prediction of negative outcomes of the dependent variable. The reliability metrics (*SAIDI* and *SAIFI*) from our dataset are strictly positive in nature, and thus censored at zero. The standard model used on panel datasets whenever the dependent variable is censored at zero is the Random Effects Tobit Model. [23]

While the Random Effects Tobit Model makes use of the strictly positive censored nature of the indices used to measure the reliability of service, the specification also assumes each

[23] The results from the estimating the model in equation (3) assuming a random-effects Tobit model specification are included in the Appendix for completeness. See Table 2. A1 in the list of tables. We estimate the random effects Tobit model using only the IHS transformed values of SAIDI and SAIFI, as these values are all strictly greater than or equal to zero, while the log-transformed values are not.

explanatory variable of interest in our model affects the probability that a customer will experience a disturbance in power system reliability in the same direction as the intensity of disturbance given that it has occurred. That is, across the two separate regimes, any included explanatory variable of interest (e.g., the net capacity generated by intermittent renewable resources or whether or not the utility can automatically detect disturbances) will impact the probability that a customer experiences a disturbance in power system reliability in the same way that it would impact the duration of the disturbance given that it has occurred.

This restriction might be slightly unreasonable in an economic setting involving utility firm-level disturbance data, where it is conceivable that given a certain level of technology, some electric utilities may be able to better manage the duration of power outages when and if they do occur. Therefore, in addition to estimating a random effects model using the log and IHS transformed values of *SAIDI* and *SAIFI* as the dependent variable, our statistical approach also employs the use of a two-step Cragg-Hurdle Model (Cragg 1971; Greene 2012). The Hurdle model, as it is more commonly known, is not only well-suited to handle the excess zeros problem (i.e., censoring of the dependent variable) but also allows us to estimate separate equations for the bounded and non-bounded outcomes. That is, it allows us to model the probability of experiencing a disturbance in power system reliability separately from the intensity of the disturbance, given it has occurred.

We employ the Hurdle model and estimate equation (1) using only the IHS transformed values of *SAIDI* and *SAIFI*, as these values are all strictly greater than or equal to zero, while the log-transformed values are not. Empirically, the Hurdle model is characterized by the following relationship:

(13) $y_{it} = s_{it} D_{it}^{*}$.

In equation (12) y_{it} represents the observed value of the dependent variable (i.e., the IHS transformed value of *SAIDI* or *SAIFI*); s_{it} represents the "participation decision" which is a binary variable that determines whether y_{it} is zero or strictly positive; and D_{it}^* is a continuously distributed, non-negative latent variable that is only observed if $s_{it} = 1$.[24]

Under the Hurdle model specification, the participation "decision" or selection variable s_{it} is structured as follows:

$$(14) \quad s_{it} = \begin{cases} = 1 & if\ \boldsymbol{Z}_{it}\gamma + v_{it} > 0 \\ = 0 & Otherwise \end{cases}.$$

Here $\boldsymbol{Z}_{it}$ represents a vector of explanatory variables believed to influence whether or not an electric utility experiences a disturbance in power system reliability; γ represents the coefficients to be estimated, and v_{it} is the standard normal error term. The variables in $\boldsymbol{Z}_{it}$ include: = $Auto_{it}$, an indicator variable equal to one if the electric utility has an outage management system that can be used to automatically detect disturbances in power system reliability; $Customers_{it}$, a continuous variable equal to the total number of customers served by each utility in a given year; and $Renewable_{it}$ an indicator variable equal to one if the utility lists wind or solar as their prime mover. These variables were chosen because based on how they are defined and recorded, it is highly probable they contribute to whether or not a utility experiences an interruption in power system reliability.

In the Hurdle model, the second regime is specific to the participation decision. That is, the second regime is used to analyze the intensity of a power system outage given that it has occurred. The intensity of the outage is modeled as follows:

[24] Other than s_{it} being binary and d_{it}^* being continuous, there is another critical difference between s_{it} and d_{it}^*: we effectively observe s_{it} because it is observationally equivalent to the indicator $1[y_{it} > 0]$. However, we only observe d_{it}^* when $s_{it} = 1$.

(15) $E[Outage_{it} \mid s_{it} = 1] = \pi_0 + \pi_1 Intermittent_{it} + \pi X_{it} + \lambda Year_{t-1} + \alpha_i + \varepsilon_{it}$,

where $Intermittent_{it}$ represents the net capacity of electricity generated by intermittent renewable resources; X_{it} is the vector of explanatory variables outlined above; $Year_{t-1}$ is used to represent a set of year dummy variables included for all but one of the five years of observation; π and λ represent parameters to be estimated; α_i is used to represent the unobserved individual-utility-level characteristic believed to influence reliability of service; and ε_{it} is the idiosyncratic error term. [25]

In the case when the Hurdle model outlined is assumed to be linear, the selection model can be viewed as a Probit model, while D_{it}^* is assumed to follow a truncated normal distribution. The model is unique in that the parameters and covariates in both models are allowed to differ. The marginal effects are assumed to estimate the marginal impact of each explanatory variable of interest on the dependent variable (e.g., transformed values of *SAIDI* or *SAIFI*) conditional on a power system outage occurring. Applying the hurdle model to examine power system interruptions is a convenient way to allow different mechanisms to account for the "participation" and "extent" of power outages experienced by end-consumers, which is the primary reason we use it in this analysis.

2.7 RESULTS AND DISCUSSION

The results from the application of the *F*-test are presented below in Table 2.4.

Table 2.4 *F*-Test Results for Hypothesis that Utility-Level Specific Effects Are Not Present

Reliability Metric	*F*-test	Degrees of Freedom (between/within)	Prob. $> F$
$\ln(SAIDI)$	7.96	(247/662)	< 0.000
$\ln(SAIFI)$	7.11	(247/662)	< 0.000
$\text{IHS}(SAIDI)$	9.12	(247/662)	< 0.000
$\text{IHS}(SAIFI)$	7.24	(247/662)	< 0.000

[25] Under the specification outline above d_{it}^* is observationally equivalent to equation (6).

Our findings suggest that both utility and year specific effects are statistically significant (at the 0.01% confidence level) for both the log-transformed and IHS transformed values of *SAIDI* and *SAIFI*. The results suggest there is a strong correlation between the individual characteristics of the utilities and the reliability metrics recorded, as well as between the reliability metrics and the year in which they were recorded. Similar to Eto et al. (2012) we assume this effect is due to the differences in reporting and monitoring practices of the individual utilities. By taking the time effects and utility-specific characteristics into account within our model, we can assume the utility-specific and time effects are not contaminating the estimated coefficients on any other explanatory variables of interest.

Equation (11) was estimated using the transformed values of *SAIDI* and *SAIFI* (i.e., the log-transformed and IHS transformed values) as dependent variables, assuming a simple linear random effects model specification with standard errors corrected for both heteroscedasticity and autocorrelation. The estimation results are presented below in Table 2.5.

Table 2.5 Log-linear and IHS Random Effects Specification Results

	$\mathbf{ln}(\boldsymbol{SAIDI})$ (1)	$\mathbf{ln}(\boldsymbol{SAIFI})$ (2)	$\mathbf{IHS}(\boldsymbol{SAIDI})$ (3)	$\mathbf{IHS}(\boldsymbol{SAIFI})$ (4)
Customers	-4.46×10^{-7}***	-3.17×10^{-7}*	-4.25×10^{-7}**	-1.66×10^{-7}**
	(1.95×10^{-7})	(2.00×10^{-7})	(2.09×10^{-7})	(8.56×10^{-8})
IEEE	1.2504***	-0.1800*	1.406***	0.2601***
	(0.2002)	(0.0964)	(0.2107)	(0.0427)
Auto	0.1211	0.0409	0.1174	0.0571*
	(0.1099)	(0.0633)	(0.1217)	(0.0316)
Circuits	0.0004*	0.0002	0.0004*	0.0001
	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Retail Sales	9.95×10^{-9}	7.03×10^{-9}	9.10×10^{-9}	2.72×10^{-9}
	(6.92×10^{-9})	(5.11×10^{-9})	(7.68×10^{-9})	(2.34×10^{-9})
Transmit	0.1662	0.1259*	0.1228	0.0596
	(0.1251)	(0.0707)	(0.1384)	(0.0447)

Distribute	0.4703 (0.4965)	0.0940 (0.1788)	0.5405 (0.5389)	0.1920* (0.1120)
Generation	-0.2400 (0.1920)	0.3002** (0.1449)	-0.2625 (0.2148)	0.0224 0.0487
Renewable W/O Hydro	-0.2922*** (0.1121)	-0.0396 (0.0725)	-0.3065*** (0.1203)	-0.0966*** (0.0372)
Renewable Generation	3.42×10^{-7}*** (1.38×10^{-7})	2.21×10^{-8} (7.72×10^{-8})	3.27×10^{-7}** (1.42×10^{-7})	4.71×10^{-8} (4.39×10^{-8})
NERC Region Dummies	Yes	Yes	Yes	Yes
Ownership Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Constant	3.96*** (0.71)	0.32 (0.30)	4.54*** (0.80)	0.81*** (0.19)
σ_u	1.189	0.632	1.318	0.330
σ_ε	0.903	0.513	0.931	0.284
ρ	0.635	0.603	0.667	0.574
Observations	**937**	**937**	**937**	**937**
Hausman test (m-value)	7.68	13.99	8.11	7.73
Hausman $\chi^2(11)$	0.7417	0.2335	0.7034	0.7375

Standard errors in parenthesis

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

The results from all but one of our regression models indicate measuring and recording values for *SAIDI* and *SAIFI* following the IEEE 1366 Standard, has a positive and significant effect on the frequency and duration of power outages recorded by utilities. Additionally, consistent with our theoretical expectations, we find utilities who transmit and distribute their own electricity can expect to have longer and more frequent power outages. However, this result is not statistically significant. Our results indicate utilities who operate more distributional circuits can expect to experience longer power outages, as measured by the transformed values of *SAIDI*.

Across all random effects model specifications, we find utilities who identify wind or solar (photovoltaics) as a prime mover of at least one of their power plants, on average can expect to experience shorter and less frequent power outages than utilities who do not identify as such. This result is not consistent with our theoretical expectation. However, this result could be related to the fact that utilities who operate at least one power plant that relies on intermittent renewable resources to generate power might be better at managing periods when wind and/or solar availability is likely to be limited. As a result, theses utilities may be more prone to take the actions necessary to avoid power system outages. For example, utilities who operate a power plant or set of power plants that primarily depend on wind and/or solar might decide to take those power plants offline during periods of severe weather and replace them with a natural gas-fired generating plant.

However, our results also indicate as the net capacity of electricity generated from intermittent renewable resources increases, longer power system outages, as measured by the transformed values of *SAIDI*, can be expected. This result is consistent with our theoretical expectation. Although increasing the net capacity of electricity generated by intermittent renewables was found to have a statistically significant impact on power system reliability according to models (1) and (3), where the log-transformed and IHS transformed values of *SAIDI* serve as the dependent variable, the result is marginal. More specifically, the random effects model results indicate, all else equal, generating one additional MWh of electricity from intermittent renewable resources results in only a 0.00003% increase at the margin in the duration of disturbances as measured by both the IHS and log-transformed values of *SAIDI*.

In addition to estimating equation (11) assuming a simple random effects model specification, we also estimate the model in equation (11) using a Cragg-Hurdle model specification, as outlined by equations (12) and (13). The results from estimating the Hurdle model, using both the IHS transformed values of *SAIDI* and *SAIFI* are listed in Table 2.6 below.

Table 2.6 Linear Hurdle Model Results

	IHS($\boldsymbol{SAIDI}$) (5)	**IHS**($\boldsymbol{SAIFI}$) (6)
Customers	2.88×10^{-6}*** (7.73×10^{-7})	4.77×10^{-7}*** (1.45×10^{-7})
IEEE	0.3477*** (0.1047)	0.0552 (0.0426)
Auto	0.2868 (0.1798)	0.0580 (0.0409)
Circuits	0.0003 (0.0001)	0.00004 (0.0001)
Retail Sales	7.94×10^{-9}*** (3.83×10^{-9})	1.52×10^{-9} (1.93×10^{-9})
Transmit	0.3634*** (0.1138)	0.0837 (0.0514)
Distribute	-0.6785*** (0.1704)	-0.1844*** (0.0644)
Generation	-0.0018 (0.1371)	0.1423* (0.0767)
Renewable w/o Hydro	-0.3360 (0.2580)	-0.0901 (0.0637)
Renewable Generation	4.44×10^{-7}*** (1.12×10^{-7})	6.03×10^{-8} (5.06×10^{-8})
NERC Region Dummies	Yes	Yes
Ownership Dummies	Yes	Yes
Year Dummies	Yes	Yes
Constant	6.4375*** (0.3077)	1.5241*** (0.1789)
Selection Model		
Auto	0.2062 (0.1843)	0.4075 (0.1824)
Customers	3.74×10^{-6}*** (8.82×10^{-7})	3.80×10^{-6}*** (8.56×10^{-7})

Renewable w/o Hydro	-0.1262	-0.2616
	(0.2641)	0.2392
Constant	0.4919***	0.3774***
	(0.1250)	0.1247
Ln(Sigma)		
Constant	-0.0968	-0.8464
	(0.0581)	0.1270
Sigma	0.9077	0.4289
	(0.0527)	(0.0545)
Log-likelihood	-1,365.43	-643.21
Observations	**937**	**937**

Standard errors in parenthesis

$* \ p < 0.10, \ ** \ p < 0.05, \ *** \ p < 0.01$

Consistent with our previous results, we find using the IEEE 1366 Standard to measure and record values of *SAIDI*, leads to longer disturbances on average being recorded by electric utility companies (firms). Additionally, we find electric utility companies (firms) who distribute their own electricity are approximately 70% less likely to experience a disturbance in power system reliability as measured by the IHS transformed value of *SAIDI,* and 20% less likely to experience a disturbance in power system reliability as measured by the IHS transformed value of *SAIFI.* Based on the results from the selection model, we can conclude electricity companies (firms) with more customers are more likely to experience an outage as measured by *SAIDI* and *SAIFI* than not.

Consistent with the linear random effects model specification results for the log-transformed values of *SAIDI* we find generating more net capacity from intermittent renewable resources leads to longer outages (i.e., disruptions in network system reliability). Based on our results we can infer, utilities who experience a power outage can on average, expect a 0.00004% increase in the duration of the disturbance as measured by the IHS transformed value of *SAIDI* if they generating one additional MWh of net capacity from intermittent renewable resources. While our findings are statistically significant, it is important to note, however, that they currently are not necessarily economically significant.

For example, in terms of customer minutes without power, across all model specifications, our results indicate increasing the net capacity of electricity generated by intermittent renewable resources will lead to power system outages that on average last less than one second for end-consumers. Moreover, in order for the reliability of service for end-consumers, as measured by IHS transformed values of *SAIDI* to decrease by 100% (a power outage lasting additional one minute), electric utility companies would need to generate on average between 2.3 to 3.1 million MWh of net capacity from intermittent renewable resources each year. Currently, only 5 (less than 1%) of the electric utility companies in our sample meet this criterion; an indication that it may be a while before large-scale power interruptions are realized.

Nevertheless, in areas where relatively large, non-marginal increases in the proportion of electricity generated from intermittent renewables are expected in the near future (or in some cases mandated) our estimation results imply that electrical power system reliability could decrease substantially. For example, currently across the United States more than 90 cities, ten counties, and two states have renewable energy policies in place requiring electric utility companies to generate 100% of their net capacity from renewable energy resources (Sierra Club 2019). These imposed increases in intermittent renewable resource capacity beg the question: “If power system reliability is already being negatively affected at the margin by the use of intermittent renewable resources, how will rapid, large-scale transitions to renewable energy resources affect network system reliability in the future both technically and economically?

To provide some perspective on the policy implications of our results, in the following section of this paper we forecast interruptions in power system reliability assuming different renewable energy policy scenarios. We then use these forecasts to predict costs associated with forecasted interruptions in power system reliability using the Interruption Cost Estimator (ICE) as suggested by Sullivan et al. (2018). Our forecasted results are based on Renewable Energy Standards (RES) which are regulatory mandates that require electric utility companies within a given area (i.e., city, state, or region) to source a certain amount of the electricity they generate from intermittent renewable resources.[26] Most RES are designed to incrementally increase the capacity of electricity generated by intermittent renewables over time.[27] More than half of all U.S. states have some type of RES policy or goal in place (SEIA 2019).

FORECASTING POWER SYSTEM RELIABILITY UNDER RENEWABLE ENERGY STANDARDS

To begin our forecasting procedure, we identified ten states that currently have some type of RES in place. The states identified include California, Texas, Hawaii, Iowa, Kansas, Michigan, New Mexico, Oregon, Vermont, and Colorado. Of these states, California, Hawaii, and Vermont were found currently to have the most aggressive RES in place. For example, California has a renewable energy policy in place requiring utilities to generate 33% of the electricity they sell from renewable energy resources by the year 2020; 40% by the year 2024; 45% by the year 2027; and 50% by the year 2030. Hawaii, on the other hand, has plans in place to source 100% of its electricity from intermittent renewable resources by the year 2045 while Vermont's plans include having 55% of its electricity generated by renewables by 2017 and 75% by the year 2032.

[26] There are many different variations of RES, including clean energy targets, which allow utilities to utilize nuclear energy resources and other low-polluting non-renewable alternatives such as natural gas, and renewable portfolio standards (RPS) which require utilities to ensure a specified amount of the electricity they sell comes from renewable energy resources and other low-pollution alternatives.

[27] For example, an RES might require utilities to increase the capacity of electricity generated by renewables by 1% each year for the next ten years, resulting in a cumulative 10% increase in renewable generation in that state. Conversely, an RES might specify that by a certain date (e.g., ten years in the future) at least 10% of the electricity generated in a given state must come from intermittent renewables.

To obtain the predicted outcomes, we use the to calculate the total net capacity associated with a given percentage increase in net capacity generated by intermittent renewable resources. For example, if the net capacity currently generated by an individual utility in a given state is 100,000-megawatt hours (MWh), then a 10% increase in capacity generated by intermittent renewables would be equivalent to 10,000 MWh. Based on current RES in place, when forecasting outages in power system reliability we considered increases in capacity from renewables ranging from 5% to 100%. The results from our forecasting procedure are listed below in Table 2.7.[28]

[28] A detailed description of the forecasting procedure used for our analysis can be found in the Chapter 2 Appendix.

Table 2.7 Predicted Outcomes for *SAIDI* and *SAIFI* assuming a Random Effects Model and Estimation Results

Percentage of Net Capacity Generated by Intermittent	California		Texas		Hawaii		Iowa		Kansas	
Renewables	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*
5%	.599	131	.755	195	1.351	147	.671	273	.895	230
10%	.610	153	.795	228	1.356	153	.694	413	.907	254
15%	.622	187	.815	268	1.362	159	.719	640	.920	286
20%	.633	238	.836	318	1.369	165	.745	1,008	.933	326
25%	.646	317	.857	381	1.375	172	.772	1,605	.946	379
30%	.658	441	.878	459	1.381	179	.801	2,578	.960	447
50%	.713	2,361	.970	1,038	1.405	210	.937	18,008	1.020	999
75%	.795	27,275	1.098	3,286	1.437	261	1.161	215,142	1.106	3,379
100%	.899	344,745	1.247	11,567	1.470	329	1.471	2,638,574	1.207	12,789

Percentage of Net Capacity Generated by Intermittent	Michigan		New Mexico		Oregon		Vermont		Colorado	
Renewables	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*	*SAIFI*	*SAIDI*
5%	.828	252	.671	135	.836	232	.747	31	.879	269
10%	.854	372	.681	156	.846	249	.747	31	.914	356
15%	.882	599	.692	183	.855	269	.748	31	.951	485
20%	.911	1,035	.703	215	.865	292	.749	32	.989	678
25%	.943	1,882	.715	256	.875	319	.749	32	1.029	967
30%	.977	3,535	.726	306	.886	350	.750	32	1.071	1,398
50%	1.144	50,726	.775	657	.928	536	.753	33	1.256	6,650
75%	1.449	1,553,039	.844	1,833	.986	1,020	.757	34	1.545	49,860
100%	1.922	4,8936,440	.922	5,303	1.048	2,122	.760	34	1.917	379,132

Before we interpret the results of the policy simulation above, it is important to recall how values for *SAIDI* and *SAIFI* are measured and recorded by electric utility companies. Both *SAIDI* and *SAIFI* are index measures that were originally developed by the IEEE to quantify sustained interruptions in power system reliability (i.e., interruptions lasting longer than 5 minutes). Based on how the indices are calculated, values for *SAIDI* can be interpreted as the total number of minutes a customer went without power during a disturbance, while values for *SAIFI* can be interpreted as the probability that a customer will experience a power system outage.

Assuming technology remains constant, based on the forecasting results we can expect as the net capacity of electricity generated by intermittent renewable resources reaches 5%, customers can expect to go without power for an additional 190 minutes per year, on average. That is, customers across all ten states can expect to experience disruptions in power system reliability that last for approximately two and half hours per year when 5% of the total net capacity is supplied by intermittent renewables. Conversely, if the net capacity of electricity generated by intermittent renewables exceeds 50%, the number of customer minutes without power is expected to exceed 1,000 in five of the ten states. Moreover, as the percentage of net capacity generated by intermittent renewables exceeds 50%, the probability a customer will experience a disruption in power system reliability reaches 1 in four separate states.

Again, assuming technology remains constant, customers in all states except Hawaii and Vermont can expect the length of time without power to nearly double as the percentage of net capacity generated by intermittent renewables increases from 30% to 50% for all but two states, Vermont and Hawaii. Lastly, given current technology, our forecasting results indicate generating between 50% to 100% of net capacity from intermittent renewable resources will on average increase the probability of experiencing a power outage by 30% as measured by *SAIFI*, across all ten states.

POWER SYSTEM INTERRUPTION COSTS RESULTS

Following the suggestion of Sullivan et al. (2018) we utilized the Interruption Cost Estimate (ICE) Calculator, co-developed by the Lawrence Berkeley National Laboratory and Nexcant Inc., to predict outage cost estimates associated with our forecasted interruptions in power system reliability resulting from increasing the capacity of electricity generated by intermittent renewable resources (see Table 2.7). The publicly-available ICE Calculator is an interactive, user-friendly tool that enables utilities, reliability planners, government organizations, and other interested parties to estimate the benefits associated with improving power system reliability and/or the costs associated with interruptions in power system reliability (Sullivan et al. 2018).

To produce outage cost estimates, the ICE calculator relies on data from 34 prior published papers that estimate the costs associated with interruptions in power system reliability (Sullivan et al. 2018). [29] More specifically, the ICE calculator relies on work by Lawton et al. (2003) and Sullivan et al. (2009) who perform a meta-analysis of customer survey data to create a single database which can be used to estimate electricity customer outage costs. As a result, the ICE Calculator contains information from 105,000 different customer surveys collected by 10 different electric utilities between the years 1989 and 2012. [30]

[29] The data provided was collected from individual utilities, who in the interest of estimating the costs associated with customer interruptions, administered a set of surveys that described hypothetical interruptions and asked customers to estimate the costs they would occur if they experienced interruptions of varying duration, at different times of the day, and during different seasons. Residential customers were asked to indicate the amount they would be willing to pay to avoid interruptions occurring under these conditions. Respondents were typically asked to estimate their costs for between four and eight hypothetical interruptions (Sullivan et al. 2009).

[30] For a complete outline of the methods used to estimate the interruption costs using customer survey data see Sullivan et al. (2009).

To utilize the ICE calculator, one simply needs to input information on the total number and type of customers (i.e., residential or non-residential) served by each utility, the geographic location of the electric utility (i.e., the state in which the utility operates in), and estimated values for the anticipated changes in reliability of service (i.e., values for *SAIDI* and *SAIFI*). [31] Using the above information, the ICE Calculator produces estimates for four key outage cost metrics: (1) the cost per interruption event; (2) the cost per average kW; (3) the cost per unserved kWh of electricity (i.e., lost load); and (4) the total cost of sustained interruptions. Based on suggestions from Sullivan et al. (2009), we estimate the economic costs of power system interruptions using metric (3) and metric (4).

While the ICE Calculator is a useful tool, it is important to note that it does have some limitations. For example, as pointed out by Sullivan et al. (2018), the formula used to calculate the cost per un-served kWh of electricity likely oversimplifies the calculation necessary to determine the quantity of electricity that would have been consumed if a power outage had not occurred. [32] In addition, the surveys used to create the meta-database were conducted sporadically over a 20-year period. Thus, their resulting cost-estimates must be must be temporally adjusted. [33] Moreover, because the underlying costs estimates are based on hypothetical interruption scenarios presented in prior surveys, the ICE it is not designed to predict costs associated with power system

[31] In addition to using values for SAIDI and SAIFI to measure changes in the reliability of service, the ICE Calculator allows utility companies or other interested parties to use values for the Customer Average Interruption Duration Index (CAIDI). CAIDI is similar to SAIFI, except the denominator is the number of customers interrupted instead of the total number of utility customers served by the utility.

[32] To determine the cost per un-served kWh of electricity, the ICE calculator uses the following formula: Cost per un-served kWh = Total Interruption Cost Per Event ($)/(Annual kWh/8760)×Duration of the Interruption

[33] The ICE calculator adjusts the estimates using the Bureau of Economic Analysis's GDP deflator and reports all cost in 2016 dollars.

interruptions that last longer than 32 hours (1,920 minutes). As a result, we are not able to predict outage costs for forecasted interruptions in power system reliability if the value of *SAIDI* > 1,920.[34]

To determine the average number of residential and non-residential customers served by utilities within a given state, we rely on historical information from the U.S. EIA. More specifically, we multiply the average total number of customers served by electric utilities within a given state from our dataset, by the average proportion of residential and non-residential customers served by utilities in the same state over the past five years. For example, from 2013 to 2017, on average 88% of the customers served by utilities in California were residential, while 12% were non-residential customers. The average number of customers served by utilities in the state of California from our data set is 935,728 customers. Therefore, when using the ICE calculator to predict outage costs, we assume there are 795,368 residential customers and 140,360 non-residential customers. Table 2.8 below outlines this procedure for all ten states.

Table 2.8 Residential and Non-Residential Customer Estimates for States with RES in Place

State	Residential	Non-Residential	# of Customers
California	795,368 (88%)	140,360 (12%)	935,728
Texas	253,743 (87%)	37,916 (13%)	291,659
Hawaii	111,567 (87%)	16,671 (13%)	128,238
Iowa	134,787 (85%)	23,786 (15%)	158,573
Kansas	124,297 (83%)	25,458 (17%)	149,755
Michigan	426,803 (89%)	52,752 (11%)	479,555
New Mexico	120,994 (86%)	19,697 (14%)	140,691
Oregon	328,427 (87%)	49,083 (13%)	377,555
Vermont	69,742 (85%)	12,307 (15%)	82,049
Colorado	500,435 (85%)	88,312 (15%)	588,747

The results from our application of the ICE Calculator to the forecast outages in power system reliability are listed below in Tables 2.9 and 2.10.

[34] The same holds true if the value of *SAIFI* ≥ 100 or if the value of *CAIDI* ≥ 960. It is important to note that the ICE Calculator assumes $CAIDI = \frac{SAIDI}{SAIFI}$ (Sullivan et. al 2018).

Table 2.9 Cost per kWh of Unserved Energy Derived from the ICE Calculator ($2016)

Percentage of Net Capacity Generated by Intermittent Renewables	California	Texas	Hawaii	Iowa	Kansas
5%	$81	$37	$61	$42	$39
10%	$81	$37	$61	$46	$40
15%	$82	$38	$59	$43	$41
20%	$85	$39	$59	-	$42
25%	$90	$40	$58	-	$43
30%	$93	$42	$57	-	$45
50%	-	-	$55	-	-
75%	-	-	$53	-	-
100%	-	-	$52	-	-
Percentage of Net Capacity Generated by Intermittent Renewables	**Michigan**	**New Mexico**	**Oregon**	**Vermont**	**Colorado**
5%	$39	$44	$42	$108	$60
10%	$42	$44	$42	$108	$64
15%	$46	$44	$43	$108	$68
20%	-	$45	$43	$106	$71
25%	-	$47	$44	$106	$62
30%	-	$49	$45	$106	
50%	-	$50	$49	$104	
75%	-	-	-	$102	
100%	-	-	-	$103	

Table 2.10 Total Costs of Sustained Power System Interruptions ($2016)

Percentage of Net Capacity Generated by Intermittent Renewables	**California**	**Texas**	**Hawaii**	**Iowa**	**Kansas**
5%	$378,389,323	$116,758,732	$48,823,689	$103,667,456	$74,007,935
10%	$441,530,518	$116,758,732	$50,085,182	$171,800,467	$82,373,782
15%	$546,919,059	$166,075,951	$51,380,177	$247,091,716	$94,661,329
20%	$721,139,099	$203,390,402	$52,709,761	-	$110,921,454
25%	$1,013,475,610	$253,100,645	$54,255,240	-	$133,882,770
30%	$1,454,679,779	$316,476,207	$55,803,719	-	$164,948,663
50%	-	-	$62,919,566	-	-
75%	-	-	$75,351,087	-	-
100%	-	-	$93,497,940	-	-
Percentage of Net Capacity Generated by Intermittent Renewables	**Michigan**	**New Mexico**	**Oregon**	**Vermont**	**Colorado**
5%	$201,712,744	$35,012,536	$177,264,496	$9,065,974	$376,627,613
10%	$323,822,946	$40,486,361	$191,960,127	$9,065,974	$527,912,925
15%	$564,339,402	$48,049,758	$209,900,097	$9,084,157	$772,761,835
20%	-	$57,706,787	$231,812,065	$9,163,558	$1,135,217,861
25%	-	$71,094,319	$257,200,934	$9,163,558	$1,394,960,580
30%	-	$88,661,864	$287,910,104	$9,163,558	-
50%	-	$197,158,491	$483,105,475	$9,280,432	-
75%	-	-	-	$9,398,390	-
100%	-	-	-	$9,435,474	

An unserved kWh of electricity is an amount of electricity demanded but not supplied due to an unplanned interruption in power system reliability. In other words, it is the amount of electricity that would have been consumed if the power outage had not occurred. Therefore, the costs per unserved kWh of electricity in Table 2.9 should be interpreted as the cost incurred by customers when their desired amount of electricity was not supplied because of an unplanned interruption.

Overall, we find that for seven out the ten states the cost per unserved kWh of electricity resulting from longer power outages increased as the percentage of capacity supplied by intermittent renewable resources increased, holding technology constant. However, we observed the opposite result in Vermont and Hawaii where the cost per un-served kWh of electricity decreased as the capacity of electricity supplied by intermittent renewable resources increased, holding technology constant. Across all ten states, the cost per unserved kWh of electricity is nearly 100 times more than the price actually paid per kWh electricity delivered.

For example, in 2016 customers in the state of Hawaii (which typically has the highest electricity rates) paid on average 23.87 cents per kWh of electricity they consumed. However, going without power is expected to cost them between $52-$62/kWh of electricity not delivered. The same holds true for customers in Texas, who paid less than 10 cents per kWh of electricity they consumed in 2016. Thus, the implication is that the costs of going without power, even for a short amount of time, far exceed the price of delivering a reliable supply of electricity to end-consumers.

The total cost of sustained power system interruptions in Table 2.10 are the overall costs of power system interruptions lasting longer than five minutes. These costs include the costs customers face when they experience a disruption in power system reliability, as well as the costs utilities incur when interruptions occur. Utility costs are referred to as component costs, and include costs such as lost production/sales, damage to equipment or materials, extra overhead, additional labor and overtime costs necessary to restore power, as well as any other costs associated with interruptions in power system reliability.

As shown in Table 2.10, for all 10 states the total costs of sustained interruptions in power system reliability resulting from longer and more frequent power outages for utilities and their customers alike increased as the percentage of capacity supplied by intermittent renewable resources increased, holding technology constant. The total costs of sustained interruptions in power system reliability appear to be the lowest in Vermont, staying within the range of $9-10 million (USD). However, as the percentage of capacity generated by renewables exceeds 20% such as in California and Colorado, the total cost of sustained interruptions in power system reliability begin to exceed $1 billion (USD).

2.8 SUMMARY AND CONCLUSIONS

The overall purpose of this chapter (essay) was to present a theoretical and empirical analysis of the effects on energy reliability of increasing the capacity of intermittent renewable resources such as wind and solar in the electric power grid. Our state-contingent theoretical model illustrates the fundamental problem faced by power utility operators who need to decide on how much intermittent resource power to bring on-line in the future under uncertainty, including uncertain weather conditions. The theoretical model suggests that if the weather and associated wind and solar conditions in the future are not as the operator expected, the result can be too little or too much power generated by intermittent renewable resources resulting in power system outages. Such outages result in less electricity delivered to end-customers and less profit to utility companies.

In the empirical portion of this study, we examine whether or not increasing the net capacity of electricity generated by intermittent renewable resources (i.e. wind and solar energy) affects the frequency and/or duration of disturbances in the electrical power generating system of the United States. Consistent with the state-contingent theoretical model, we used two measures of power interruptions or outages to end-customers; the System Average Interruption Duration Index (*SAIDI)* and the System Average Interruption Frequency Index (*SAIFI*) collected from two annual surveys administered by the U.S. Energy Information Administration (EIA). Our econometric model of energy reliability uses end-user interruptions or outages as measured by *SAIDI* and *SAIFI* as proxies for in-service reliability, such that higher values of *SAIDI* and *SAIFI* indicate lower levels of service reliability.

Using an unbalanced panel of 276 U.S. electric utility companies from 2013 to 2017, we modeled reliability of service (as measured by *SAIDI* and *SAIFI*) as a function of the net capacity of electricity generated by intermittent renewable resources, while controlling for other operational level characteristics of individual utilities believed to influence the reliability of service. Given the panel nature of our data, we conducted a Hausman test, the results of which indicated estimating the model assuming a random effects model specification, rather than a fixed effect specification, was preferred. Overall, our empirical results suggest increasing the net capacity of electricity generated by intermittent renewable resources has a statistically significant negative marginal effect on electrical system reliability. The magnitude of this marginal effect was relatively small.

More specifically, our findings suggest that over a one-year period, increasing the net capacity of electricity generated by intermittent renewable resources, on average has led to less than a 1% increase in the duration of outages for end-consumers, as measured by *SAIDI*. In terms of customer minutes without power, we find increasing the net capacity of electricity generated by intermittent renewable resources leads to power outages that on average last less than one second.

Although our regression model results showed a statistically significant positive marginal effect of increasing intermittent resource capacity on electric grid reliability, this negative effect may not be economically significant at the margin because of its relatively small in magnitude. However, our analysis suggests that non-marginal increases in net capacity generated by intermittent renewables of 25% or more may result in substantial increases in power system outages. Thus, intermittent resource-inducted power interruptions could become more problematic in the future, especially in regions of the United States experiencing rapid, non-marginal increases in intermittent renewable capacity driven by public policy such as Renewable Energy Standards.

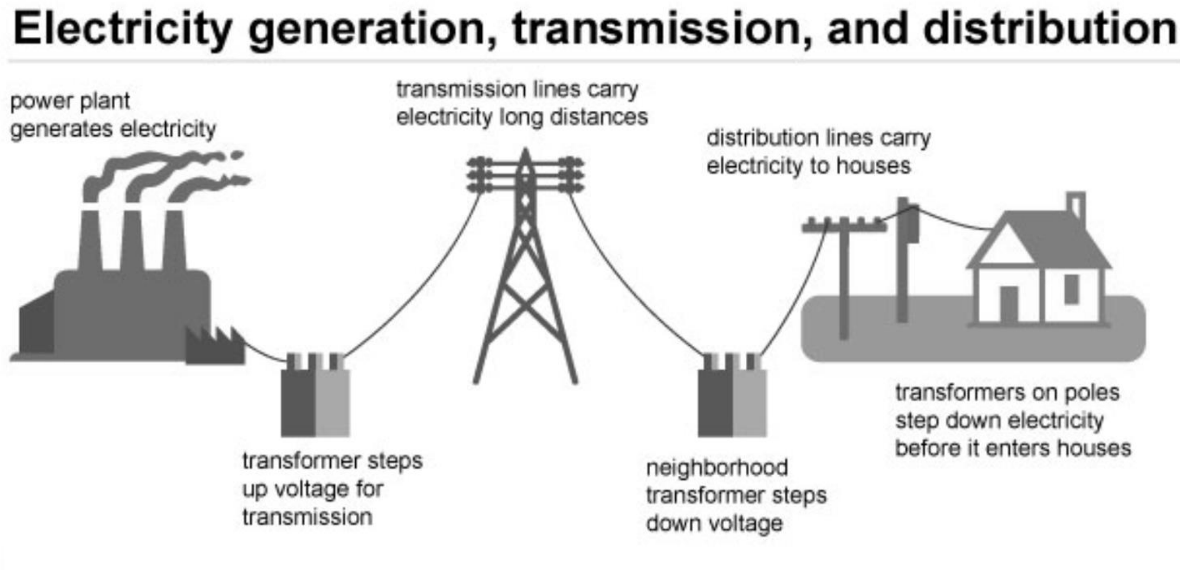
To provide some perspective on the economic and policy implications of our results, we forecast outages in power system reliability assuming different renewable energy policy scenarios for ten different states. We use these forecasts to predict the costs associated with forecasted power outages using the Interruption Cost Estimator (ICE). Our forecasted results indicate that as the percentage of power grid capacity generated by intermittent renewable resources increases, so too will the frequency and duration of power system interruptions. Moreover, once the capacity supplied by intermittent renewables exceeds 50%, outages will go from lasting a few hours to lasting a few days, for one-half of the states, assuming technology remains the same.

Moreover, the costs associated with un-delivered kWh of electricity will increase as the percentage of capacity generated by intermittent renewables increases. In all cases, these costs far exceed what customers currently pay per kWh of electricity delivered thereby making electricity less affordable to consumers. More expensive electricity can reduce private and commercial business profitability and the ability of households to obtain adequate levels of energy security and avoid energy poverty. As a society such as the United States transitions to a renewable energy economy over time (e.g., exhaustible, nonrenewable fossil fuels will eventually run out, at least economically), there will be a continual need to assess both the benefits and costs of increased reliance on intermittent energy resources such as wind and solar.

A limitation of our analysis is the lack of an empirical measure for labor. The number of workers (i.e., technicians) needed to manage a power system operation likely remain constant from year to year. However, larger utilities are likely to employ more units of labor than smaller utilities. Depending on the returns to scale of labor, having access to more employees could help prevent interruptions in power system reliability. Additionally, access to skilled workers, especially those who are knowledgeable on how to account for the variable nature of intermittent renewable resources, could have a direct impact on the frequency and duration of power system interruptions.

It is our hope that this study will help stimulate additional theoretical and empirical investigations of how power grid reliability is being managed as the capacity of intermittent renewable resources in the U.S. electrical grid increases. Additionally, we hope to contribute to the overall policy discussion on power grid reliability, including how to properly manage the influx of renewable energy resources. Moving forward, we plan on extending the policy implications of our results by forecasting power system outages using the estimation results from our two-step Cragg Hurdle model. Additionally, we plan to explore alternatives to the ICE calculator for estimating the economics losses (i.e., costs) customers experience from power interruptions using our empirical model and data.

APPENDIX CHAPTER 2



Source: U.S. Energy Information Administration Modified from the National Energy Education Development Project

Figure 2. A1 Three Phases of the Electricity Production Process

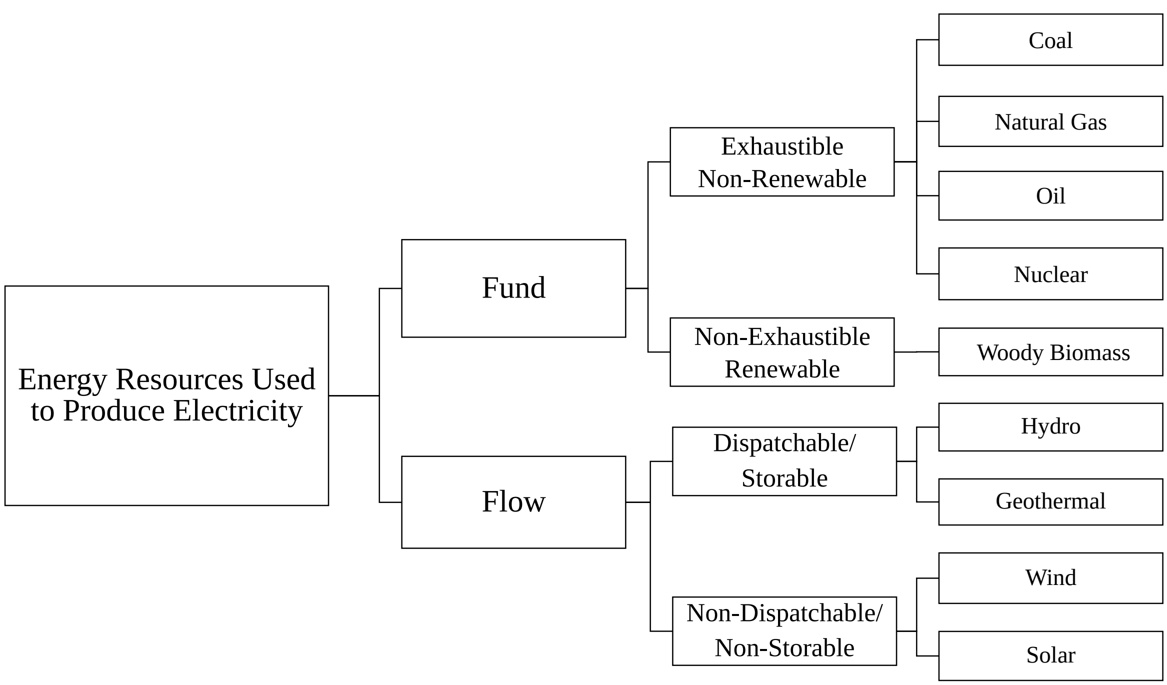
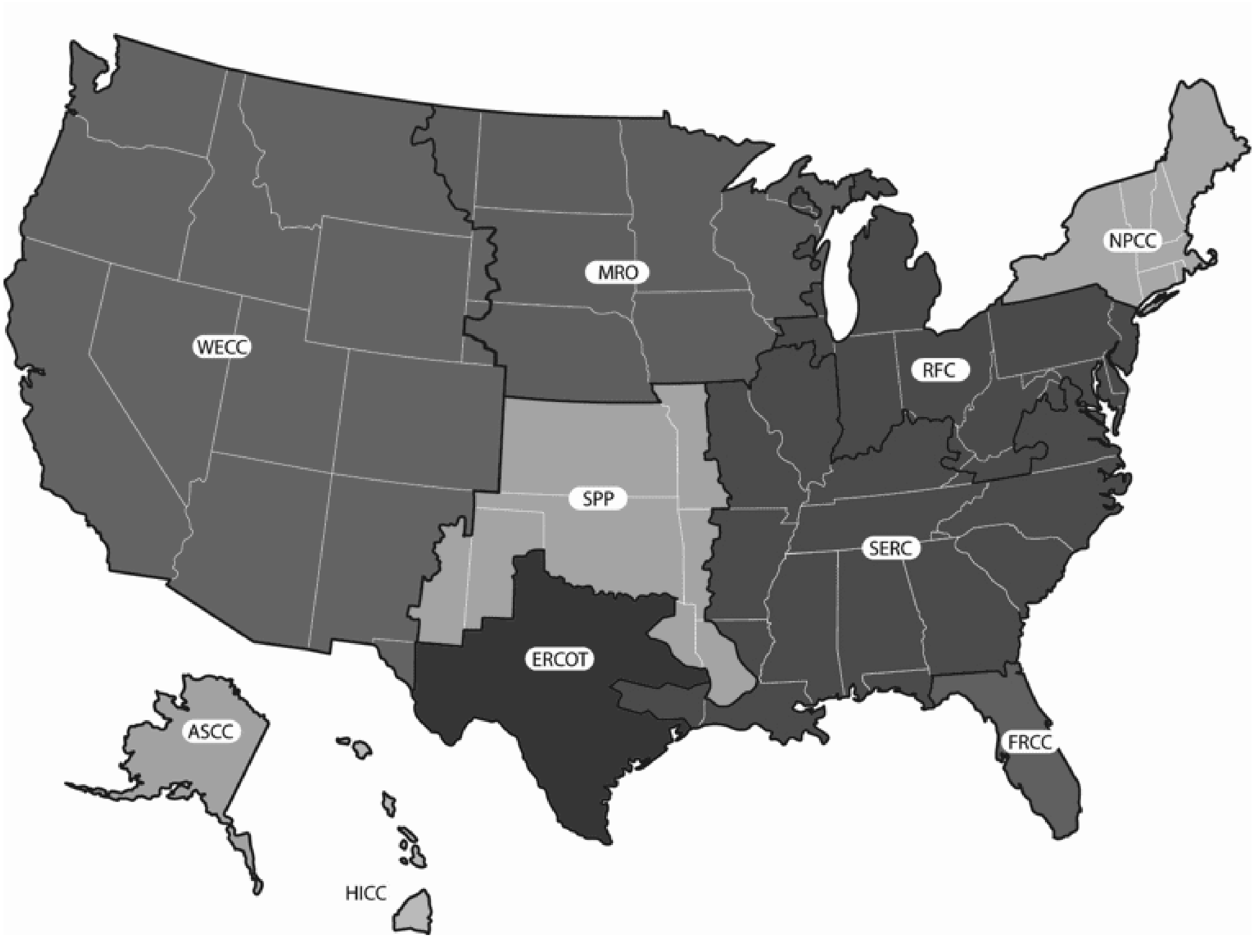


Figure 2. A2 Energy Resources Used to Produce Electricity



Source: U.S. Environmental Protection Agency EGrid Map

Figure 2. A3 North American Reliability Council (NERC) Regions

Table 2. A1 Random Effects Tobit Model

	IHS *SAIDI*	**IHS *SAIFI***
Customers	-5.15×10^{-7}	-2.05×10^{-7}
	4.33×10^{-7}	1.15×10^{-7}
IEEE	1.6362***	0.3801***
	0.1535	0.0455
Auto	0.0494	0.0660
	0.1413	0.0418
Circuits	0.0005	0.0002
	0.0004	0.0001
Retail Sales (MWh)	9.69E-09	2.76E-09
	1.54E-08	4.09 E-9
Transmit	0.1431	0.0865
	0.1831	0.0543
Distribute	-0.4603	-0.1077
	0.9193	0.2760
Generation	-0.4985*	-0.0408
	0.2952	0.0840
Renewable W/O	-0.3353	-0.1131
Hydro	0.2104	0.0611
Renewable Generation	3.23×10^{-7}	5.76×10^{-8}
	3.86×10^{-7}	1.11×10^{-7}
Ownership Dummies	Yes	Yes
NERC Region Dummies	Yes	Yes
Year Dummies	Yes	Yes
Constant	5.9610***	1.1704***
	0.1831	0.3372
σ_u	1.6056	0.4109
	0.0935	0.0251
σ_ε	1.0137	0.3116
	0.0302	0.0092
ρ	0.7150	0.6349
	0.0269	0.0320
Log-Likelihood	-1,464.42	-492.79
Observations	924	924

Standard errors in parenthesis

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

A Solution for Predicted Values of y_{it}

The data generating process assumed for a random effects model specification where an inverse hyperbolic sine transformation has been applied to the dependent variable (y_{it}) can be expressed as follows:

(A.1) $\ln\left(y_{it} + \sqrt{y_{it}^2 + 1}\right) = X_{it}'\beta + a_i + u_{it}.$

Here X_{it}' represents the individual observations; β represents a vector of parameters to be estimated; a_i represent the unobserved individual-level effect (i.e., unobserved heterogeneity); and u_{it} represents the idiosyncratic error term. Suppose we are interested in obtaining predicted outcomes (i.e., forecasts) for the transformed value of the dependent variable, $\ln\left(\widehat{y_{it} + \sqrt{y_{it}^2 + 1}}\right)$. In addition, assume we know the true value for each of the parameters in the vector β.

To obtain the predicted outcomes, we would simply need to substitute in values for the known parameters and then take the expected value of the right-hand side, conditioning on the known observations X_{it} as follows:

(A.2) $\ln\left(\widehat{y_{it} + \sqrt{y_{it}^2 + 1}}\right) = E[X_{it}'\beta + a_i + u_{it}|X_{it}].$

Because the β's are constant and the values for each individual observation are known, $E[X_{it}'\beta] = X_{it}'\beta$. Moreover, given that we included an intercept and assumed the idiosyncratic error term was normally distributed, we can assume $E[a_i] = 0 \;\forall\; i$ without loss of generality and $E[u_{it}|X_{it}] = 0 \;\forall\; i \; and \; t$. Therefore, the predicted value of a dependent variable that has been transformed using an IHS can be expressed as follows:

(A.3) $\ln\left(\widehat{y_{it} + \sqrt{y_{it}^2 + 1}}\right) = X_{it}'\beta \;\; \forall \; combinations \; of \; i \; and \; t.$

The situation for predicting values for the non-transformed value of y_{it}, however, is a bit more complicated. To predict values for y_{it}, we must first solve for y_{it} by taking the inverse of the inverse hyperbolic sine. Assuming the same data generating process as before

(A.4) $\ln\left(y_{it} + \sqrt{y_{it}^2 + 1}\right) = X'_{it}\beta + a_i + u_{it}$

and letting $\varepsilon_{it} = a_i + u_{it}$ represent the composite error term, we can rewrite the proceeding equation as follows

(A.5) $\ln\left(y_{it} + \sqrt{y_{it}^2 + 1}\right) = X'_{it}\beta + \varepsilon_{it}.$

To solve for y_{it} we first exponentiate both sides of the equation

(A.6) $\exp\left[\ln\left(y_{it} + \sqrt{y_{it}^2 + 1}\right)\right] = \exp[X'_{it}\beta + \varepsilon_{it}],$

which produces the following result

(A.7) $y_{it} + \sqrt{y_{it}^2 + 1} = \exp[X'_{it}\beta + \varepsilon_{it}].$

Substituting in $w\ for\ X'_{it}\beta + \varepsilon_{it}$, subtracting y_{it} from the left hand side, and squaring both sides of the equation we obtain the following:

(A.8) $y_{it}^2 + 1 = \exp[w] * \exp[w] - 2\exp[w] * y_{it} - y_{it}^2.$

Next, by subtracting y_{it}^2 from both sides of the equation we can rearrange the equation, and solve for y_{it} as follows:

(A.9) $y_{it} = \frac{\exp[w]*\exp[w]-1}{2\exp[w]}$

or equivalently

(A.10) $y_{it} = \frac{\exp[w]}{2} - \frac{1}{2\exp[w]}.$

Next, pulling out the common factor of $1/2$ and substituting $X'_{it}\beta + \varepsilon_{it}$ back in for w we obtain the following expression for values of y_{it}:

(A.11) $y_{it} = \frac{1}{2}(\exp[X'_{it}\beta + \varepsilon_{it}] - \exp[-X'_{it}\beta - \varepsilon_{it}])$.

To obtain the predicted outcomes (i.e., values for $\hat{y}_{it}$) we proceed as before and take the expected value of the right-hand side, conditioning on the set of known covariates X_{it},

(A.12) $\hat{y}_{it} = E\left[\frac{1}{2}(\exp[X'_{it}\beta + \varepsilon_{it}] - \exp[-X'_{it}\beta - \varepsilon_{it}])|X_{it}\right]$.

Because $1/2$ is a constant its expected value is just its value. Moving the expectation operator through the equation, we obtain the following

(A.13) $\hat{y}_{it} = \frac{1}{2} * (E[\exp(X'_{it}\beta + \varepsilon_{it})] - E[\exp(-X'_{it}\beta - \varepsilon_{it})])$

Using the rules of exponents, we can rewrite the proceeding equation as follows

(A.14) $\hat{y}_{it} = \frac{1}{2} * (E[\exp(X'_{it}\beta) * \exp(\varepsilon_{it})] - E[\exp(-X'_{it}\beta) * \exp(-\varepsilon_{it})])$.

As before, because the variables X'_{it} are constant for each observation i at any given time period t, and we know the true values for the parameters of interest β, we can assume the value for $\exp(X'_{it}\beta)$ is also a constant and therefore, $E[\exp(X'_{it}\beta)] = \exp(X'_{it}\beta)$. The same holds true if we replace the true values of β with their estimates (i.e., $\hat{\beta}$).

Moreover, because $\exp(X'_{it}\beta)$ is a constant it is assumed to be independent of $\exp(\varepsilon_{it})$, we can separate the two terms and as a result, the equation for obtaining the predicted values for y_{it} becomes

(A.15) $\hat{y}_{it} = \frac{1}{2} * \left[\exp\left(X'_{it}\hat{\beta}\right) * E[\exp(\varepsilon_{it})\right] - \exp\left(-X'_{it}\hat{\beta}\right) * E[\exp(-\varepsilon_{it}))]]$.

We'd like to be able to write,

(A.16) $\hat{y}_{it} = \frac{1}{2} * [\exp\left(X'_{it}\hat{\beta}\right) - \exp\left(-X'_{it}\hat{\beta}\right)]$

Which looks like it should work considering the mean of the composite error term is assumed to be zero and $\exp(0) = 1$. The problem, however, is that

(A.17) $E[\exp(\varepsilon_{it})] \neq \exp(E[\varepsilon_{it}])$.

The expected value of the formula $\exp(\varepsilon_{it})$ depends on the assumed distribution of ε_{it} and because ε_{it} represents a composite error term we must consider the distribution of each part. Because we included an intercept it is safe to assume $E[u_{it}|X_{it}] = E[a_i|X_{it}] = 0$. Additionally, to solve for the predicted values of y_{it}, for simplicity we assume $[u_{it}^2|X_{it}] = \sigma_u^2$, $E[a_i^2|X_{it}] = \sigma_a^2$, and $E[u_{it}a_i|X_{it}] = 0$. As a result, the variance of the composite error term ε_{it} is equal to $\sigma_u^2 + \sigma_a^2$.

To solve for the $E[\exp(\varepsilon_{it})]$ we need to be aware of the following relationship between the normal and log-normal distribution. In general, if a variable $A \sim N(\mu, \sigma^2)$ then $B = \exp(A)$ is assumed to be $\sim log - N(m, v)$ where $m = \exp[\mu + {\sigma^2}/{2}]$. In our case $\mu = E[\varepsilon_{it}] = 0$ and $\sigma^2 = var\ (\varepsilon_{it}) = \sigma_u^2 + \sigma_a^2$. Therefore, we can solve for $E[\exp(\varepsilon_{it})] = \frac{1}{2}[\sigma_u^2 + \sigma_a^2]$. Now we can generate our predictions ("fitted values") for y_{it} using the following equation:

(A.18) $\hat{y}_{it} = \frac{1}{2} * ([\exp(X_{it}'\hat{\beta}) + .5\sigma_a^2 + .5\sigma_u^2] - [\exp(-X_{it}'\hat{\beta}) - .5\sigma_a^2 - .5\sigma_u^2])$

CHAPTER 3

ESSAY 2: ESTIMATING AND COMPARING EMPIRICAL MEASURES OF HOUSEHOLD ENERGY INSECURITY*

*Harker Steele, Amanda. To be submitted to *Contemporary Economic Policy*

ABSTRACT

While a large body of the applied economics literature has been dedicated to accurately measuring food insecurity, relatively little has been done so far to define and measure household energy insecurity. One reason household energy insecurity has been largely ignored is that no consistent and universally accepted index measure has been developed to measure it. The objective of this chapter is to help fill this gap in the literature by estimating and comparing alternative empirical measures of household energy insecurity. Using household responses from the 2015 Residential Energy Consumption Survey we compare and contrast five different classification procedures and empirical measures of household energy insecurity. We discuss the validity of each of the measures in terms of their ability to achieve construct, content, and convergent validity. We conclude that a measure of energy insecurity based on the Rasch model results provides a conceptually and empirically strong and valid measure of household energy insecurity. Therefore, the Rasch model results are used to create the preferred energy insecurity index, a four group energy insecurity index measure.

Keywords: Energy Insecurity, Empirical Measures Index, Rasch Model

3.1 INTRODUCTION

Being able to meet basic, daily household energy service needs is becoming exceedingly difficult for many families living in the United States (Hernandez 2016). According to the World Health Organization (WHO), just as food, water, and shelter are considered to be physiological requirements for human survival, having access to enough energy to provide adequate household energy services is considered to be an essential component for sustaining a healthy quality of life (WHO 2006). The consequences of not having access to adequate household energy services include being unable to keep indoor air temperatures at a “safe” level, increased susceptibility to illnesses, and reducing and forgoing expenditures on other basic household necessities such as food and/or healthcare in order to afford energy services (Thomson 2017).

Despite the hardships having inadequate access to energy services cause households, little to no attention has been paid to this issue in the applied economics literature (Hernandez 2016). One factor contributing to this lack of attention is the absence of a single, consistent, and universally accepted metric that can be used to quantify energy service-related hardships faced by individual households (Hernandez 2016). The objective of this chapter is to help fill this gap in the literature by estimating and comparing alternative empirical measures that are used to gauge whether or not a household is energy insecure.

In this chapter we compare and contrast five different approaches for assessing whether or not households face energy service-related hardships. We define the term energy service-related hardship as any circumstance that prevents a household from being able to maintain consistent and adequate access to basic, daily household energy service needs. Examples of energy service-related hardships include financial constraints, broken service equipment, or unforeseen circumstances beyond the control of the household.

We define the term “energy security” as a state where a household has consistent physical and economic access to a sufficient, safe, and affordable energy supply to meet each household members’ most basic daily energy service needs. [35] This definition of energy security is similar to the definition of “food security,” which refers to a state where individuals or households have consistent, reliable access to a sufficient quantity of affordable, nutritious food (Nord 2009). Households who are able to maintain consistent, reliable access to adequate energy services are considered “energy secure,” while households who are unable to maintain consistent, reliable access to adequate energy services are considered to be “energy insecure” (O’Mera 2016; Hernández 2016; Middlemiss and Gillard 2015).

Thus, the state of being energy secure implies the existence of little or no energy-related hardships (i.e., little or no circumstances that prevent a household from being able to maintain access to basic, daily energy service needs). The state of being energy insecure however, implies the existence of at least one to many energy-related hardships (i.e., several circumstances that prevent a household from being able to maintain access to basic, daily energy service needs). Based on our definition of energy security, households who face more energy-related hardships are considered to be worse off than households who face fewer energy-related hardships and therefore are considered to be less energy secure.

Consistent with the previous literature, we assume the term “energy services” can be used to describe any function performed inside the household that uses energy (e.g. electricity, natural gas, or propane) as an input to produce a desired output (Fell 2017; Fowlie, Greenstone, and Wolfram 2018). Examples of household energy services include maintaining safe, comfortable

[35] The definition provided for household energy security is based on the definition of food security as defined by the United States Department of Agriculture (USDA). More specifically, the USDA defines household Food Security as having access at all times to enough food for an active, healthy life.

indoor room temperatures, having access to cold or hot food and beverages, or sufficient indoor room lighting for leisure activities such as reading and/or writing. Energy services are generated from household appliances such as air conditioners, refrigerators, lamps, and stoves – and the operation of these appliances by members of the household.

Maintaining consistent access to household energy services has become increasingly difficult for many low-income and other economically marginalized households, including those with elderly and differently abled individuals living in the home (Wilkinson et al. 2004; Hernández 2013; Drehobl and Ross 2016; O'Mera 2016). From an economics point of view, low-income and other economically marginalized households may struggle to maintain consistent and reliable access to energy services due to poor housing conditions, financial constraints, and/or unreasonably high home energy costs. As a result, these types of households are overwhelmingly more likely to identify as being energy insecure (Hernandez 2016; 2014).

While the notion of "energy insecurity" exists in the literature, the phenomenon is not well understood (Murray and Mills 2012; Hernandez 2016). [36] Furthermore, a review of the household energy insecurity literature revealed inconsistencies in the way household energy insecurity is both defined and measured. For example, households who are energy insecure have historically been defined in other previous studies as households who are experiencing "fuel poverty," where 10% or more of a household's disposable income is being spent on fuel (Boardman 1991). Existing

[36] There have been numerous other terms used in the literature to describe a household's energy security status. For example, the terms "energy poverty" and "fuel poverty" have both been used to describe households who are energy insecure. The term fuel poverty has been used to describe situations where a household is unable to keep their home sufficiently warm at a reasonable cost, given their income. Conversely, the term energy poverty has been used to refer to a situation where a household is unable to access modern-day energy services (World Watch Institute 2019).

studies have also utilized the term “energy insecure” to refer to the United States’ reliance on foreign energy imports (International Energy Agency [IEA] 2019).

However, as Murray and Mills (2012) suggest, the IEA’s definition of energy insecurity disregards the importance on a household’s overall well-being of having access to enough energy resource inputs to produce adequate energy services. For example, when a household does not have access to energy resource inputs, it may be unable to properly heat and/or cool the home. Houses that are improperly heated are often cold and damp, and as a result are more susceptible to mold and dust which can activate or exasperate respiratory conditions in members of the household (Dear and McMichael 2011).

To understand more fully the different dimensions of household energy insecurity, we estimate and compare empirical measures of energy insecurity using responses from the 2015 Residential Energy Consumption Survey (RECS). In Section 3.2 we first explore several prior published techniques that have been used to determine whether or not a household faces any energy-service related hardships (i.e., is energy insecure). In Section 3.3 we outline our general methodological approach to determine whether or not a household is energy secure or energy insecure.

Section 3.4 discusses the data used for our analysis in detail. Section 3.5 outlines the five different techniques we employ to determine a household’s energy security status. Section 3.6 presents the model results from the five techniques described previously in Section 3.5. In Section 3.6 we compare and contrast the results from applying all five approaches, and discuss the content, construct, and convergent validity of our different results. We conclude this chapter (essay) in Section 3.7.

3.2 MEASURING HOUSEHOLD ENERGY SECURITY

Historically, a household's energy security status (i.e., whether they are energy insecure or energy secure) has been determined using three main approaches: (1) the expenditure approach; (2) the prediction approach; and (3) the self-report, subjective survey approach (O'Mera 2016). The expenditure approach is directly linked to the amount of money a household spends on energy/fuel. Households who spend more than 10% of their disposable income on energy/fuel are considered to be energy insecure (Boardman 1991). However, as suggested by Hills (2012) this method directly excludes low-income households that might end up spending less than 10% of their disposable income on energy/fuel simply because they are unable to afford to spend more.

The prediction approach estimates how much money a household would need to spend on energy/fuel to achieve a "livable" indoor air temperature, as suggested by the WHO. According to the WHO, suggested "livable" indoor air temperatures are 21°C (69.8°F) in the living room and 18°C (64.4°F) in any other occupied rooms. If spending the required amount leaves the household with a remaining income that places the household below the designated national poverty line, then the household is considered to be energy insecure (O'Mera 2016; Department of Energy and Climate Change [DECC] 2013). [37] While this method directly avoids excluding households who spend less than 10% of their income on fuel costs but are still energy insecure, it still assumes one uniform level of comfort exists across households of a similar size and make-up, such that household's preferences for indoor air temperatures are homogenous.

The third approach used to determine a household's energy insecurity status is the self-report, subjective survey approach wherein households respond to questions directly related to

[37] The forecasted required amount is based on the size and demission of the dwelling, the price of fuel inputs, the household's income, and the median level of fuel expenditures for a household of a similar size.

their ability or inability to maintain consistent access to adequate household energy services. For example, previous large-scale surveys such as the European Union's Survey on Income and Living Conditions (SILC) and the Spanish Living Conditions Survey (SLCS) have included a subset of questions that specifically ask households to report on circumstances they are currently facing or have faced that have prevented them from maintaining access to adequate energy services (O'Mera 2016).

Questions in the SILC specifically ask households to report on whether or not they have had to go without heating or cooling in the home for an extended period of time, been unable to afford to heat or cool their home "adequately," or afford their home energy bills (O'Mera 2016). Households who respond affirmatively to the questions are considered to be energy insecure. However, as Watson and Maitre (2015) point out, one potential issue with using the self-report, subjective survey approach to determine a household's energy security status is that it relies on an individual's perception of their personal, current energy security situation.

This type of subjective indicator could lead to households with higher incomes reporting they are energy insecure, even though they do not necessarily lack the financial resources to afford their home energy bills and/or maintain consistent access to adequate household energy services. For example, in some cases, higher income households may experience difficulty affording energy/fuel expenses due to other constraints on their income such as debt payments (i.e. car payments, mortgages), higher household maintenance costs, and health care expenditures (O'Mera 2016). These households might respond affirmatively to questions related to struggling to afford their home energy bills and therefore, be considered "energy insecure."

In addition, whatever the cause, some people just naturally do not feel as secure as others. For example, from an objective perspective a person may actually have adequate food, water, and shelter, but still does not feel secure from a personal, subjective perspective. While the three approaches outlined above provide a useful way to place households in binary energy security/insecurity categories (e.g., “1 = energy secure; 0 = energy insecure”) the development of an index that provides a continuum of household energy insecurity “ratings” would provide more information on the relative levels of household energy security/insecurity that exist on a larger scale (Murray and Mills 2012).

We are currently only aware of three prior attempts to define and create a specific index to measure the extent/level of energy insecurity being experienced in the United States. In the first study Colton (2003) created a Home Energy Insecurity Scale (HEIS) in an effort to estimate the extent to which an energy assistance program could help improve a household’s energy security status. The HEIS contains five different energy security status thresholds: 1) Thriving; 2) Capable; 3) Stable; 4) Vulnerable; and 5) In-Crisis. [38] Households are placed at different thresholds, along the HEIS based on their responses to a set of survey questions. The survey questions ask households to report on “how often” they have had to adjust their energy use consumption patterns over the past twelve months due to financial strain (Colton 2003).[39]

Along the HEIS, thriving households are households who can participate in the full range of home energy uses without needing outside assistance and without financial strain. A capable household is a household that may have some energy-related debts because they cannot afford to

---

[38] In their simplest form, a threshold is defined as individual points along the HEIS. Thresholds can be thought of as benchmarks along a scale that indicate a household’s current energy security status (Colton 2003).

[39] Colton (2003) relies on household responses to questions on the HEIS survey. Questions included in the HEIS survey are adapted from questions used by the U.S. Department of Agriculture (USDA) to measure “food insecurity” in the United States (Hamilton et al. 1997). Once the Home Energy Insecurity Scale survey has been completed by each household, households are assigned to an energy security status threshold based on their individual responses. For example, all respondents found to exhibit the indicators of a “thriving” household are assigned to the “thriving” threshold (Colton 2003).

pay their energy bills, but debts are only accumulated on occasion. Stable households are households who may have a need to access outside assistance to pay current energy bills, but this need does not arise more than "sometimes." Energy-related debts are accumulated by stable households more frequently than they are accumulated by capable households.

Vulnerable households are households who occasionally face energy-related choices that require members of the household to compromise not merely on comfort and/or convenience, but also on basic household energy needs such as food and/or medicine. Vulnerable household accumulate energy-related debts more frequently and debts are often large. An in-crisis household is a household who faces immediate energy needs that threaten the household's physical and/or emotional safety. An in-crisis household may experience recurring periods of going without home energy services. In-crisis households routinely engage in energy use choices that compromise other basic household needs (Colton 2003).

Based on the household's pattern of responses to the questions posed in the Home Energy Insecurity Scale Survey, the household is placed at one of the five thresholds. Because outcomes along the HEIS are defined in terms of changes in a household's energy security status from receiving energy assistance, placing households along the scale requires information to be collected before receiving energy assistance (to establish baseline within the data) and after receiving energy assistance to determine to what extent, if at all, the household energy security status has changed as a result of receiving energy assistance.

Thresholds are defined such that households either fit or do not fit within a specific threshold based on their responses to the individual questions. If a household does not fit into one of the thresholds, it is eliminated from that threshold. This process repeats itself until all households have been sorted into at least one threshold. However, as Murray and Mills (2012) point out, while the development of the HEIS represents a significant contribution to the literature, the broad nature of the five thresholds used to measure energy insecurity is a limitation. Also, the HEIS does not derive a numerical measure for the thresholds used to represent levels of energy insecurity. Therefore, the results can be difficult to interpret (Division of Energy Assistance 2008).

In the second study, Cook et al. (2008) developed an energy insecurity index using responses to a set of four questions from a household survey administered as part of the ongoing Children's Sentinel Nutrition Assessment Program (C-SNAP). Households who answered "no" to all of the questions were considered "energy secure." Households who answered "yes" to only the first question were considered to be moderately energy insecure, whereas households who answered yes to any additional question(s) (questions 2, 3, or 4) were categorized as severely energy insecure. While notable, the index developed by Cook et al. (2008) is based solely on the number of affirmative responses by households to a set of four questions.

As Balistreri (2016), Dutta and Gundersen (2007) and Gundersen (2008) point out, an index based on the number of affirmative responses at an aggregate level neglects to take full advantage of the information available in each specific question asked. [40] Additionally, using only the number of affirmative responses to define an energy security index could disregard several other key factors that could strongly indicate the presence of energy insecurity in the household. For example, households who are energy insecure may keep indoor air temperatures at unsafe levels to avoid high electric utility bills. As a result, they may be forced to seek medical attention.

[40] The objective of the paper by Gundersen (2008) is to measure the extent of food insecurity being experienced by households across the United States. As Gundersen (2008) points out, if the objective when examining food insecurity is to determine the level of severity of "food insecurity" being experienced by a household, then solely counting the number affirmative responses assumes households who respond affirmatively to having skipped meals are no more food insecure than households responded affirmably to not eating balanced meals.

Therefore, responding affirmatively to a question of this nature should bear more weight than responding "yes" to whether or not a household has received a shut-off notice.

Inspired by the creation of the food insecurity index developed by the USDA (Hamilton et al. 1997), in the third study Murray and Mills (2012) generate a household energy insecurity index using two different types of Rasch models: a Dichotomous Rasch model and a Polytomous Rasch model. They rely on data from the 2005 Residential Energy Consumption Survey (RECS) and generate an energy insecurity index that is consistent with the USDA food insecurity index. While their approach is notable, as Murray and Mills (2012) point out, questions from the 2005 RECS data set are quite broad.

Moreover, not all households surveyed in the 2005 RECS had a chance to respond to questions about their energy insecurity status. Instead, only households who were eligible to receive funding from the Low-Income Home Energy Assistance Program (LIHEAP) or whose incomes were less than 150% of the federal poverty level were asked to respond. As a result, the energy insecurity index developed by Murray and Mills (2012) is limited in scope. Furthermore, the primary objective of the analysis by Murray and Mills (2012) was to compare and contrast the severity parameter estimates (i.e., item calibrations) produced from the Polytomous Rasch model and Dichotomous Rasch model.

While they classify households as being "energy secure" or "energy insecure," based on severity parameter estimates produced from the two models, Murray and Mills (2012) make no attempt to estimate the extent or overall severity of energy insecurity being experienced by the larger population. They fail to utilize suggestions from Dutta and Gundersen (2007) and Gundersen (2008) to explore the information contained in the individual questions in more detail to create an aggregate index measure. Lastly, Murray and Mills (2012) only considered two statistical analysis techniques to create their energy insecurity index, a Dichotomous Rasch model and a Polytomous Rasch Model. We consider five.

Their use of the Polytomous Rasch model was motivated by the only response options available to households in the 2005 RECS being "how often." In more recent iterations of the survey (i.e., the 2009 and 2015 iterations) the response options available to households are a combination of yes/no and "how often." Therefore, the Polytomous Rasch model is not applicable. To address the need for an updated index measure of household energy insecurity, we generate an energy insecurity index using data from the 2015 RECS, considering five separate methods used previously to determine whether or not a household is energy secure or energy insecure. An outline of our general methodological approach, considering each of the five different alternatives, is outlined below.

3.3 METHODOLOGICAL APPROACH

Following the literature on food security, one way we could determine whether or not a household is energy insecure is by assessing whether not the household is consuming enough fuel to provide a standard level of household energy services. In the case of a household whose primary fuel source is electricity, we would need to determine the total number of kilowatt-hours (kWh) of electricity necessary to provide a standard level of household energy services. For example, we would need to determine how many kWh are required to produce a standard number of cooked meals, reach a standard indoor air temperature, or generate a standard amount of clean laundry.

For households who consume fewer kWh of electricity than deemed necessary, the extent of a household's level of insecurity (i.e. lack of security) could be inferred by calculating the difference between the number of kWh actually consumed by the household and the number of kWh necessary to provide this pre-determined standard level of energy services. However, based on our dataset, it is only possible for us to observe the number of kWh actually consumed by the household. As a result, we are not able to determine whether or not a household is energy secure or energy insecure by examining the number of kWh consumed.

Therefore, in an attempt to accurately gauge each household i's level of energy insecurity we use a proxy, a unique index measure of household energy insecurity. To construct our energy insecurity index, we first let $n = \{1,2,\ldots,N\}$ denote a set of households, such that n is equivalent to the total number of households within a given set (Dutta and Gundersen 2007; Gunderson 2008; Balistreri 2016). For each household $i \in N$, we establish a value s_i which is used to denote any "energy service-related hardships" faced by the individual household i, such that a higher value of s_i corresponds with more energy service related hardships being experienced by members of the household.

As mentioned in Section 3.1, the term "energy service-related hardship," refers to any circumstance or situation inside the home that might prevent the household's from being able to produce and consume its adequate level of energy services (i.e., maintain access to heating and cooling, cook meals at home). Therefore, based on our definition of household energy security, more energy service-related hardships within a household correspond with higher levels of household energy insecurity. Furthermore, we assume that for all $i \in N$, values for s_i lie within the interval $[0, S]$, such that a value 0 denotes the complete absence of any energy service-related hardships (i.e., no energy insecurity), while a value of S denotes the existence of multiple energy service related hardships (i.e., the most energy insecurity) (Balistreri 2016).

To gauge each household i's energy security status based on its ability to maintain access to an adequate level of energy services, we construct individual values for s_i using responses by households to questions included in the 2015 RECS. The questions used to construct individual values of s_i change depending on which of the five metrics are used. In general, individual values for s_i are constructed as follows:

(1) $s_i = f(a_j) \quad \forall\, j = 1, \dots, J$

where s_i is used to represent the energy service related hardships faced by the individual household i; a_j is used to represent the set of J questions from the 2015 RECS used to construct the individual index measure; and $f(\cdot)$ is a general "function" used to represent one of the five different techniques we apply in this chapter to construct the energy insecurity index: the expenditure approach, whether the household has received home energy assistance, cluster analysis, principal components analysis, or a Dichotomous Rasch model.

To generate the energy insecurity index based on values of s_i produced from using the different methods, we first have to establish a benchmark (i.e., a cut point) τ such that a household is considered to be energy insecure if and only if $s_i > \tau$ and energy secure if and only if $s_i \leq \tau$ (Gundersen 2008). In this case, τ can be thought of simply as the dividing line between what energy service related hardships experienced by households separate energy secure households from energy insecure households.[41]

In general, for each household i, the energy insecurity index which we label as $EISINDEX_i$ produced by any of the five different methods is defined to be 0 if $s_i \leq \tau$ and defined to be equal to $(s_i - \tau)$ if $s_i > \tau$.[42] Therefore, an energy insecurity index value equal to zero ($EISINDEX_i = 0$) is an indication that the household has experienced fewer or exactly the number of energy service-related hardships implied by the threshold value of τ and therefore, is considered to be

[41] As Gundersen (2008) points out, τ is akin to the federal poverty line, which divides household's in poverty from households not in poverty.
[42] Similar to how incomes of households above the poverty line are not reflected in measures of overall poverty, the energy security status of energy secure households is not reflected in the energy insecurity measure used (Gunderson 2008).

"energy secure." An energy insecurity index value equal to something besides zero implies the household has faced more energy service related hardships than the number of energy service related hardships implied by the value of τ and therefore, is considered to be "energy insecure."

As the extent of energy insecurity experienced by the household increases, the greater is the distance between s_i and τ. Thus, similar to how the food insecurity index is constructed, the constructed energy insecurity index $EISINDEX_i$ measures the degree to which a household is energy insecure, rather than energy secure. What is left to choose is an appropriate threshold value of τ, which will depend on which of the five metrics is used to determine a households energy security status. [43] Once a threshold value of τ has been determined, households can divided into mutually exclusive energy insecurity categories based on the individual values of s_i they receive. The different energy insecurity categories generated can then be used to produce the energy insecurity index $EISINDEX_i$ value for each household.

3.4 DATA

Data for constructing our energy insecurity index comes from responses by households to a subset of questions included in the 2015 Residential Energy Consumption Survey (RECS). The RECS is a national multiphase survey administered by the United States Energy Information Administration (U.S. EIA) once every three years. Households surveyed in the RECS are assumed to provide a representative sample of the U.S. population as a whole. The RECS samples only households that are occupied as primary residences. It therefore excludes secondary homes (i.e., vacation homes), vacant units, and military barracks.

[43] For example, following the literature on food security when applying the Rasch Model to distinguish between households that are energy insecure or energy secure, a threshold value of τ equal to zero is most appropriate, because it implies zero affirmative responses to any of the survey questions that indicate any energy related hardships have being experienced in the home (Hamilton et al. 1997; Gundersen 2008). Gundersen (2008) describes an unfavorable food situation in the home as the existence of a struggle to maintain access to enough food due to circumstances faced by the household.

Information from the 2015 RECS was collected from respondents via telephone surveys and web-based questionnaires. The surveys and questionnaires are identical and asked households to reveal information about energy use consumption patterns. We specifically utilize data from Sections K and L of the 2015 RECS. Section K includes information on household level characteristics including household size, income, age, and employment status. Information was collected from Section L of the 2015 RECS because this section of the survey specifically asks households to report on challenges they may have faced over the past twelve months in paying their energy bills or maintaining access to heating and/or cooling inside the home.

We interpret these questions as indicators that a household was unable to produce an adequate level of energy services and therefore is energy insecure, to some degree. Table 3.1 below provides a complete description of the seven questions used to capture energy insecurity.

Table 3.1 2015 Residential Energy Consumption Survey Questions (Section L)

2015 RECS SECTION L: ENERGY INSECURITY and ASSISTANCE	**Question/Item Label**
1. In the last year, how many months did your household reduce or forego expenses for basic household necessities, such as medicine or food, in order to pay an energy bill?	Reduce
2. In the last year, has your household kept your home at a temperature that you felt was unsafe or unhealthy?	Unsafe
3. In the last year, how many months did your household received a disconnection notice, shut off notice, or non-delivery notice for an energy bill?	Notice
4. In the last year, was there ever a time your household was unable to use your main source of heat or air conditioning because you could not afford the fuel source and it was disconnected?	No Fuel
5. In the last year, was there ever a time your household was unable to use your main source of heat or air conditioning because equipment was broken and you couldn't afford to pay to repair or replace the equipment?	HVAC
6. In the last year, has anyone in your household needed medical attention because your home was too hot or too cold?	Medical

7. About how many days over the past year, has your household gone without heat and/or air conditioning over the past year?	Days

Unlike prior iterations of the RECS, all households surveyed in the 2015 RECS had the opportunity to respond to the questions included in Table 3.1. [44] Observations with missing responses for one or more of the seven energy assistance questions included in Table 3.1 were dropped from the final RECS data set.

Additionally, response choices for questions in the 2015 RECS varied. For questions 1, 2, and 3 households were instructed to state "how often" over the past twelve months they have dealt with an energy-related hardship described in the questions. Response options for questions 1, 2, and 3 included "never," true for "some months," true for "only one or two months," or true for "almost every month." Response options for questions 4, 5, and 6 included only "yes" or "no." Question 7 asked households to report on the number of days over the past twelve months they went without heating or cooling.

To create the energy insecurity index, we collapse responses to questions 1 to 7 into dichotomous questions and coded responses as binary variables following the standard methods used in the food security literature (Bickel et al. 2000). More specifically, for questions requiring a "how often" response, a response of "almost every month" or "some months" were coded as affirmative (value = 1), while a response of "only 1 or 2 months" or "never" were coded as negative (value = 0). For yes/no responses, "yes" was coded as a 1 and "no" was coded as a 0. For responses

[44] In prior iterations of the RECS, only households who were eligible to receive funds from the LIHEAP were asked questions related to energy insecurity (Murray and Mills 2012).

on the number of days, “days ≥ 36” were coded as a 1 and “days < 36” were coded as a 0. The data structure for questions 1 through 7 is illustrated in Table 3.2.

Table 3.2 The Data Structure

	Responses to Questions						
Respondent	**Q1** Reduce	**Q2** Unsafe	**Q3** Notice	**Q4** No Fuel	**Q5** HVAC	**Q6** Medical	**Q7** Days
1	0	1	1	0	1	0	0
2	0	0	0	0	0	1	1
3	1	0	1	0	0	0	0
4	1	1	0	1	1	1	0
5	0	1	0	0	0	1	1
6	0	1	1	1	1	0	0
7	0	0	0	1	1	0	1
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
n	1	1	0	0	0	1	0

Because the energy insecurity questions are all negative in nature, households who respond affirmatively more frequently are considered to be more energy insecure (less energy secure) than households who respond negatively more frequently. Because there are only seven questions, the total number of affirmative responses by a household, based on recoding the questions as dichotomous choice questions, can range from zero to seven. The precise severity level of each question is unknown *a priori*.

3.5 ENERGY INSECURITY MODELS

To create a consistent energy insecurity index, we begin by comparing and contrasting five different classification procedures used to determine if a household is energy insecure or energy secure: 1) whether or not the household has applied for and received home “energy assistance” to help pay energy bills and/or cover expenses to fix broken HVAC equipment; 2) whether or not, over the past year the household has spent more than 6 or 10 percent of their disposable income on fuel/energy (i.e., the expenditure approach); 3) Cluster Analysis; 4) Principal Components Analysis; and 5) a Dichotomous Rasch model. Because not all of the questions included in 2015 RECS required “how often” responses, we do not explore the Polytomous Rasch Model suggested and explored previously by Murray and Mills (2012).

ENERGY ASSISTANCE

The first metric used to determine if a household is energy insecure or energy secure is based on whether or not the household applied for and received “energy assistance.” Using this metric, we consider households to be energy insecure if they received home energy assistance to help restore heating and/or cooling or to help fix broken HVAC equipment sometime over the past twelve months. To determine the energy insecurity status of all household's surveyed, we focus on individual household's responses to the following three questions in the 2015 RECS:

1. Has your household participated in a home energy assistance program that helps pay energy bills or fix broken equipment?
2. Did your household apply for and receive home energy assistance to help pay your energy bill as a result of receiving a disconnection notice?

3. Did your household apply for and receive home energy assistance to help restore your heating or cooling because either your equipment was broken and you could not afford to pay for the repair/replacement or you could not pay for the fuel inputs, and as a result were disconnected from service?

We separated households into mutually exclusive groups based on their responses to the proceeding questions. We examined overall responses from households and also differentiated household responses by housing type. In terms of the general methodological approach described in Section 3.3, a household is said to be facing energy service-related hardships if it responded "yes" to any of the three questions listed previously. The energy service related hardships, s_i faced by the individual household are constructed as binary response variables as follows:

(2) $$s_i = \begin{cases} = 1 \text{ [Household provided a "Yes" reponse to 1 or more of the 3 questions]} \\ = 0 \text{ [Otherwise]} \end{cases}.$$

Using the receipt of energy assistance as a proxy for energy service-related hardships, the energy insecurity index $EISINDEX_i$ index is constructed as follows,

(3) $$EISINDEX_i = \begin{cases} = 1 \text{ if } s_i = 1 \\ = 0 \text{ if } s_i = 0 \end{cases}.$$

Here households who receive an energy insecurity index value, $EISINDEX_i$, equal to one are considered to be energy insecure, while household who receive and energy insecurity index value, $EISINDEX_i$, equal to zero are considered to be energy secure.

We interpret questions 1, 2, and 3 above as indicators of whether or not the household participated in the LIHEAP or some other type of home energy assistance program. The LIHEAP is a federally funded program that provides financial assistance to help households pay to heat and/or cool their homes (U.S. Department of Health and Human Services 2018). Funds from the LIHEAP are paid directly to utilities, on behalf of households who apply for and receive assistance.

The amount of financial assistance a household is eligible to receive depends on the size of the household, the household's overall income, and the type of fuel used to heat and/or cool the home.[45] To be eligible to receive energy assistance from the LIHEAP, the household must have an income that is below 150 percent of the Federal Income Poverty Guideline (Department of Health and Human Services 2018).

Since its inception in 1981, the LIHEAP has helped millions of low-income families pay their home energy bills (Edison Electric Institute [EEI] 2018). However, funds in general are limited and have decreased substantially over the past five years (United States Department of Health and Human Services 2017). For example, in 2011 funding for the Low-Income Home Energy Assistance Program (LIHEAP),[46] was $4.7 billion (USD). Since then, funding has been cut by nearly 25% and the projected budget for LIHEAP in 2019 is estimated to be only $3.69 billion (USD) (Edison Electric Institute [EEI] 2018). [47] Furthermore, unlike some other federal assistance programs, for example Medicaid, simply being eligible for the LIHEAP does not guarantee a household will receive benefits (Perl 2018). Benefits are limited by the amount that Congress appropriates each year. The number of households who receive benefits depends both on the total appropriations and how states decide to disperse funds (Perl 2018). [48]

While the receipt of home energy assistance provides a proxy for being energy insecure, as pointed out by Murray and Mills (2012), the share of households who receive home energy

---

[45] To be eligible to receive assistance the household must have an income that is below 150 percent of the Federal Income Poverty Guidelines (Department of Health and Human Services 2018).
[46] LIHEAP IS a federally funded program that helps low-income, impoverished households pay for heating or cooling and make home energy efficiency upgrades (i.e., weatherization).
[47] Furthermore, it is estimated that only one in five households who are eligible to receive assistance from LIHEAP are actually able to receive it, indicating the need for energy assistance far exceeds the available supply.
[48] The number of households who apply for the LIHEAP consistently exceeds the number of households for which funding is available (DeNavas-Walt and Proctor 2014; Perl 2018).

assistance (i.e., participate in the LIHEAP) are disproportionately non-rural. [49] Recent work by Ross, Drehobl, and Stickles (2018) finds that compared to their metropolitan counterparts, home energy burdens (defined as the percentage of disposable income a household spends on its energy bills) faced by rural households are much higher. [50] Thus, it appears that rural households who need energy assistance the most (i.e., are more energy insecure) are not applying for or participating in an energy assistance program. As a result, using participation in a home energy assistance program to determine a household energy insecurity status could provide an inaccurate representation of the true severity of energy insecurity across the United States.

THE EXPENDITURE APPROACH

The second metric used to determine a household's energy insecurity status is the expenditure approach. As stated earlier, the expenditure approach to measuring household energy insecurity is directly related to the amount of money a household spends on fuel (Boardman 1991). According to the expenditure approach, households who spend above a certain percentage of their disposable income on energy/fuel costs are considered energy insecure (O'Mera 2016).

Following Boardman (1991) we assume the threshold beyond which expenditures on energy/fuel cease to be affordable is 10 % of net disposable income. Therefore, households who spend more than 10% of their annual income on energy/fuel expenditures are considered to be energy insecure while households who spend less than 10% of their annual income on energy/fuel are considered to be energy secure.

---

[49] Rural households across the United States tend to live farther away from city centers (Scally et al. 2018). As a result, applying for energy assistance to restore heating or cooling may be more difficult.

[50] Rural households have a median home energy burden of 4.4%, compared to the national average of 3.3% (Ross, Drehobl, and Stickles 2018).

Using data from the 2015 RECS, we determine a household's energy security status as follows:

$$(4) \qquad EISINDEX_i = \begin{cases} = 1 \text{ if} \left(\frac{Annual\ Fuel\ Expenditures\ (\$)}{Median\ Annual\ Income\ (\$)} * 100\right) \geq 10 \\ = 0 \text{ if} \left(\frac{Annual\ Fuel\ Expenditures\ (\$)}{Median\ Annual\ Income\ (\$)} * 100\right) < 10. \end{cases}$$

Here the energy service related hardships, s_i faced by the individual household are represented by the proportion of annual income the household spends on energy/fuel as represented by,

$$(5) \qquad s_i = \left(\frac{Annual\ Fuel\ Expenditures\ (\$)}{Median\ Annual\ Income\ (\$)} * 100\right).$$

If the value of $s_i \geq 10$, then the household is said to be facing energy related hardships that are preventing it from maintaining consistent access to energy services. Under this framework, households who receive an energy insecurity index value, $EISINDEX_i$, equal to one (i.e., households who spend more than 10% of their income on energy/fuel) are considered to be energy insecure, while household who receive and energy insecurity index value, $EISINDEX_i$ equal to zero are considered to be energy secure.

Following suggestions by Chandler (2016), we also consider the case where a household's expenditures on energy/fuel exceed 6% of their net disposable income. We consider households who spend more than 6% of their annual income on energy/fuel expenditures to be energy insecure while households who spend less than 6% of their annual income on energy/fuel are considered to be energy secure. Under this case, the energy security status of a household is determined as follows:

$$(6) \qquad EISINDEX_i = \begin{cases} = 1 \text{ if} \left(\frac{Annual\ Fuel\ Expenditures\ (\$)}{Median\ Annual\ Income\ (\$)} * 100\right) \geq 6 \\ = 0 \text{ if} \left(\frac{Annual\ Fuel\ Expenditures\ (\$)}{Median\ Annual\ Income\ (\$)} * 100\right) < 6. \end{cases}$$

Again, under the framework specified above, households who receive an energy insecurity index value, $EISINDEX_i$, equal to one are considered to be energy insecure, while household who receive and energy insecurity index value, $EISINDEX_i$ equal to zero are considered to be energy secure. Using the expenditure approach as outlined above is the preferred metric for measuring household energy insecurity in the United Kingdom and the Republic of Ireland (Boardman 2010).

However, as pointed out by Hills (2012) using the expenditure approach to measure household energy insecurity can directly exclude households who spend less than 10 percent of their income on fuel expenditures, simply because they are unable to afford to spend more. For example, it may be the case that even if the amount of the household's fuel bill exceeds 10 percent of their overall income, the amount the household actually ends up paying toward the bill is less than 10 percent of their disposable income because they can only afford to pay a fraction of the cost of their total utility bill. Therefore, if the expenditure approach is used to determine a household's energy security status, households who spend below the threshold of 10% because they are unable to afford to spend more would be considered to be energy secure.

Furthermore, as pointed out by Hills (2012) and Drehobl and Ross (2016) using the expenditure approach to measure household energy insecurity does not account for specific household characteristics, such as differences in household composition (i.e., number of people living in the home), square footage of the household, or the thermal energy efficiency rating of the residence occupied. For example, families who rent might consume more fuel on average than families who own residences of a similar size and make-up because the decision to make an in-home energy efficiency investment to reduce fuel consumption (i.e., purchasing an *Energy Star®* certified appliance or adding insulation) is beyond their control (Carliner 2013).

Also, when measuring household energy insecurity using the expenditure approach, it is important to control for the composition of the household, especially when determining what amount of money is “reasonable” for a household to have to spend on energy/fuel to provide energy services (Hills 2012). Similar to how households with more individuals living in the home will require more food to be considered "food secure," households with more individuals living in the home will likely need to consume more energy/fuel to provide the level of household energy services that results in the household being energy secure. In addition, residences with more square feet will likely require more energy/fuel to provide the same level of warmth that is achievable in a smaller household while consuming less energy/fuel.

Thus, when determining a household energy security status, it is critical to consider the fuel efficiency rating of a household, the household’s size and composition, and the amount of money a household spends on energy/fuel costs. Failure to control for household level characteristics could lead to biased results. Therefore, in addition to using the expenditure approach to determine a household’s energy security status, we also consider three additional multivariate techniques which allow us to control for household level characteristics: (1) Cluster Analysis; (2) Principal Components Analysis (PCA); and (3) a Dichotomous Rasch Model.

These three techniques examine how households responded to the questions from the 2015 RECS included in Table 3.1. The idea underlying the use of responses to multiple questions to determine a household’s energy security status is that, just as in the case of food insecurity, no single question can accurately portray the full concept of energy insecurity (Gundersen 2008). Each of the questions included in Table 3.1 are believed to reveal some aspect of household energy insecurity. Therefore, by examining each household’s pattern of responses to the questions, we should be able to understand more clearly what factors contribute to a household identifying as being energy insecure versus being energy secure.

A general question faced by practitioners of applied economics who use surveys such as the RECS, is how to effectively organize data from the survey into meaningful structures. Surveys, by their nature, result in data structures that are multivariate (Abeyssekera 2003). That is, all of the data collected in a survey, by construction, should be related to the individual outcome variable of interest. The use of multivariate methods to analyze survey data allows for a deeper exploration into possible patterns that might exist within the data (Abeyasekera 2003). Multivariate methods also allow many variables to be considered simultaneously (Abeyasekera 2003).

CLUSTER ANALYSIS

The third method we used to determine a household's energy security status is cluster analysis. Cluster analysis is an exploratory, statistical analysis classification technique which aims to divide observations from a single data set into different groups, such that the degree of association between observations in the same group is maximized and minimized otherwise. Groups are referred to as "clusters." While the use of cluster analysis can lead to the discovery of different structures within the data, cluster analysis provides little to no explanation of why the individual clusters identified exist. The only thing assumed by the application of cluster analysis is that observations in an individual cluster are more similar to one another than observations in any other cluster (Everitt et al. 2001). [51]

To generate an energy insecurity index using cluster analysis, two decisions have to be made. First, one must determine the distance measure to be used which specifies the degree of

[51] For more information on the methods of cluster analysis please refer to the appendix.

similarity or dissimilarity desired between observations within an individual cluster. The distance measure is a numerical value reflecting the degree of "closeness" between each pair of observations. In our case, similarity measures for binary responses to the questions included in Table 3.1 are based on four different values from the cross-tabulation of responses to the questions by individual household's i and j.

The cross-tabulation for comparing individual households i and j is displayed below in

		Observation j	
		1	0
Observation i	1	w	x
	0	y	z

Figure 3.1 Cross-Tabulation for Matching Households

Here w is the number of questions for which individual households i and j both responded in the affirmative (i.e., both received a value of 1); z is the number of questions for which individual households i and j both did not respond affirmatively (i.e., both received a value of zero); x is the number of questions which individual household i responded in the affirmative but individual household j did not; and y is the number of questions which individual household j responded in the affirmative but individual household i did not (Abeyaskera 2003).

Given the survey responses being considered for our analysis are coded as binary variables, we rely on the matching technique discussed in Zubin (1938) where the degree of similarity between individual observations is determined by the proportion of matches between any two observations in the same cluster. The degree of similarity between any two observations is calculated as follows:

(7) $\frac{w+z}{w+x+y+z}$.

Once the distance measure has been determined, what is left is to determine which method of clustering is most appropriate given the structure of our data (Abeyaskera 2003). Because responses to the questions in Table 3.1 are coded as binary variables, to create an energy insecurity index using cluster analysis, we utilize hierarchical clustering. With hierarchical clustering, it is assumed at the start that each observation (household) belongs to its own unique cluster. Households are assumed to be clustered individually. Individual clusters are combined sequentially with other clusters based on the degree of similarity between the households in the individual clusters. The partitioning of households into different clusters continues until the predesignated number of clusters as specified by the researcher is met, or until no observations change groups.

Based on how food insecurity is measured in the United States, we set the predesignated number of clusters to four. [52] As a result, we have four different energy insecurity index ($EISINDEX_i$) categories. By applying hierarchal clustering, we were able to separate households into mutually exclusive clusters (i.e., groups) such that each cluster represents a distinct energy insecurity group. The energy related hardships s_i faced by individual households in the same cluster are assumed to be similar and are based on household responses to the questions included in Table 3.1.

[52] The latent food security status of individual household is assumed to lie along a continuum, which extends from high food security to very low food security. This continuum is divided into four mutually exclusive categories: (1) high food security; (2) marginal food security; (3) low food security; (4) very low food security.

PRINCIPAL COMPONENTS ANALYSIS

The fourth method we used to determine a household's energy insecurity status is principal components analysis (PCA). PCA is a statistical analysis technique that considers how different individual variables can be linked together to measure a single outcome variable of interest. The outcome variable of interest in our case is whether the household is energy secure or insecure. PCA is normally used when collecting first-hand data on the outcome variable of interest is outside the scope of the project, either because it is too time-consuming or requires extensive resources.

In our case, PCA is applied to the seven questions included in Table 3.1. The seven questions are assumed to contain information about specific conditions, experiences, and behaviors that indicate a household is energy insecure. Responses to the seven questions are combined together to generate values of s_i for each individual household i as follows:

(8) $$s_i = f(a_1, a_2, \dots, a_7) \quad \forall\, j = 1, 2, \dots, 7$$

Recall, values of s_i are used to represent the energy service-related hardships faced by each individual household, such that higher values of s_i correspond with more energy service-related hardships.

Because each of the seven questions included in Table 3.1 measures some facet of household energy insecurity (i.e., an inability of the household to produce and consume an adequate level of household energy services), it is assumed that responses to the questions are likely to be correlated with one another. The application of PCA attempts to capture this correlation and use it to establish a set of weights for each of the questions of interest. These weights are then used to create individual principal-components such that each individual principal component represents a different linear weighted combination of the initial variables of interest (Vyas and Kumaranayake 2006).

All principal components produced contain the same information as the original variables of interest. However, the information is partitioned over the components in a particular way such that earlier components contain more information than later components. By using the PCA results to generate values for s_i, we are assuming the reason households respond differently to the questions of interests is that the energy service-related hardships they face varies and as a result, so does their energy security status. Using the responses to the seven questions, labeled here as $a_1$through a_7, the application of PCA produces the following:

(9) $$PC_m = w_{m1}a_1 + w_{m2}a_2 + \cdots + w_{m7}a_7 \qquad \forall m = 1, \ldots, 7,$$

where w_m is used to represent the individual weight assigned to each specific question a for the m^{th} corresponding principal component. The first principal component produced $m = 1$ is considered to be the linear combination that explains as much variation as possible among the variables of interest (Abeyaskera 2003).

The first step in the application of PCA to create a single index measure from a set of correlated variables is to estimate a correlation matrix for the individual variables of interest. Most PCA procedures calculate this first step using Pearson correlations, which assume the variables of interest are normally distributed. Responses to the questions included in Table 3.1 and used to create our index measure for energy insecurity, however, are binary. As a result, to create an energy insecurity index using PCA we first have to generate a tetrachoric correlation matrix for each of the seven binary response variables included in our analysis. In its simplest form, a tetrachoric correlation matrix can be thought of as a matrix of the Pearson correlation coefficients for a set of bivariate normally distributed variables.

In addition to being coded as binary response variables, an essential characteristic of the questions included in Table 3.1 is that the severity of the questions varies. The severity of the conditions identified by the different questions is somewhat intuitive from reviewing them. For example, the question, "*In the last year, has anyone in your household needed medical attention because your home was too hot or too cold?*" is a more severe hardship and represents less favorable energy service circumstances inside the home compared to the hardship represented by the question, "*In the last year, how many months did your household receive a disconnection notice, shut off notice, or non-delivery notice for an energy bill?*" However, the precise severity level of each specific question is unknown *a priori*.

For our analysis, we interpret the weights produced in the first principal component as factor scores. Based on how responses to the survey questions included in Table 3.1 are coded, if a household responds affirmatively to an individual question then it is assumed that sometime over the past twelve months the household has faced the particular hardship identified by the question. Therefore, to obtain values for s_i, the factor scores $\widehat{w}_{1i}$ associated with the specific questions the household responds affirmatively to are added together.

The result of this summation is outlined below

(10) $$s_{i,pca} = \widehat{w}_{11}a_1 + \widehat{w}_{12}a_2 + \cdots + \widehat{w}_{17}a_7,$$

Where $s_{i,pca}$ corresponds with the individual value for s_i produced specifically from the application of PCA.

Based on how the values of $s_{i,pca}$ are constructed we can interpret $s_{i,pca}$ as each individual household's "energy insecurity score," where higher energy insecurity scores indicate more energy service hardships within the household. Furthermore, based on how the values of $s_{i,pca}$ are constructed, values of $s_{i,pca}$ produced that are equal to zero ($s_{i,pca} = 0$) indicate the household did not respond affirmatively to any of the questions listed in Table 3.1 (i.e., has not faced any hardships over the past twelve months). Conversely we assume households who receive a positive energy insecurity score (i.e., a value for $s_{i,pca} > 0$) have faced at least one hardship that has prevented access to adequate energy services within the household.

Thus, households facing more hardships, both in terms of quantity and level of severity, are associated with larger energy insecurity scores being received (i.e., greater values of $s_{i,pca}$ being produced). To construct an energy insecurity index using the results from PCA, we assume a threshold value of $\tau = 0$. We then use two separate techniques to partition households into mutually exclusive energy insecurity categories and assign energy insecurity index values. The different energy insecurity categories are designed to separate energy insecure households from households who identify as being energy secure. The energy insecurity category to which each household is assigned depends on the value of the individual energy insecurity score ($s_{i,pca}$) the households receives and how that value relates to the value of the chosen threshold, $\tau = 0$.

Under the first method of partitioning, households with an energy insecurity score ($s_{i,pca}$) equal to zero, were assigned an energy insecurity index value equal to zero or $EISINDEX_i = 0$. Conversely, households who received an energy insecurity value of $s_{i,pca} > 0$ were assigned an energy insecurity index value equal to one or $EISINDEX_i = 1$. Mathematically, this method of partitioning households into different energy insecurity categories to construct an energy insecurity index is represented as:

$$(11) \qquad EISINDEX_i = \begin{cases} = 1 \ \ if \ s_{i,pca} > 0 \\ = 0 \ \ if \ s_{i,pca} = 0 \end{cases}.$$

Using this method of partitioning we have only two index values to measure household energy security: $EISINDEX_i = 0$ and $EISINDEX_i = 1$.

Recall that each individual household i's energy insecurity score $\left(s_{i,pca}\right)$ is determined by the number of questions in Table 3.1 to which the household responds affirmatively. An affirmative response to any of the questions is an indicator that a household was not able to access an adequate level of energy services inside the household during sometime over the past twelve months. We consider households who are unable to produce adequate household energy services (i.e., meets each household members' most basic daily energy service needs) as energy insecure. Therefore, following the first method of partitioning households who receive an $EISINDEX_i = 0$ are considered "Energy Secure," while households who receive $EISINDEX_i = 1$ are considered "Energy Insecure."

Because we assume all values of $s_{i,pca} > 0$ indicate the household has faced energy service hardships over the past twelve months that have prevented the household from maintaining access to adequate energy services, we also create a second energy insecurity index using the PCA results which allows us to measure the extent of energy insecurity experienced by households. To create the second energy insecurity index, we assume if $s_{i,pca} > 0$, then the difference between $s_{i,pca}$ and the chosen value of threshold $\tau = 0$ is equivalent to the extent of energy insecurity experienced by the household.

Under this framework, households are either energy secure or energy insecure. The extent of energy insecurity for households who identify as "Energy Insecure" however, is determined by the difference between the threshold value of τ set equal to zero and the value of $s_{i,pca}$ received by the household. As before we assign households who received an $s_{i,pca}$ value equal to 0 and an energy security index value $EISINDEX_i = 0$. Households who received value of $s_{i,pca} > 0$ are assigned an energy security index value $EISINDEX_i = (s_{i,pca} - 0)$. Mathematically this method of partitioning households into different energy insecurity categories to construct an energy insecurity index is represented as follows:

$$(12) \quad EISINDEX_i = \begin{cases} = (s_{i,pca} - 0) \text{ if } s_{i,pca} > 0 \\ = 0 \text{ if } s_{i,pca} \leq 0 \end{cases} .$$

In this case, households who receive an energy insecurity index value of zero or $EISINDEX_i = 0$ are considered energy secure. Households who receive an energy insecurity index value equal to $(s_{i,pca} - 0)$ however, are considered energy insecure. The degree of severity of energy insecurity experienced by the household is determined by the absolute difference between the individual energy insecurity score received $s_{i,pca}$ and τ, the chosen value of the threshold, which is assumed to be zero. The energy insecurity index, for energy insecure households ranges in value from zero to the highest energy insecurity score received by an individual household such that $s_{i,pca} \in [0, S_{pca}]$ Here S_{pca} represents the highest energy insecurity score a household is able to receive from the application of PCA. Using this method of partitioning, the higher the energy insecurity index value assigned to the household, the less energy secure the household.

DICHOTOMOUS RASCH MODEL

The fifth method we used to determine a household's energy insecurity status is the Dichotomous Rasch Model. The Dichotomous Rasch Model is a type of item response theory (IRT) model used currently by the USDA to measure food insecurity in the United States (Opsomer, Jensen, and Pan 2003). [53] IRT models are mathematical models that attempt to explain relationships between latent,

[53] For a detailed explanation of how the Dichotomous Rasch Model is applied by the USDA to measure food insecurity, see the Appendix.

unobservable traits and observed outcomes. [54] The theory underlying all IRT models is that the probability an individual will respond affirmatively to an item/question is determined by the difference between the item/question's level of difficulty (level of severity) and the individual person's unobserved ability (position or trait) (Kilanowski and Lin 2012). Questions are assumed to become increasingly more difficult (i.e., more severe) as fewer individuals respond affirmatively to them.

All IRT models make the following key assumptions: (a) individuals differ from each other based on their unobserved latent traits (i.e., unobserved ability); (b) the probability an of an individual responding affirmatively to an item/question of interest is a function of their unobserved latent traits; (c) an individual person's responses to different items/questions are assumed to be independent of one another (i.e., responses are assumed to be locally independent); and (d) responses from different individuals are assumed to independent of each other (Lalor, Wu, and Yu 2016).

If the local independence assumption is not violated and the items/questions under consideration are a "good fit," then the estimated severity levels of the individual items can be used to create a scale. The scale is assumed to measure the full extent of the unobserved latent trait. [55] Items/questions are placed along that scale based on their estimated level of difficulty (severity) (Opsomer et al. 2003). Respondents can be placed along the same scale based on their

[54] They have been used extensively in the educational testing industry to measure student's academic abilities (Yang and Kao 2014; Ames and Penfield 2015).

[55] Infit and outfit statistics are used to determine how "good" the questions identified are at measuring an underlying construct (Murray and Mills 2012). The outfit statistic is an unweighted fit statistic based on standardized residuals (Hamilton et al. 1997). The infit statistic is a weighted fit statistic based on standardized residuals (Hamilton et al. 1997). As the infit and outfit statistics of items of interest deviate from their expected value, they are considered as candidates for removal (Hamilton et al. 1997).

overall ability (i.e., latent unobserved trait) (Opsomer et al. 2003). The latent trait under consideration for our study is a household's true energy insecurity status.

Similar to how the experience of being food insecure is considered to be too complex to be captured by any single indicator (Bickel et al. 2000), the experience of being energy insecure is also too complex to be captured by any single indicator. Furthermore, similar to how there is no commonly used language that describes the entire continuum of what it means food insecure (Nord 2014), there is also no commonly used language the describes the entire continuum of what it means to be energy insecure (Murray and Mills 2012). Energy insecurity is a latent trait and not directly observable. People do not say, on a scale from 1 to 10, my energy insecurity is at level 3 (Nord 2014). Instead, people reveal information about their level of energy insecurity based on how they respond to different questions regarding energy service-related hardships they have faced or are currently facing.

Information about these experiences can be elicited from survey questions (Nord 2014). To determine a household's energy insecurity status, we apply the Dichotomous Rasch model to the seven questions from the 2015 RECS listed in Table 3.1. The Dichotomous Rasch model allows us to consider how responses to this set of survey questions can be combined together to create an energy insecurity scale (Nord 2014). The energy insecurity scale represents the full continuum of what it means for a household to be energy insecure. That is, it captures information on the set of experiences faced by households, ranks those experiences in terms of severity. Households are placed along the energy insecurity scale based on information they reveal from their responses to the questions.

As stated earlier in Section 3.3, responses to the questions in Table 3.1 are coded as binary response variables and each household receives the opportunity to respond to all seven questions.

We assume each household in our sample responds to each question based on its underlying (e.g., latent) true level of energy insecurity. The more energy insecure the household, the higher the probability the household will give a positive (e.g., yes) response to the questions. In the Dichotomous Rasch model, each question that is asked represents a different level of energy service-related hardships being experienced by the household.

Therefore, each question is assumed to indicate a different severity of energy insecurity being experienced by the household, such that questions representing higher severity levels are more likely to be answered negatively (e.g., "no" answer). Conversely, questions representing lower levels of severity are more likely to be responded to affirmatively (e.g., "yes" answer), which indicates more severe energy service-related hardships are occurring less frequently in the home. The Dichotomous Rasch model can be expressed as a logistic model of the form:

$$(13) \qquad Prob\left(I_{ij} = 1 \middle| \alpha_i, \theta_j\right) = \frac{\exp(\alpha_i - \theta_j)}{1 + \exp(\alpha_i - \theta_j)},$$

such that the probability a household responds affirmatively to household energy insecurity question $I_{ij} = 1$ is conditional on the individual household respondent's ability (underlying latent energy insecurity status), α_i, and the individual question's level of hardship severity (difficulty), θ_j. The indicator variables I_{ij} are assumed to be independent of each other, conditional on the parameters α_i and θ_j.

An easy way to interpret the model in equation (13) is to note when $\alpha_i = \theta_j$, the individual household i has a 50% chance of responding to the question j affirmatively. When $\alpha_i > \theta_j$, the probability that individual household i responds affirmatively to question j is greater than 50% and, conversely when, $\alpha_i < \theta_j$ the probability a household responds affirmatively to question j is less than 50% (Opsomer et al. 2003). The model for all questions and households aggregates to the following:

$$(14) \qquad Prob\left(I_{ij} = 1 \middle| \alpha_i, \theta_j\right) = \prod_i^n \prod_j^m \frac{\exp(\alpha_i - \theta_j)}{1+\exp(\alpha_i - \theta_j)},$$

such that n represents the number of households in the sample and m represents the number of questions, which in our case is seven. Although, the model in equation (14) is an exponential model, it cannot be fitted directly via maximum likelihood estimation unless appropriate constraints are added to prevent over parameterization (Opsomer et al. 2003). To get unique parameter estimates, $\sum_{j=1}^{m} \theta_j$ must be set equal to zero (Opsomer et al. 2003).

Conditional maximum likelihood estimation of equation (14) produces item severity parameter estimates ($\hat{\theta}_j$) for each of the $J = 7$ items/questions we consider. The severity parameter estimates are consistent with observed household responses to the items/questions (Nord 2014). In the food security literature, severity parameter estimates are referred to as "item calibrations" (Hamilton et al. 1997). The severity parameter estimates (i.e., item calibrations) for the questions are combined together to create an energy insecurity scale (i.e., a continuous interval-level measure of household energy insecurity).

The energy insecurity scale is a continuous, linear scale designed to measure the degree of energy insecurity experienced by an individual household in terms of a single numerical scale (Bickel et al. 2000). The range of values on the scale are assumed to expresses the full range of severity of energy insecurity as observed in all households across the United States (Bickel et al. 2000). The unit of measure for the scale is a matter of convention (Bickel et al. 2000). [56] Following

[56] Application of the Rasch model initially assigns scale values in a range that yields a mean of zero, which can produce both positive and negative estimated severity parameter estimates. Because the presence of both positive and negative values on a scale can be difficult to interpret, it is conventional to transform values into a range such as 0 to 10 or 0 to 100. The estimated severity level of each parameter estimate is transformed to a mathematically equivalent value of the chosen scale (Bickel et al. 2000).

the literature on food security (Bickel et al. 2000), we let the energy insecurity scale range in value from zero to the most severe level of energy insecurity, which is numerically equivalent to the severity parameter estimate of the most severe question we consider.

A household's placement along the scale is determined by the number of increasingly severe indicators of energy insecurity that the household has experienced, which is determined by the number of questions the household responds affirmatively to (Bickel et al. 2000), also known as the household's "raw score" (Bickel et al. 2000). The household's "raw score" is used to determine where on the energy insecurity scale the household will fall and what energy insecurity score will be received by the household. Because there are seven questions in Table 3.1, there are at most eight possible energy insecurity scores households can receive.

As a result, there are at most eight different locations along the scale where households can be placed. Households who respond to none of the questions included in Table 3.1 are considered to be fully energy secure (i.e., not energy insecure) and as a result are assigned a scale value of zero and placed at the bottom of the scale. Households who respond to all seven questions in the affirmative are considered to be completely energy insecure. These households are assigned a scale value equal to the parameter estimate of the most severe indicator of household energy insecurity (i.e., the parameter estimate associated with the most difficult questions) and placed at the top of the scale.

Other households are placed along the scale based on the number of questions they respond affirmatively. They are assigned a scale value that is equivalent to the estimated severity parameter estimate that corresponds to the number of questions to which they responded in the affirmative. For example, if the parameter estimate produced for the 4th most severe indicator of household energy insecurity is equal to 3.0, then households who respond affirmatively to four questions are assigned a scale value (i.e., energy insecurity score) equal to 3.0.

Households can be ranked, based on the energy insecurity scale value they are assigned, which represents the severity of energy insecurity experienced within the household. If the scale is subdivided into groups or categories, which is what is commonly done in the food security literature (see Hamilton et al. 1997), then households can also be ranked based on which of the energy insecurity categories they are assigned to (Nord 2003). Each category represents a distinct range of values along the energy insecurity scale. Households are separated into energy insecurity categories, based on which range of values along the scale include the value of the energy insecurity score received by the household.

For example, following the literature on food security (Bickel et al. 2000) we could classify households as being "Very Low Energy Secure," "Low Energy Secure," "Marginally Energy Secure," or "High Energy Secure" by first establishing those ranges along the scale and then determining to which category a household belongs based on their current position along the energy insecurity scale.[57] However, as Gundersen (2008) and Balistreri (2016) point out, while classifying households into different energy insecurity categories is convenient, as an aggregation technique it does not provide a complete and accurate representation of the extent, depth, and severity of energy insecurity being experienced by households living across the United States.

Because households are placed into different categories based on their position along the scale and the range of values on scale is established from the severity parameter estimates of the

[57] The USDA classifies households into one of four groups (very low food security, low food security, marginal food security, and high food security) based on the number of questions from the 18-itm food security scale the household responds affirmatively to. For example, based on the 18-item scale, if a household responds affirmatively to four, five, six, seven, or eight questions then the household is classified as being "Marginally Food Secure" (Nord 2014).

different items, households who respond affirmatively to a similar number of items are all classified in the same group. Even though the energy service related hardships between households within the same group could vary substantially. For example, assume positive responses to three, four, or five questions results in a household being categorized as being "Low Energy Secure." It is obvious that these households are not full energy secure or fully energy insecure based on the number of questions they response positively to.

However, categorizing all of these household as being "Low Energy Secure," does not account for the fact that the energy service-related hardships faced by these households is likely to vary widely. Therefore, when examining the extent, depth, and severity of energy insecurity being experienced by households living across the United States, it is important to consider the individual energy insecurity scores received by households within the same group. To establish a more accurate aggregate measure of energy insecurity, while considering the information contained in each of the questions used to determine a household's energy insecurity score, Dutta and Gundersen (2007), Gundersen (2008), and Balistreri (2016) suggest creating a normalized insecurity index.

The normalized energy insecurity index provides an estimate for the full extent, depth, and severity of energy insecurity being experienced by households living in the U.S. The normalized insecurity index is adapted from the index measure for general poverty (Dutta and Gundersen 2007; Gundersen 2008; Balistreri 2016). We follow this advice and create a normalized energy insecurity index (NEII) using the Rasch model results, assuming three different ways of partitioning households into different energy insecurity categories. The NEII is a normalized version of the energy insecurity index outlined in Section 3.3.

Recall from Section 3.3. that the extent of energy insecurity experienced by each household i depends on the energy service-related hardships faced by the household. The value s_i is used to denote the "energy service-related hardships" faced by an individual household i, such that higher values of s_i corresponds with more energy service-related hardships being experienced by members of the household (i.e., the latent value of energy insecurity experienced by each household i). Values for s_i are assumed to exist within the interval $[0, S]$, such that 0 represents no energy service-related hardships being experienced by the household and S represents the most energy service-related hardships possible being experienced by the household.

Households are considered to be energy insecure if the energy service-related hardship they face, exceed the threshold of energy service-related hardships considered to be acceptable (i.e., $s_i > \tau$ where τ is used to represent the threshold). Conversely households are considered energy secure if $s_i \leq \tau$ (Gundersen 2008). If the objective is to classify households into different energy insecurity groups, while still taking advantage of the information contained in each question, the group to which the household is classified should depend on the distance between the energy insecurity score the household receives and the threshold value chosen. Households who are of similar distance away from the chosen threshold can be placed in the same category. The higher the value of the $EISINDEX_i$, the farther the household is from the threshold of being energy secure (Gundersen 2008, Balistreri 2016).

No matter the number of groups considered, the normalized $EISINDEX_i$, represents an aggregation of energy security levels for all different household living across the United States (Gundersen 2008; Balistreri 2016). The normalized energy insecurity index is denoted here as d_i and calculated as follows:

(15) $d_i = \frac{s_i - \tau}{z - \tau}$ if $s_i > \tau$; $d_i = 0$ if $s_i \leq \tau$.

Here d is denotes the degree of energy insecurity suffered by the group of all households N (Gundersen 2008; Balistreri 2016); s_i is the energy insecurity score received by the household from applying the Dichotomous Rasch model; z is the maximum possible energy insecurity score able to be received by a household from the application of the Rasch model (i.e., the severity parameter estimate associated with the most severe question in Table 3.1); and τ is the chosen value of the threshold.

Following Gundersen (2008) we assume d is real valued function of $d_1, d_2, \dots, d_n$ that is the "rule" for aggregating household's energy insecurity levels. The aggregation rule more specifically is a function $D: [0,1]^n \to R^n$ such that D aggregates the energy insecurity levels $d_1, d_2, \dots, d_n$ of the different households groups on an index, d (Gundersen 2008). The normalization of the index requires that d be zero when the normalized energy insecurity index is zero for all households. In addition, it requires d be equal to one when the normalized energy insecurity index is one for all households.

Following Gundersen (2008) and Balistreri (2016) we use three different aggregation rules for the function D and create three separate energy insecurity indices using the Rasch model results and the following formula:

(16) $d^{\alpha} = \frac{\sum_{i=1}^{n} (d_i)^{\alpha}}{n}$.

Here n denotes the total number of in the population of interest (i.e., households living across the United States) and d_i represents the normalized energy insecurity index value received by the households following equation (15). When $\alpha = 0$, d defines the energy insecurity rate of the population, or the proportion of households living in the United States that are energy insecure

(Gundersen 2008). When $\alpha = 1$, equation (16) represents the energy insecurity gap or depth of energy insecurity experienced.

One can think of the energy insecurity gap as the average proportionate gap in energy insecurity, or the amount on average that households fall below the energy insecurity threshold chosen (Balistreri 2016). When $\alpha = 2$, equation (16) is assumed to measure the severity of energy insecurity being experienced by households living in the United States (Gundersen 2008; Balistreri 2016). These different measures are useful for understanding how when one household becomes more energy insecure, the prevalence of energy insecurity is unchanged in the United States, but the gap between energy security and energy insecurity becomes increasingly worse.

3.6 RESULTS

Results from the five different approaches considered to determine a household's energy security status are presented below.

ENERGY ASSISTANCE RESULTS

Tables 3.3 through 3.7 present estimates for the number of households who identify as being "energy insecure," according to how households responded to the three questions in the RECS survey related to home energy assistance. Households are considered to be energy insecure if they responded affirmatively to any of the three questions. Table 3.3 presents results for all housing types, considering single family attached homes, single-family detached homes, apartments, and mobile homes.

Table 3.3 Energy Assistance (All Housing Types)

Energy Assistance Questions from the 2015 RECS	Number of Households	Percentage (%)
Household Participated in a Home Energy Assistance Program that helps Pay Energy Bills and/or Replace/Fix Broken HVAC	345	6.07

Household Applied for and Received Energy Assistance to Help Pay Home Energy Bills after Receiving a Disconnection Notice	116	2.04
Household Applied for and Received Energy Assistance to Help Restore Heating and/or Cooling in the Home	41	0.25
Total Number of Households	N = 5,686	

Using participation in an energy assistance program as a proxy for a household's energy security status, 345 households identify as being energy insecure. Approximately 2% of the households in our sample responded affirmatively to applying for and receiving home energy assistance to help pay their home energy bills after receiving a disconnection notice. Information on applicants that applied for but did not receive home energy assistance to help pay their home energy bills is not available. Only 41 of the households in our sample responded affirmatively to receiving energy assistance to help restore heating and/or cooling inside the home, either due to equipment failure or an inability to afford energy resource inputs (i.e., electricity, propane, heating oil).

Tables 3.4 through 3.7 below provide estimates for the number of households living in different types of housing units (i.e., single-family detached and attached homes, apartment buildings, and mobile homes) who identify as being "energy insecure," according to whether or not the household received home energy assistance.

Table 3.4 Energy Assistance (Single Family Detached Homes)

Energy Assistance Questions from the 2015 RECS	Number of Households	Percentage
Household Participated in a Home Energy Assistance Program that helps Pay Energy Bills and/or Replace/Fix Broken HVAC	179	4.77
Household Applied for and Received Energy Assistance to Help Pay Home Energy Bills after Receiving a Disconnection Notice	58	1.55

Household Applied for and Received Energy Assistance to Help Restore Heating and/or Cooling in the Home	18	0.48
Total Number of Households living in Single Family Detached Homes	N = 3,752	

Table 3.5 Energy Assistance (Single Family Attached Homes)

Energy Assistance Questions from the 2015 RECS	Number of Households	Percentage
Household Participated in a Home Energy Assistance Program that helps Pay Energy Bills and/or Replace/Fix Broken HVAC	33	6.89
Household Applied for and Received Energy Assistance to Help Pay Home Energy Bills after Receiving a Disconnection Notice	12	2.51
Household Applied for and Received Energy Assistance to Help Restore Heating and/or Cooling in the Home	4	0.84
Total Number Households living in Single Family Attached Homes	N = 479	

Table 3.6 Energy Assistance (Apartments)

Energy Assistance Questions from the 2015 RECS	Number of Households	Percentage
Household Participated in a Home Energy Assistance Program that helps Pay Energy Bills and/or Replace/Fix Broken HVAC	100	8.55
Household Applied for and Received Energy Assistance to Help Pay Home Energy Bills after Receiving a Disconnection Notice	29	2.48
Household Applied for and Received Energy Assistance to Help Restore Heating and/or Cooling in the Home	10	0.86
Total Number of Households living in Apartments	N = 1,169	

Table 3.7 Energy Assistance (Mobile Homes)

Energy Assistance Questions from the 2015 RECS	Number of Households	Percentage

Household Participated in a Home Energy Assistance Program that helps Pay Energy Bills and/or Replace/Fix Broken HVAC	33	11.54
Household Applied for and Received Energy Assistance to Help Pay Home Energy Bills after Receiving a Disconnection Notice	17	5.94
Household Applied for and Received Energy Assistance to Help Restore Heating and/or Cooling in the Home	9	3.14
Total Number of Households living in Mobile Homes	N = 286	

A single family detached home is defined as a standalone house or a free-standing residential building. A single family attached home is a home that shares at least one wall with another home. Examples of single family attached homes include townhouses, condos, row houses, and duplexes. An apartment is considered to be a residence in a building with at least two or more units. Mobile homes include manufactured and mobile residences that are occupied and used as permanent living accommodations throughout the year.

Seventy percent of the households surveyed in the 2015 RECS indicated they live in single-family detached houses (i.e., stand-alone homes). The second largest housing category was apartment living. Examining household types individually and using the receipt of energy assistance as a proxy for energy insecurity, we find the percentage of households who identify as energy insecure is highest among households who live in mobile homes. Mobile homes are constructed in factories, delivered in pieces and assembled on site. By design, they have poorly insulated floors, walls, and ceilings (Department of Energy 2010).

Poor insulation can result in more energy consumption by households to maintain comfortable indoor air temperatures. As energy consumption increases, households could face higher home utility bills, which could cause some households, especially those with lower incomes, to apply for home energy assistance. The percentage of households who apply for and receive home energy assistance and live in single-family detached homes is lower than the percentage of households who apply for and receive home energy assistance from any other housing type.

THE EXPENDITURE APPROACH RESULTS

Table 3.8 below presents estimates for the number and percentage of households who identify as being "Energy Insecure" according to the expenditure approach considering all household types.

Table 3.8 Expenditure Approach Results (All Households)

	Number of Households	Percentage
Spend More than 6% of Disposable Income	999	17.57
Spend More than 10% of Disposable Income	487	5.86
Total Number of Households	N = 5,686	

Following the suggestions of Boardman (1991) and Chandler (2016) we consider households who spend more than 10% or 6% of their income on energy/fuel to be "Energy Insecure." Considering all household types (i.e., single-family detached and attached homes, apartments and mobile homes) approximately 999 households surveyed in the 2015 RECS spend more than 6% of their annual income on energy/fuel, while approximately 487 households spend more than 10%. Using 6% of income as the threshold beyond which expenditures of energy/fuel cease to be affordable, 17.57% of the households surveyed in the 2015 RECS identify as being energy insecure.

Tables 3.9 through 3.12 below provides estimates for the number and percentage of households living in different types of housing units (i.e., single-family detached and attached homes, apartment buildings, and mobile homes) who identify as "Energy Insecure" according to the expenditure approach. Based on how the individual values of s_i were calculated, households who spend more than 10% of their income on energy/fuel also spend more the 6% of the income of energy/fuel.

Table 3.9 Expenditure Approach Results (Single Family Detached Homes)

	Number of Households	Percentage
Spend More than 6% of Disposable Income	514	13.70
Spend More than 10% of Disposable Income	274	7.30
Total Number of Households Living in Single Family Detached Homes	N = 3,752	

Table 3.10 Expenditure Approach Results (Single Family Attached Homes)

	Number of Households	Percentage
Spend More than 6% of Disposable Income	83	17.33
Spend More than 10% of Disposable Income	35	7.31
Total Number of Households Living in Single Family Attached Homes	N = 479	

Table 3.11 Expenditure Approach Results (Apartments)

	Number of Households	Percentage
Spend More than 6% of Disposable Income	270	23.10
Spend More than 10% of Disposable Income	119	10.18
Total Number of Households Living in Apartments	N = 1,169	

Table 3.12 Expenditure Approach Results (Mobile Homes)

	Number of Households	Percentage
Spend More than 6% of Disposable Income	132	46.15
Spend More than 10% of Disposable Income	86	30.07
Total Number of Households Living in Mobile Homes	N = 286	

Approximately 23% of the households surveyed in the 2015 RECS who live in an apartment identified as being energy insecure when 6% of one's annual income was assumed to be the threshold beyond which energy/fuel expenditures cease to be affordable. While fewer households overall identified as living in mobile homes as permanent residents, the percentage of households living in mobile homes who identified as energy insecure according to the expenditure approach is the highest among all the different housing types considered.

Of the 286 households who live in mobile homes, 132 of them spend more than 6% of their disposable income of energy/fuel each year. Moreover, close to one-third of these households spend more than 10% of their annual income on energy/fuel. For each of the different household types, the number of households who identify as being energy insecure because they spend more than 6% of their income on energy/fuel is nearly double the number of households who identify as being energy insecure because they spend more than 10% of their annual income on energy/fuel. These results indicate that determining the energy security status of households following the expenditure approach is sensitive to value of the chosen threshold.

CLUSTER ANALYSIS

The results from the application of cluster analysis are presented below in Tables 3.13 and 3.14.

Table 3.13 Cluster Analysis Results: Energy Insecurity Groups

Energy Insecurity Groups	Number of Households
1	144
2	4,832
3	634
4	76
Total	**N = 5,686**

Table 3.14 Cluster Analysis Results Summary Statistics

Energy Insecurity Group		Reduce	Unsafe	Notice	No Fuel	HVAC	Medical	Days
	Min	0	0	0	0	0	0	0
1	Mean	0.701	0.118	0.285	0.111	1	0.056	0.506
	Max	1	1	1	1	1	1	1
	Min	0	0	0	0	0	0	0
2	Mean	0	0.023	0.024	0.010	0.031	0.005	0.001
	Max	1	1	1	1	1	1	1
	Min	1	0	0	0	0	0	0
3	Mean	1	0.226	0.222	0.098	0	0.036	0.005
	Max	1	1	1	1	0	1	1
	Min	0	0	0	0	0	0	0
4	Mean	0.882	0.974	0.513	0.237	0.934	0.368	0.158
	Max	1	1	1	1	1	1	1
	Min	0	0	0	0	0	0	0
Total	Mean	0.141	0.061	0.059	0.025	0.064	0.014	0.016
	Max	1	1	1	1	1	1	1

The objective of applying cluster analysis was to understand how the individual households could be naturally grouped together based on their responses to the questions included in Table 3.1. The clusters shown in Table 3.13 are all exclusive, as they assign each household to only one cluster. Table 3.15 below presents the means for each of the variables used from Table 3.1 for each of the four different energy insecurity groups identified by the application of cluster analysis.

Table 3.15 Cluster Analysis Results Average Individual Energy Insecurity Groups

Variable	Group 1	Group 2	Group 3	Group 4
Reduce	0.701	0	1	0.882
Unsafe	0.118	0.023	0.226	0.973

Notice	0.285	0.024	0.222	0.513
No Fuel	0.111	0.010	0.098	0.237
HVAC	1	0.031	0	0.934
Medical	0.056	0.005	0.036	0.368
Days	0.507	0.001	0.005	0.158

Overall, we find all of the households who belong to Group 2 did not respond affirmatively to reducing or foregoing expenses for other basic household necessities such as medicine or food in order to pay their home energy bill at any time over the past twelve months (Question 1 in Table 3.1). Conversely, all of the households in Group 3 did respond affirmatively to reducing or foregoing expenses for other basic household necessities to pay their home energy bill. Similarly, all of the households who belong to Group 1 responded affirmatively to being unable to use their main source of heat or air conditioning because their HVAC equipment was broken and they could not afford to pay to repair or replace the equipment (Question 5 in Table 3.1). Conversely all of the households in Group 3 did not respond in the affirmative to Question 5.

Overall, the average number of affirmative responses for each question is highest for Energy Insecurity Groups 1 and 4. However, the number of households in these two groups is the smallest with 144 households being assigned to Group 1 and only 63 households being assigned to Group 4. The average number of affirmative responses to all questions is the lowest for Group 2. To determine the energy insecurity status of each of the four groups identified by the application of cluster analysis, we examine the individual characteristics of the different households identified in each of the four groups.

We consider the following key household level characteristics: median household income measured in US Dollars (*INCOME*); an indicator variable (*MOBILE)* equal to one if the residence occupied by the household is a mobile home and zero otherwise; an indicator variable (*DETACHED*) equal to one if the residence occupied by the household is a single family detached home and zero otherwise; an indicator variable (*ATTACHED*) equal to one if the residence occupied by the household is a single family attached home and zero otherwise; an indicator variable (*APARTMENT*) equal to one if the residence occupied by the household is a unit in an apartment building and zero otherwise; an indicator variable (*OWNERSHIP*) equal to one if the primary residence is owned by the household and zero otherwise (i.e., rented); an indicator variable (*EMPLOYMENT*) equal to one if the head of the household is employed either part-time or full time and zero otherwise; a continuous variable equal to the number of people living inside the home (*MEMBERS*); a continuous variable equal to the number of children aged 16 and younger living inside the home (*CHILDREN*); an indicator variable (*EDUCATION*) equal to one if the head of the household has a college degree or beyond and zero otherwise; a continuous variable equal to the number of bedrooms inside the home (*BEDROOMS*); and a continuous variable equal to the number of complete bathrooms inside the home (*BATHROOMS*). Summary statistics for each of the different energy insecurity groups are listed below in Tables 3.16 through 3.19.

Table 3.16 Summary Statistics for Group 1

Variable	Mean	Std. Dev.	Min	Max
INCOME	41,458.33	34,152.45	10,000	140,000
MOBILE	0.16	0.37	0	1
DETACHED	0.62	0.49	0	1
ATTACHED	0.07	0.26	0	1
APARTMENT	0.15	0.36	0	1
OWNERSHIP	0.61	0.49	0	1
EMPLOYED	0.44	0.50	0	
MEMBERSRS	3.06	1.68	1	10

CHILDREN	0.92	1.18	0	5
EDUCATION	2.76	1.08	1	5
BEDROOMS	2.85	1.01	0	6
BATHROOMS	1.68	0.62	1	4
Observations		N=144		

Table 3.17 Summary Statistics for Group 2

Variable	Mean	Std. Dev.	Min	Max
INCOME	66,916.39	42,997.92	10,000	140,000
MOBILE	0.04	0.19	0	1
DETACHED	0.68	0.47	0	1
ATTACHED	0.08	0.27	0	1
APARTMENT	0.20	0.40	0	1
OWNERSHIP	0.72	0.45	0	1
EMPLOYED	0.49	0.50	0	1
MEMBERS	2.52	1.40	1	12
CHILDREN	0.56	1.01	0	10
EDUCATION	3.22	1.14	1	5
BEDROOMS	2.87	1.11	0	10
BATHROOMS	1.79	0.76	0	6
Observations		N=4,832		

Table 3.18 Summary Statistics for Group 3

Variable	Mean	Std. Dev.	Min	Max
INCOME	36,293.38	28,815.19	10,000	140,000
MOBILE	0.09	0.28	0	1
DETACHED	0.50	0.50	0	1
ATTACHED	0.11	0.31	0	1
APARTMENT	0.30	0.46	0	1

OWNERSHIP	0.48	0.50	0	1
EMPLOYED	0.41	0.49	0	1
MEMBERS	2.81	1.49	1	8
CHILDREN	0.80	1.13	0	6
EDUCATION	2.60	1.00	1	5
BEDROOMS	2.57	1.06	0	6
BATHROOMS	1.49	0.61	0	5
Observations		N=634		

Table 3.19 Summary Statistics for Group 4

Variable	Mean	Std. Dev.	Min	Max
INCOME	30,526.32	23,545.40	10000	130000
MOBILE	0.24	0.43	0	1
DETACHED	0.54	0.50	0	1
ATTACHED	0.09	0.29	0	1
APARTMENT	0.13	0.34	0	1
OWNERSHIP	0.61	0.49	0	1
EMPLOYED	0.37	0.49	0	1
MEMBERS	3.25	1.79	1	10
CHILDREN	1.09	1.39	0	6
EDUCATION	2.33	1.06	1	5
BEDROOMS	2.83	0.94	1	6
BATHROOMS	1.53	0.60	1	3
Observations		N=76		

Overall, we find the median average income is the lowest for households in Energy Insecurity Groups 3 and 4. On average, more households in Group 4 live in mobile homes, but more homes on average in Group 4 identify as being homeowners. The average number of bedrooms and bathrooms is consistent across all four groups, indicating that the size of the household in terms of the number of rooms, provides little information on the energy insecurity status of the different household groups. Furthermore, the average number of children in each group is approximately equal to one child, while the average number of individuals living inside the home in each group is approximately equal to three individuals.

Using the classifications of food security established by the USDA, based on the pattern of responses by households to the questions included in Table 3.1 and the individual characteristics of the households we assume those in Energy Insecurity Group 2 are “High Energy Secure,” those in Energy Insecurity Group 3 are “Marginally Energy Secure,” those in Energy Insecurity Group 1 are “Low Energy Secure,” while those in Energy Insecurity Group 4 are “Very Low Energy Secure.” Table 3.20 below lists the number of households included in each of the energy insecurity groups considered.

Table 3.20 Energy Insecurity Groups

Energy Insecurity Groups	**Number of Households**	**Percentage of Households**
High Energy Secure	4,832	85%
Marginally Energy Secure	634	11%
Low Energy Secure	144	3%
Very Low Energy Secure	76	1%
Total	**N=5,686**	

While the use of cluster analysis allows us to divide the households into four distinct groups, the validity of the results is difficult to determine given that very different clusters can be formed from the same data depending on how the analysis is executed. For example, if a different similarity measure was used in the analysis, then the number and characteristics of the households identified in each group could change drastically.

PRINCIPAL COMPONENTS ANALYSIS RESULTS

The results from applying PCA to create an index measure of household energy insecurity are presented below in Table 3.21 and Table 3.22.

Table 3.21 Individual Principal Components One through Four

Variable	Component 1	Component 2	Component 3	Component 4
Reduce	0.4105	0.3111	-0.1182	-0.3290
Unsafe	0.3893	0.0477	-0.4590	-0.5519
Notice	0.3542	0.4779	0.3267	0.0368
No Fuel	0.3659	0.3265	0.3699	0.3640
HVAC	0.4048	-0.4759	0.2108	0.0804
Medical	0.3547	-0.0972	-0.6251	0.6483
Days	0.3618	-0.5745	0.3109	-0.1627
Eigenvalue	4.0997	1.0732	0.7564	0.4072
Proportion	0.5857	0.1533	0.1081	0.0582

Table 3.22 Individual Principal Components Five through Seven

Variable	Component 5	Component 6	Component 7
Reduce	-0.1669	0.7574	-0.1046
Unsafe	0.3146	0.4792	-0.0481
Notice	-0.5566	0.4100	0.2455
No Fuel	0.6914	-0.0209	-0.1078
HVAC	-0.2500	0.1073	-0.6962
Medical	-0.1192	-0.0325	0.1961
Days	0.0930	-0.1246	0.6259
Eigenvalue	0.3733	0.2436	0.0466
Proportion	0.0533	0.0348	0.0067

Compared to the other principal components produced, principal component number one (labeled as Component 1 in Table 3.21) explains the most variation. Therefore, to construct the energy insecurity index, we utilize the weights produced from the first principal component. Interpreting the weights produced by the first principal component as factor scores $\widehat{w}_{mi}$, we determine each individual household i's energy insecure score $s_{i,pca}$ as follows:

(17) $s_{i,pca} = 0.4105 \cdot \text{Reduce} + 0.3893 \cdot \text{Unsafe} + 0.3542 \cdot \text{Notice} + 0.3659 \cdot \text{No Fuel} + 0.4048 \cdot \text{HVAC} + 0.3547 \cdot \text{Medical} + 0.3618 \cdot \text{Days}.$

The variables Reduce, Unsafe, Notice, No Fuel, HVAC, Medical, and Days are indicator variables that take on a value of one if the individual household i responded in the affirmative to the question that corresponds with the individual indicator and zero otherwise (see Table 3.1). For example, assume an individual household j only responded affirmatively to questions 1, 3, and 5. Following the application of PCA, their individual energy insecurity score $s_{i,pca}$ would be calculated as follows:

(18) $s_{i,pca} = 0.4105(1) + 0.3893(0) + 0.3542(1) + 0.3659(0) + 0.4048(1) + 0.3547(0) + 0.3618(0).$

Therefore, the energy insecurity score for household j is 1.1695. Based on equation (17) the maximum energy insecurity score a household is able to receive is equal to 2.6448. The PCA energy insecurity scores across all household types range in value from 0 to 2.2865, indicating that none of the households in the survey responded affirmatively to all seven questions.

Following the first method of partitioning discussed previously, we assign households who received an energy insecurity score ($s_{i,pca}$) equal to zero an energy insecurity index value equal to zero, $EISINDEX_i = 0$. Conversely, households who received a positive energy insecurity score

($s_{i,pca} > 0$) are assigned an energy insecurity index value equal to one or $EISINDEX_i = 1$. As stated earlier, using this method of partitioning, only two index values for household energy insecurity are considered: $EISINDEX_i = 0$, which indicates the household is "energy secure," and $EISINDEX_i = 1$ which indicates the household is "energy insecure." Based on this framework, 4,442 (78%) of the households included in our sample are considered "Energy Secure," while 1,244 (22%) households are considered "Energy Insecure."

In addition to the energy insecurity index outlined above, using PCA we created a second energy insecurity index to measure the extent of energy insecurity being experienced by households with a positive energy insecurity score ($s_{i,pca} > 0$). More specifically, we assign households who receive an energy insecurity score $s_{i,pca} = 0$ an energy insecurity index value $EISINDEX_i = 0$. As before, these households are considered to be "Energy Secure." Households who receive positive energy insecurity scores ($s_{i,pca} > 0$) however, are considered to be "energy insecure." Because the factor scores produced from the first principal component are all positive, more affirmative responses to the questions in Table 3.1 lead to higher energy insecurity scores being produced. The energy insecurity index value they receive $EISINDEX_i$ is equal to the absolute difference between the individual energy insecurity score received $s_{i,pca}$ and 0. Therefore, the extent of energy insecurity experienced by households increases as the value of the PCA energy insecurity score increases.

As stated earlier, the energy insecurity scores produced from the application of PCA range in value from 0 to 2.2865. Of the 5,686 households included in our sample, 4,442 received an energy insecurity score of zero ($s_{i,pca} = 0$) indicating they did not respond affirmatively to any of the questions included in Table 3.1. As a result, these households are considered completely

"Energy Secure." The remaining households received energy insecurity score, $s_{i,pca}$ that ranged in value from 0.3542 to 2.22865 indicating progressive levels of energy insecurity. Households are divided up based on the energy insecurity score they received. Results presented in Figure 3.2.

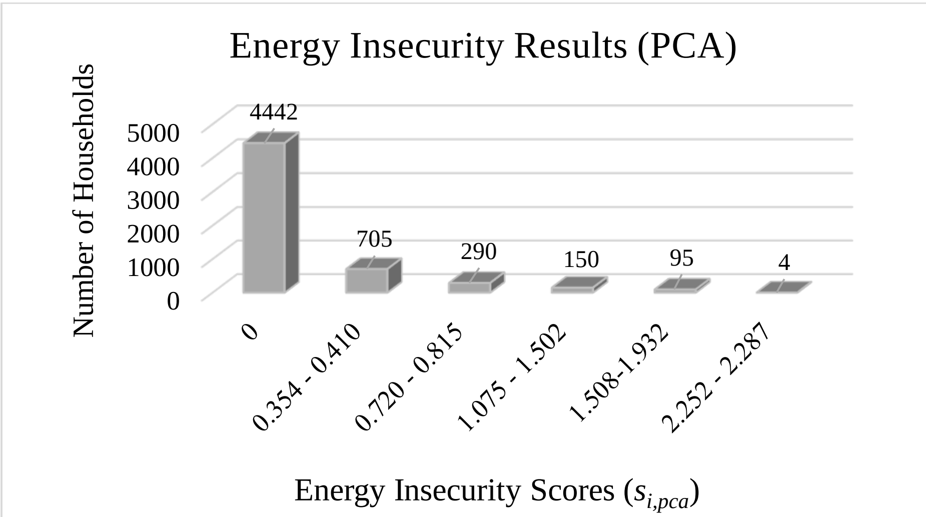


Figure 3.2 Energy Insecurity Index Results (PCA)

Based on the factor scores produced in the first principal component from the application of PCA (see Table 3.20), we can see the majority of households who responded affirmatively to any of the questions included in Table 3.1 responded affirmatively to only one question, which results in an energy insecurity score that ranges in value from 0.354 to 0.410. Of the households included in our sample, 290 households received energy insecurity scores than ranged in value from 0.720 to 0.815, which indicates affirmative responses to either two or three questions. Only four households received an energy insecurity score greater than 2.25. None of the households received an energy insecurity score greater than 2.682, which indicates none of the households included in our sample responded in the affirmative to all seven questions.

DICHOTOMOUS RASCH MODEL RESULTS

The Dichotomous Rasch model analysis was based on all seven questions that are believed to affect the extent and severity of household energy insecurity experienced across the United States.

The results from the application of the Dichotomous Rasch Model are outlined below in Tables 3.23, 3.24, and 3.25. Table 3.23 presents summary statistics on the responses by households to each of the seven questions listed in Table 3.1.

Table 3.23 Summary Statistics Reponses

Question	Variable	Mean	St. Dev.	Minimum	Maximum
1	Reduce	0.141	0.348	0	0
5	HVAC	0.063	0.244	0	0
2	Unsafe	0.061	0.239	0	0
3	Notice	0.059	0.236	0	0
4	No Fuel	0.025	0.157	0	0
7	Days	0.016	0.125	0	0
6	Medical	0.014	0.119	0	0
Observations		N = 5,686			

The questions in Table 3.23 are listed in decreasing order of the proportion of households who responded affirmatively to them. Approximately 14% of the households surveyed responded affirmatively to having reduced or forgone expenditures on other basic household necessities such as medicine or food, in order to pay their home energy bill. Only 0.14% of households responded affirmatively to having sought medical attention because their home was too hot or too cold. Approximately the same percentage of household surveyed responded affirmatively to having kept temperatures at an unsafe level (Question 2) and having received a disconnection notice sometime over the past twelve months (Question 3).

Table 3.24 below presents the "raw scores" for households included in our sample.

Table 3.24 Raw Scores for Questions

Raw Score	Number of Households	Percent
0	4,442	78.12
1	705	12.40
2	290	5.10
3	146	2.57
4	77	1.35
5	22	0.39

6	4	0.07
7	0	0

A household's raw score is equal to the number of questions in Table 3.1 the household responded affirmatively to. Analysis of the data revealed 4,442 households did not respond affirmatively to any of the questions. As a result, these households received a raw score of zero. No households responded affirmatively to all seven questions. As a result, no households in our sample received a raw score of seven. Approximately 12.4% of the households surveyed responded affirmatively to at least one question, which generates a raw score of one for these households. Only four households responded affirmatively to six of the seven question listed in Table 3.1. These households received a raw score of six. The results from estimating equation (14) via conditional maximum likelihood are presented in Table 3.25.

Table 3.25 Estimated Severity Level of the Energy Insecurity Questions

Question	**Item**	**No. of Households Responded "Yes"**	**Item Calibration**	**St. Error**
Q1	Reduce	802	3.08	0.08
Q5	HVAC	363	4.38	0.11
Q2	Unsafe	347	4.44	0.10
Q3	Notice	338	4.48	0.10
Q4	No Fuel	143	5.65	0.13
Q7	Days	91	6.20	0.15
Q6	Medical	81	6.34	0.15
Observations	**N= 5,686**			

The questions (i.e., items) are listed in Table 3.25 according to their estimated level of severity. The severity estimates of each question are listed in column four, which is labeled as "Item Calibration." An individual item's calibration represents the point on the energy insecurity scale where there is at least a 50% chance that any given household will respond "yes" to the specific item (Bickel et al. 2000). Households with higher values on the scale than a particular item's calibration score have more than a 50% chance of responding affirmatively to the individual item, while households with lower values on the energy insecurity scale have less than a 50% chance of responding affirmatively to the item in question.

Based on the Dichotomous Rasch model results, reducing or forgoing expenditures on other household necessities such as food and/or medicine to pay energy bills (Question 1) appears to be the least difficult item for households to affirm. Seeking medical attention because the home was too hot or too cold (Question 7) appears to be the most difficult item to confirm. These results make intuitive sense. Item clumping seems to be present between some questions. Questions 2, and 5 are grouped closely near 4.4 along the latent energy insecurity scale. This result is not surprising considering an inability to use one's main source of heat or air conditioning because equipment was broken (Question 5) likely leads to indoor air temperatures feeling unsafe or unhealthy.

The central function of the Rasch model is to assign each responding household a value on the energy insecurity scale. The energy insecurity scale is constructed from the individual item calibrations produced (see column 4 in Table 3.25). It ranges in value from the lowest level of energy insecurity experienced (the item calibration corresponding to Question 1) to the highest level of energy insecurity experienced (the item calibration corresponding to Question 6). The household's scale value is based on the number of questions the household responded affirmatively to (i.e., a count of the number of affirmative responses to the questions included in the scale). The energy insecurity scale constructed from the item calibrations in Table 3.25 is presented below in Figure 3.3.

Energy Secure	Item Severity	Number of Households
	0	4,442
	3.08	705
	4.38	290
	4.44	146
	4.48	77
	5.65	22
	6.20	4
Energy Insecure	6.34	0

Figure 3.3 Energy Insecurity Scale

There are no excessively large gaps between less severe questions and more severe questions. The gap between most questions is less than one logit. The difference between the easiest "least severe" question (Question 1) and the hardest "most severe" question (Question 6) is approximately 3.26 logits. The majority of the households (4,442 of the 5,686 households) surveyed in the 2015 RECS identify as being "energy secure." These households received an energy insecurity score of zero. 705 households surveyed received an energy insecurity score of 3.08. The remaining households received an energy insecurity score greater than 3.08, which places them at the more severe end of the energy insecurity scale.

Using the household scores produced from the application of the Dichotomous Rasch model it is possible to identify household level characteristics of the most energy insecure households. We consider the "most energy insecure" households to be those who receive an energy insecurity score of 5.65 or greater, given these are the households who respond affirmatively to more than half of the questions included in the 2015 RECS.

Table 3.26 Household Characteristics of Energy Insecure Households

Household Characteristic	**Mean**	**Mean All Households**
Received Energy Assistance	19%	0.06
Median Income ($)	37,692	62,371
Employed	27%	0.48
Post-Secondary Education	31%	0.69
Renter	54%	0.69
Number of Children	0.81	0.61
Head of Household Hispanic	38%	0.13
Head of Household African American	23%	0.10
Single Family Detached Home	62%	0.66
Single Family Attached Home	12%	0.08
Mobile Home	23%	0.05
Apartment	4%	0.21
Observations	**26**	**N = 5,686**

Several significant difference exist between households who are severely energy insecure and all other household types. First and foremost, nearly 20% of the households who are considered to be severely energy insecure applied for and received home energy assistance to help restore heating and/or cooling in their home. Compared to all other household types, severely energy insecure households are five times more likely to reside in mobile homes.

The median average income of severely energy insecure households is about half the average income of all other households. A larger percentage of households who are severely energy insecure are headed by individuals who are Hispanic or African American. There is very little difference in the number of children who live in severely energy insecure households compared to all other household types. Approximately the same percentage of severely energy insecure households and all other households live in single family detached homes.

To create an energy insecurity index, $EISINDEX_i$ for each household i, we choose two different threshold values for τ. The first threshold value is $\tau =$ zero. No affirmative responses to the questions included in Table 3.1 is an indication that the household has not faced an energy service-related hardships over the past month and therefore, is fully energy secure. The energy insecurity index value for each household, assuming the threshold value of $\tau = 0$ is calculated as follows:

(19) $EISINDEX_i = s_i - 0.$

Here, just as before s_i is equal to the energy insecurity score received by the household from the Rash model. Possible values for $s_i \in (0, 3.08, 4.38, 4.44, 4.48, 5.65, 6.20, 6.34)$. In this case, the energy insecurity index $EISINDEX_i$ can take on one of the eight different values included in the set. A household's energy insecurity index value is equivalent to the numerical value associated with its placement along the energy insecurity scale.

The normalized energy insecurity index d_i under the first case where the threshold value of $\tau = 0$, is calculated as follows:

(20) $d_i = \frac{s_i - 0}{6.34 - 0}$ if $s_i > 0$; $d_i = 0$ if $s_i \leq 0$.

Under the first case, all the values of $s_i \geq 0$. Therefore, the normalized energy insecurity index ranges from 0 to 1. Table 3.27 below lists the results.

Table 3.27 Normalized Energy Insecurity Index ($\tau = 0$)

Questions Responded Affirmatively To	s_i	d_i	No. of Households
0	0	0	4,442
1	3.08	0.49	705
2	4.38	0.69	290
3	4.44	0.70	146
4	4.48	0.71	77
5	5.65	0.89	22
6	6.20	0.98	4

7	6.34	1	0
Observations	**N = 5,686**		

Using the formula in equation (16) and letting N=5,686 we can aggregate the index values in Table 3.27 to determine the proportion of households who are energy insecure, the energy insecurity gap (i.e., depth of household energy insecurity being experienced), and the severity of energy insecurity experienced.

The proportion of households who identify as energy insecure is equal to

(21) $$d^0 = \frac{\sum_{i=1}^{5,686}(d_i)^0}{5,686} = \frac{1,244}{5,686} = 0.2188.$$

When the threshold value $\tau = 0$, the energy insecurity gap is equal to

(22) $$d^1 = \frac{\sum_{i=1}^{5,686}(d_i)^1}{5,686} = \frac{725.97}{5,686} = 0.1277$$

and the severity of energy insecurity is equal to

(23) $$d^2 = \frac{\sum_{i=1}^{5,686}(d_i)^2}{5,686} = \frac{605.35}{5,686} = 0.1065.$$

Overall, we find approximately 22% of the households surveyed in the 2015 RECs identify as being energy insecure. The average household gap in energy insecurity is about 12%, which indicates we can expect 12% of all households, on average, to fall below the energy security threshold. Households who fall below the energy security threshold can expect to be 11% more energy insecure than households who are above the energy security threshold.

The second threshold value of τ we choose is based on the threshold value of τ used to create the food insecurity index (FII) (Hamilton et al. 1997; Bickel et al. 2000; Nord 2003; Gundersen 2008; Balistreri 2016). To estimate a food insecurity index value for households living across the United States, the USDA created a food security scale from responses to 18 items/questions included in the Current Population Survey (CPS) (Hamilton et al. 1997; Bickel et al. 2000). The severity of each question was estimated using the Rasch model (Hamilton et al. 1997; Bickel et al. 2000). The range of the severity parameter estimates (i.e., difference between the most severe question and least severe question) determines the range of food insecurity on the scale.

Based on the Rasch model results, the USDA considers households to be "food secure" if they respond affirmatively to three or fewer questions (Nord 2003; Balistreri 2016). The threshold value of τ in this case corresponds to severity parameter estimate for the second most severe question in the scale (Nord 2003; Balistreri 2016). The FII for each household i is calculated as the difference between the severity estimate for the number of questions to which the household responds affirmatively and the severity parameter estimate of the second most severe survey question.

Using a similar approach, for the second energy insecurity index we created using the Rasch model results, we consider households to be "energy insecure" if and only if they respond affirmatively to two or more questions included in section L of the 2015 RECS (see Table 3.1). In this case the threshold value of τ is assumed to be equal to the severity parameter estimate of the first question (Q1 Reduce) which is equal to 3.08. The energy insecurity index value for each household i is now equal to

(24) $EISINDEX_i = s_i - 3.08$

for all values of $s_i > 3.08$. The energy insecurity index value received can now take on one of seven values such that $EISINDEX_i \in (0, 1.30, 1.36, 1.40, 2.57, 3.12, 3.26)$.

The normalized energy insecurity index d_i under the second case when the threshold value of $\tau = 3.08$, is calculated as follows:

(25) $d_i = \frac{s_i - 3.08}{6.34 - 3.08}$ if $s_i > 3.08$; $d_i = 0$ if $s_i \leq 3.08$.

Table 3.28 lists the results.

Table 3.28 Normalized Energy Insecurity Index ($\tau = 3.08$)

Questions Responded Affirmatively To	s_i	d_i	No. of Households
0	0	0	4,442
1	3.08	0	705
2	4.38	0.40	290
3	4.44	0.42	146
4	4.48	0.43	77
5	5.65	0.79	22
6	6.20	0.96	4
7	6.34	1	0
Observations	**N = 5,686**		

Again, the normalized energy insecurity index again ranges from 0 to 1. The proportion of households who identify as energy insecure is equal to

(26) $d^0 = \frac{\sum_{i=1}^{5,686} (d_i)^0}{5,686} = \frac{539}{5,686} = 0.0948.$

When the threshold value $\tau = 3.08$, the energy insecurity gap is equal to

(27) $d^1 = \frac{\sum_{i=1}^{5,686} (d_i)^1}{5,686} = \frac{231.65}{5,686} = 0.0407$

and the severity of energy insecurity is equal to

(28) $d^2 = \frac{\sum_{i=1}^{5,686} (d_i)^2}{5,686} = \frac{103.75}{5,686} = 0.0182.$

Changing the threshold value of τ to 3.08, that is considering households as energy secure if they respond affirmatively to either no questions or only one question, the percentage of households who identify as energy insecure decreases by over half. Rather than 22% of households identifying as energy insecure, now only 10% of the households identify as energy insecure. Our results now indicate only 4% of households will fall below the energy security threshold, and those that do are only 2% more energy insecure than households who on average are above the energy security threshold.

Because the severity parameter estimates for some questions (e.g., Question 5 [HVAC], Question 2 [Unsafe], and Question 3 [Notice]) are difficult to differentiate from one another, following the literature on food security (Hamilton et al. 1997; Bickel et al. 2000) we combine household responses to combinations of questions with similar severity parameter estimates to create four mutually exclusive energy security categories. The severity parameter estimates produced from the Dichotomous Rasch model are determined by the percentage of households who respond positively to the questions. Therefore, severity estimates for individual questions that are statistically indistinguishable from one another, are likely being responded to in the affirmative by approximately the same number of households in our sample.

Under the method of partitioning described above, households are placed into one of four groups. Households who respond affirmatively to none or only one question are considered "High Energy Secure" and are assigned an energy insecurity index value = 1 ($EISINDEX_i = 1$). Households who respond affirmatively to two, three, or four questions are considered "Marginally Energy Secure" and are assigned an energy insecurity index value = 2 ($EISINDEX_i = 2$). Households who respond affirmatively to five questions are considered "Low Energy Secure" and are assigned an energy insecurity index value = 3 ($EISINDEX_i = 3$). Lastly, households who respond affirmatively to six or seven questions are considered "Very Low Energy Secure" are assigned an energy insecurity index value = 4 ($EISINDEX_i = 4$).

The energy insecurity index values correspond to different areas along the energy insecurity scale. Figure 3.4 below provides a graphical representation of this method of partitioning households into different groups.

Energy Secure	**Item Severity**	**Number of Households**	
	0	4,442	High Energy Secure $EISINDEX_i = 1$
	3.08	705	
	4.38	290	Marginally Energy Secure $EISINDEX_i = 2$
	4.44	146	
	4.48	77	
	5.65	22	Low Energy Secure $EISINDEX_i = 3$
	6.20	4	Very Low Energy Secure $EISINDEX_i = 4$
Energy Insecure	6.34	0	

Figure 3.4 Energy Insecurity Scale Results with Four Groups

Because responses to either none or zero questions indicate the household is "High Energy Secure" (i.e., not energy insecure) the threshold value of τ is as before set equal to 3.08. The energy insecurity score the household receive is still equal to the difference between their exact placement along the scale and the threshold value of $\tau = 3.08$. The energy insecurity index value ($EISINDEX_i$) however, is no longer equal to difference between the energy insecurity score s_i and the threshold value of $\tau = 3.08$.

Instead, four different energy insecurity index values $EISINDEX_i$ are assigned to households who have scores s_i within a certain range. If the energy insecurity score $s_i \leq 3.08$, then $EISINDEX_i = 1$. If the energy insecurity score is within the range $3.08 < s_i \leq 4.48$, then $EISINDEX_i = 2$. If the energy insecurity score is within the range $4.48 < s_i \leq 5.65$, then

$EISINDEX_i = 3$, and if the energy insecurity score is within the range $5.65 < s_i$, then $EISINDEX_i = 4$. Table 3.29 below provides an outline of these results.

Table 3.29 Normalized Energy Insecurity Index ($\tau = 3.08$) Four Energy Security Categories Identified

Questions Responded Affirmatively To	s_i	d_i	No. of Households	$EISINDEX_i$	Category
0	0	0	4,442	1	High Energy Secure
1	3.08	0	705		
2	4.38	0.40	290	2	Marginally Energy Secure
3	4.44	0.42	146		
4	4.48	0.43	77		
5	5.65	0.79	22	3	Low Energy Secure
6	6.20	0.96	4	4	Very Low Energy Secure
7	6.34	1	0		
Observations	N = 5,686				

When only four energy insecurity index values are considered, because the individual energy insecurity scores (s_i) of the household's remain the same, the normalized energy insecurity d_i also remains the same. Therefore, the extent, depth, and severity of energy insecurity remains the same as before. 10% of the households identify as either Marginally, Low, or Very Low Energy Secure. Only 4% of households are expected to fall below the energy security threshold. Those households that do are approximately 2% more energy insecure than households who on average are above the energy security threshold.

3.7 DISCUSSION OF INDEX RESULTS

Because energy insecurity is unobserved, latent trait we examine the index results above in terms of their ability to achieve content, construct, and convergent validity. By examining the validity of the results, we can evaluate their ability to provide a consistent and accurate representation of the experience of being energy insecure. An index that has content, construct, and convergent validity is preferred to an index that does not. Table 3.30 outlines the results from each of the energy insecurity index measures applied above. The number and percentage of households who identify as being "Energy Insecure," according to each individual index measure are listed in the third column.

Table 3.30 Energy Insecurity Index Results: Number of Households in Study Sample (N=5,686) Identified as Energy Insecure

Energy Insecurity Metric	Description	Number of Households (%)
Energy Assistance	Applied for and Received Home Energy Assistance (LIHEAP)	345 (6%)
Expenditure Approach	$>$ 10% of Income	487 (6%)
	$>$ 6% of Income	999 (18%)
Cluster Analysis	High Energy Secure	4,832 (85%)
	Marginally Energy Secure	634 (11%)
	Low Energy Secure	144 (3%)
	Very Low Energy Secure	76 (1%)
Principal Components Analysis	Energy Insecurity Score $s_{i,pca} > 0$	1,244 (22%)
Dichotomous Rasch Model ($\tau = 0$)	Energy Insecurity Scale Value > 0	1,244 (22%)
Dichotomous Rasch Model ($\tau = 3.08$)	Energy Insecurity Scale Value > 3.08	539 (9%)
Dichotomous Rasch Model Four Groups	High Energy Secure	5,147 (91%)
	Marginally Energy Secure	513 (8.53)%
	Low Energy Secure	22 (0.4%)
	Very Low Energy Secure	4 (0.07%)

A quick comparison of the results in Table 3.29 reveals that the number of households who identify as "Energy Insecure" changes drastically depending on which index metric is used. For example, when using the expenditure approach ($>$ 6% of Income) to identify energy insecure households, nearly three times as many households identify as being "Energy Insecure" than when the receipt of home energy assistance is used to identify energy insecure households. Furthermore, the number of households who identify as "Energy Insecure" because they spend more than 6% of the income of energy/fuel is nearly double that of the number of households who identify as "Energy Insecure" because they spend more than 10% of their income on energy/fuel.

The cluster analysis results suggest nearly 15% of the households surveyed in the 2015 RECS identified as being either low, very low, or marginally energy secure, while 85% identify as being highly (i.e., completely) "Energy Secure." Households who identify as low, very low, or marginally energy secure are all considered to be energy insecure to some degree. Very low energy secure households are considered the most energy insecure. Only 76 households in our sample were identified as "Very Low Energy Secure" households. While twice as many households were identified as "Low Energy Secure" households, the two together (144 households Low Energy Secure households and 76 Very Low Energy Secure households) represent roughly only half the number of households who are identified as being "Energy Insecure" when the receipt of home energy assistance is used to identify energy insecure households.

The results from the application of cluster analysis are somewhat consistent with the results from the application of PCA and the Dichotomous Rasch Model. However, the application of PCA and the Dichotomous Rasch model suggest more households in our sample (22% of those surveyed when the threshold value of τ is set equal to 0) identify as being "Energy Insecure." When the threshold is changed to 3.08, the number of households who identify as being "Energy

Insecure" according to the Rasch model decreases by more than half; an indication that the choice of threshold has an effect on which households are considered "Energy Insecure" and which households are considered "Energy Secure." As stated earlier, one of the objectives of this chapter is to identify a single, universally accepted metric that can be used to determine if households are energy secure or energy insecure. To determine which of the preceding index measures should be used we examine the "validity" each index measure.

The general concept of validity is defined as "the degree to which a test measures what it claims, or purports, to be measuring" (Brown, 1996, p. 231). Therefore, an index should be considered "valid" if it measures the underlying construct it was designed to measure, which in our case is a household's true level of energy insecurity (Bucher 2014, DeVellis 2003).[58] In order for the energy insecurity index measures we estimate to be considered valid they must provide a consistent and accurate representation of what it means for a household to be energy insecure. That is, they must provide insight into whether or not a household is able to maintain consistent physical and economic access to a sufficient, safe, and affordable energy supply to meet each household member's most basic daily energy service needs.

To assess the validity of the different index measures outlined above, we consider how they perform in terms of achieving: 1) Content Validity; 2) Construct Validity; and 3) Convergent Validity. Content validity is a pre-requisite for both construct and convergent validity (Aravamudhan and Krishnaveni 2015). Therefore, content validity should be considered first when determining whether or not an index measure for household energy insecurity is "valid." Content validity refers to how well an instrument (i.e., a test or set of survey questions) measures the

[58] For example, a math test (instrument) would be considered valid if it accurately measured a student's mathematical ability.

theoretical construct the instrument intends to measure. For example, how well does a math test measure a student's mathematical ability. [59] Content validity is a critical step in the development of a new measurement scale, as it can help to establish a mechanism (e.g., specific set of survey questions) that links an abstract, latent concept, such a household's level of energy insecurity, with observable traits and measurable indicators (Wynd, Schmidt, and Schaefer 2003). To test the content validity of an index measure, typically recognized subject matter experts are assigned to evaluate whether or not the items used to construct the specific index provide an accurate representation of the underlying, latent construct (Lawshe 1975; Aravamudhan and Krishnaveni 2015). Items are evaluated by experts based on their relevance and representativeness of the latent construct (Lawshe 1975; Beck and Gable, 2001; Aravamudhan and Krishnaveni 2015). It is assumed that the higher the rating of relevance and representativeness of the items provided by the expert, the higher the content validity of any single item (Aravamudhan and Krishnaveni 2015).

Using each experts rating of the items, a content validity index and content validity ratio for each individual item j from the subset of items can be constructed. The content validity index for each individual item j refers to the percentage of experts who rate the item as both relevant and representative. Items are considered relevant if they provide an appropriate depiction of the underlying construct. For example, questions related to reading comprehension would not be considered relevant for a mechanism designed to measure mathematical ability. Items are considered representative if they provide an overall depiction of the underlying construct. For example, all items on a math test related to addition would be representative of a student's ability to add numbers.

[59] Other examples might include how well an IQ test measures a person's true intelligence or the Test of Economic Literacy (TEL) which is a nationally-normed and standardized test for measuring the economic understanding of U.S. high school students (Walstad, Rebeck, and Butters 2013).

A content validity index value of 1 is considered evidence of content validity if fewer than five experts are examining an item (Polit, Beck, and Owen 2007). A content validity index value of .8 should be considered proof of an item's content validity if five or more experts are evaluating the item (Polit, Beck, and Owen 2007). The content validity of an individual item j can also be judged according to its content validity ratio (CVR) (Lawshe 1975). The CVR for an individual item j is calculated as follows:

(29) $$CVR_j = \frac{N_e - (N/2)}{(N/2)}$$

where N_e is the number of experts who agree that the individual item is "essential," and N is the total number of experts (Lawshe 1975). Content validity ratios range in value from -1 to +1 (Lawshe 1975). A negative CVR value indicates fewer than half of the experts rated the item as "essential," while a positive CVR values indicates more than half the experts rated an item as "essential" (Lawshe 1975). While the CVR and content validity index provide two statistical approaches to examine the content validity of questions from the 2015 RECS we used to construct the different energy insecurity indices presented in Table 3.30, access to an expert panel was not feasible for our study.

Therefore, instead of using an expert panel, we examine the content validity of the individual questions used to construct each index in terms of their ability to provide the most accurate representation of the experience of being energy insecure. Recall from Section 3.1, that a state of being energy secure is defined as a state of having *consistent physical* and *economic* access to a *sufficient, safe, and affordable* energy supply to meet the basic daily energy service needs of members of the household. Based on this definition, a state of being energy insecure refers to a state of *not* having consistent physical and economic access to a sufficient, safe, and affordable energy supply to meet basic daily energy service needs. For an energy insecurity index measure to have content validity, each of the items used to construct it must be relevant and provide an accurate representation of a household's *inability* to maintain both physical and economic access to a sufficient, safe, and affordable energy supply to meet daily energy service needs.

The items used to construct the energy insecurity index for the receipt of home energy assistance and the expenditure approach only refer to the household's ability to maintain consistent *economic* access to an *affordable* energy supply to meet daily household energy service needs. Thus, the questions used to construct these two index measures for household energy insecurity do not provide insight into the sufficiency or safety of the household's energy supply. Nor do they provide insight into whether or not adequate energy service needs were met. Rather, they focus only on whether or not households struggled to *afford* their home energy bills. Therefore, we infer that the content validity of the expenditure approach and receipt of home energy assistance as a measure of energy insecurity is inadequate.

Cluster analysis, PCA, and the Dichotomous Rasch model are all applied to household responses to the set of questions listed in Table 3.1. These questions cover a wide range of circumstances that could impact a household's ability to maintain consistent physical and economic access to a sufficient, safe, and affordable energy supply. For example, Question 2 (Unsafe) asks households to report on whether or not they have kept their home at a temperature they felt was unsafe or unhealthy. Households who respond affirmatively to this question are unable to consistently maintain access to a *safe* energy supply.

Question 1 (Reduce) and Question 3 (Notice) relate to the affordability of a household's energy supply, Question 4 (No Fuel), Question 5 (HVAC), and Question 7 (Days) relate to the sufficiency of the household's energy supply, and Question 6 (Medical) relates to the safety of the household's energy supply. Together, responses to these questions cover the entire construct of household energy insecurity. That is, a household's inability to maintain consistent *physical and economic access* to a *sufficient, safe, and affordable* energy supply to meet basic daily household energy service needs. It is important to note that questions from prior iterations of the RECS have also been used to construct an energy insecurity index (see Murray and Mills 2012). We consider this an indication that the questions contain at least some content validity.

In addition to content validity, an index is assumed to have construct validity if it accurately measures the theoretical, unobservable construct or trait it intends to measure. In our case, construct validity can be used to assesses how each individual energy insecurity index produced corresponds to the theoretical construct of household energy insecurity. Given energy insecurity is an unobservable (i.e., a latent household trait) examining the construct validity of the different index measures is challenging - as it is difficult to determine just how accurately each energy insecurity index gauges the extent and severity of the experience of being truly energy insecure.

To test the construct validity of the different index results listed in Table 3.30 we use a multitrait-multimethod matrix (MTMM), which was first proposed by Campbell and Fiske (1959). An MTMM is a matrix of correlation coefficients that can be used to assess the construct validity of different instruments that are designed to measure the same unobserved trait (Campbell and Fiske 1959; Bagozzi, Yi, and Phillips 1991). If measurement techniques for multiple unobserved traits are being examined, then the diagonal of an MTMM is a measure of the reliability of each technique (Campbell and Fiske 1959; Trochim 2006).

In order for construct validity to be achieved, both convergent and discriminant validity must also be achieved (Campbell and Fiske 1959; Trochim 2006). Discriminant validity is the extent to which concepts that should not be theoretically related to one another are in fact, unrelated to each other (Campbell and Fiske 1959; Trochim 2006). Convergent validity is a subtype of construct validity, that can be used to assess whether or not the individual index measures converge to the same results. Indices that are assumed to measure the same construct should converge to the same results when applied to a single data set.

Achieving convergent validity implies construct validity has also been achieved (Bishop and Boyle 2019). If the results produced from two different index measures are statistically indistinguishable, then this is evidence of convergent validity (Bishop and Boyle 2019). [60]We can assess both discriminant and convergent validity using the MTMM (Trochim 2006). If two indices are presumed to measure the same construct/trait, then they should be related to each other and a positive correlation between them is expected. [61] If two indices are not designed to measure the same construct/trait, then they should not be related to one another and a negative correlation between them is expected.

All of the indices included in Table 3.30 are designed to measure the underlying construct/trait of household energy insecurity. Therefore, a positive correlation among each index measures is expected (Campbell and Fiske 1959). Because we are interested in measuring only one unobserved trait, a household's true level of energy insecurity, the MTMM used for our study is simply a square, symmetric matrix of correlation coefficients. Table 3.31 lists the correlation coefficients for the different index measures produced from the MTMM.

---

[60] It is important to note however, that convergent validity between two indices does not necessarily imply that the indices provide a valid measure of household energy insecurity. It is possible for two or more indices to be statistically indistinguishable from another, but produced biased estimates of the underlying construct they intend to measure (Bishop and Boyle 2019). Similarly, finding a statistically significant difference between the results of two competing indices does not provide conclusive evidence that one or both of the indices are invalid (Bishop and Boyle 2019).

[61] For example, one way to examine the construct validity of an aptitude test (i.e., IQ test) would be to see how correlated the outcomes on the aptitude test are with outcomes from another, similar test believed to measure aptitude.

Table 3.31 Correlation Coefficients: Energy Insecurity Index Measures

	Exp. >10%	Exp. >6%	Energy Assistance	Cluster Analysis	PCA	Rasch $\tau = 0$	Rasch $\tau = 3.08$	Rasch 4 Groups
Exp. > 10 %	1.00	-	-	-	-	-	-	
Exp. > 6%	0.663	1.00	-	-	-	-	-	
Energy Assistance	0.117	0.165	1.00	-	-	-	-	
Cluster Analysis	0.104	0.134	0.131	1.00	-	-	-	
PCA	0.205	0.223	0.201	0.479	1.00	-	-	
Rasch $\tau = 0$	0.198	0.225	0.201	0.504	0.922	1.00	-	
Rasch $\tau = 3.08$	0.180	0.189	0.164	0.339	0.894	0.756	1.00	
Rasch 4 Groups	0.178	0.187	0.160	0.339	0.884	0.751	0.998	1.00

As expected, each of the index measures we examine is positively correlated with the other index measures. The index results from both of the Rasch models are highly correlated with the index results from PCA. If an index measure has construct validity (i.e., provides an accurate representation of what it means to be energy insecure), then it should in practice, behave the way we anticipate it to conceptually. That is, the index should be positively related to variables that are positively related to energy insecurity and negatively related to variables that are negatively related to energy insecurity.

Because energy insecurity refers to a state of not having consistent access adequate energy services, we theoretically anticipate an energy insecurity index measure will be negatively related to median household income. In addition, we anticipate, based on prior research (see Drehobl and Ross 2016) that the energy insecurity index results will be positively correlated with Hispanic and African American households and households with more children (Hernandez 2016). Lastly, we assume energy insecurity will be negatively related to home ownership, as households who are owners typically have higher incomes making them less likely to struggle to afford energy services.

Table 3.32 below presents a set of correlation coefficients between median household income and the different energy insecurity index results from Table 3.30.

Table 3.32 Correlation Coefficients: Energy Insecurity Indices and Household Characteristics

	Median Income	African American	Hispanic	Ownership	Children
Exp. Approach (10%)	-0.367	0.108	0.023	-0.111	0.032
Exp. Approach (6%)	-0.506	0.150	0.032	-0.164	0.016
Energy Assistance	-0.196	0.075	0.031	-0.091	0.056
Cluster Analysis	-0.180	0.080	0.058	-0.120	0.061
PCA	-0.258	0.136	0.098	-0.136	0.110
Rasch Model ($\tau = 0$)	-0.282	0.148	0.109	-0.160	0.118
Rasch Model ($\tau = 3.08$)	-0.202	0.091	0.078	-0.106	0.075
Rasch Model 4 Groups	-0.200	0.088	0.078	-0.105	0.073
Observations	5,686				

Overall, we find median household income is negatively related to household energy insecurity. The correlation is strongest between households who spend more than 6% of their income on energy/fuel. Each of the energy insecurity indices is positively related to households who identify as African-American or Hispanic, and the number of children under the age of 16 living in the home. The energy insecurity indices are all negative correlated with home ownership. All of these relationships are consistent with conceptual expectations. The observed expected relationships are stronger when the Rasch model results are used to classify households as “Energy Secure” or “Energy Insecure.”

Convergent validity can also be tested by examining whether or not the different index measures we produce identify the same households as being energy insecure. Table 3.33 below provides information on the number of households in our study sample (N=5,686) who identify as being "Energy Insecure" according to each of the different index measures we estimate. Each cell in Table 3.33 can be interpreted as the number of households who identify as "Energy Insecure" across the two corresponding measures of energy insecurity. For example, using the expenditure approach, assuming a 10% threshold beyond which energy/fuel expenditures cease to be affordable, and the receipt of home energy assistance to identify energy insecure households results in the same 74 households being identified as energy insecure.

Table 3.33 Number of Households in Study Sample (N=5,686) Identified as Energy Insecure Using Each of the Different Energy Insecurity Metrics

	Exp. >10%	Exp. >6%	Energy Assistance	Cluster Analysis	PCA	Rasch $\tau = 0$	Rasch $\tau = 3.08$	Rasch 4 Groups
Exp. > 10 %	487	-	-	-	-	-	-	
Exp. > 6%	487	999	-	-	-	-	-	
Energy Assistance	74	146	345	-	-	-	-	
Cluster Analysis	167	307	135	854	-	-	-	
PCA	227	409	183	854	1,244	-	-	
Rasch $\tau = 0$	227	409	183	854	1,244	1,244	-	
Rasch $\tau = 3.08$	132	217	99	485	539	539	539	

Rasch 4 Groups	132	217	99	485	539	539	539	539

The application of PCA and the Dichotomous Rasch model both identify the same 1,244 households in our sample as being energy insecure. However, the "weights" applied to the subset of questions listed in Table 3.1 and used to generate the different index results, vary drastically from one another. In fact, they are almost the exact opposite of one another. The "weights" produced from the application of the Dichotomous Rasch model are interpreted as severity parameter estimates of the individual questions. The more difficult/severe the question, the higher the severity parameter estimate associated with the specific question. Severity parameter estimates are added together to produce energy insecurity scores for each individual household i. A household's energy insecurity score determines its energy insecurity status. Higher energy insecurity scores are assumed to be associated with higher levels of household energy insecurity.

The application of PCA also produces a set of "weights" for each of the individual questions/items j. These weights are interpreted as factor scores. Similar to the Dichotomous Rasch model, the factor scores are combined together to produce an energy insecurity score for each household i, which is used to determine a household' energy insecurity status. However, the "weights" produced from the application of PCA do not provide an estimate of the severity of question to which they correspond. Instead, the weights are simply correlation coefficients. Therefore, the weights measure how household responses to the individual questions vary with one another. Higher/larger weights imply more households responded affirmatively to the same question.

Similarly, the Rasch model results identify the same number of households (539 households) as being energy insecure when the threshold value of $\tau = 3.08$ was used and the households are divided into four different insecurity categories. This result is a function of the fact that the threshold value of $\tau = 3.08$ does not change between the two index measures. The only difference between the two index measures is the number of energy insecure categories identified. Similarly, 539 of the 1,244 households identified as energy insecure when the threshold value of $\tau = 0$ was used also identify as energy insecure when the threshold value of $\tau = 3.08$ was used and/or the households are divided into four different energy insecurity groups.

Cluster analysis identifies only 220 households as energy insecure (i.e., low energy secure or very low energy secure). Cluster analysis is applied to responses by households to the same questions as PCA and the Dichotomous Rasch model. The application of cluster analysis identifies 854 of the 1,244 households identified as energy insecure from PCA and the Dichotomous Rasch model also as energy insecure. Thus, the energy insecurity index produced from cluster analysis lacks convergent validity with PCA and the Dichotomous Rasch model results.

The index measure results produced from the expenditure approach and the receipt of home energy assistance do not converge to the same results. Using the expenditure approach assuming 6% of disposable income as the threshold, we identify 487 households as energy insecure. Only 15% of these 487 households (74 households) responded affirmatively to having applied for and received home energy assistance. Comparing these results to the results of cluster analysis, we find 307 households who identify as being marginally energy insecure, low energy secure, or very low energy secure spent more than 6% of their disposable income on energy/fuel. Only 135 of these same households applied for and received home energy assistance. Across all the different index measures, there are large disparities between the number of households who identify as energy insecure, suggesting that the index results do not converge to the same results.

In addition to examining the content, construct, and convergent validity of the different index results, we also examine the internal consistency of the different survey questions used to construct each index measure. For internal consistency to hold, two or more questions designed to measure the same concept should produce similar results (Sullivan 2011). The seven questions included in Table 3.1 that are used in the application of PCA, the Dichotomous Rasch model, and cluster analysis are the only questions of their kind included in the 2015 RECS. Therefore, we are unable to check the internal consistency of these questions.

There are however three separate questions that ask households whether or not they have received home energy assistance over the past twelve months. They include:

1. Has the household participated in a home energy assistance program that helps pay energy bills and/or replace/fix broken HVAC? (yes/no)
2. Has the household applied for and received energy assistance to help pay home energy bills after receiving a disconnection notice? (yes/no)
3. Has the household applied for and received energy assistance to help restore heating and/or cooling in the home? (yes/no)

For the receipt of home energy assistance to be an internally consistent measure for energy insecurity, then the same pattern of responses by households should be consistent across the three questions. Results in Tables 3.3 to 3.7 show inconsistencies in the number of households who respond affirmatively to all three questions; an indication that the receipt of home energy assistance is not a consistent measure for energy insecurity.

3.8 CONCLUSIONS

Accurate identification of energy insecure households can lead to the implementation of more effective policies and programs that better address the needs of these households. In this chapter, we compare and contrast five different approaches for identifying energy insecure households, in hopes of creating a single, uniform index measure of household energy insecurity. The five different approaches we consider include: 1) whether or not the household has applied for and received home "energy assistance;" 2) whether or not the household has spent more than 10% or 6% of their disposable income on fuel/energy; 3) cluster analysis; 4) principal components analysis (PCA); and 5) a Dichotomous Rasch model.

We rely on household level data from the 2015 RECS to construct an energy insecurity index measure using each of the five approaches outlined above. Using the receipt of home energy assistance as a measure of household energy insecurity, only 345 households (6%) of the households in our sample were identified as energy insecure. When the expenditure approach is used to determine if a household is energy insecure or energy secure assuming 10% as the threshold beyond which energy/fuel expenditures cease to be affordable, results in almost 20% of the households in our sample being identified as energy insecure.

However, as mentioned in Section 3.7, using the expenditure approach or the receipt of home energy assistance to determine a household's energy insecurity status only focuses on one aspect of household energy insecurity- affordability. Therefore, these two measures fail to provide a complete and accurate representation of what it means for a household to be energy insecure versus energy secure. In an effort to provide to a more accurate representation of the construct of household energy insecurity, we rely on cluster analysis, PCA, and the Dichotomous Rasch model. We apply cluster analysis, PCA, and the Dichotomous Rasch model to a subset of questions included from the 2015 RECS (see Table 3.1) that are believed together to measure the underlying construct of household energy insecurity.

Cluster analysis separates households into different energy insecurity groups based on their pattern of responses to a set of questions. Households whose responses are most similar to one another are placed in the same group. From the application of cluster analysis, we find the majority (85%) of the households in our sample identify as being "High Energy Secure." The remaining households in our sample identify as being either "Marginally Energy Secure," "Low Energy Secure," or "Very Low Energy Secure." While the application of cluster analysis is convenient because it allows us to separate households into four mutually exclusive groups, the accuracy of the results is difficult to determine given that very different clusters can be formed from the same data depending on the similarity measure specified.

The fourth and fifth approaches, PCA and the Dichotomous Rasch model produce energy insecurity scores for each individual household i. These energy insecurity scores are used to determine if households are energy insecure or energy secure. The underlying assumption of the Dichotomous Rasch model is that the probability that a household responds affirmatively to any given question depends on the degree and extent of the household's latent energy insecurity status. The scale scores produced from the Dichotomous Rasch model ranged from 0 (no affirmative responses) to 6.34 (7 affirmative responses). Conversely, the assumption underlying the application of PCA is that responses to questions that are strongly correlated with one another, vary together.

Overall, we find more households identify as being energy insecure when we consider household responses to more than one question. For example, when using the receipt of home energy assistance as a proxy for energy insecurity, only 6% of the households in our sample identify as being energy insecure. This number nearly quadruples when responses to multiple questions are considered and used to determine a household's energy insecurity status following the PCA and Dichotomous Rasch Model approaches. The seven questions included in Table 3.1 seem to provide a detailed and accurate description of what it means for a household to be energy insecure. Together the questions cover aspects of safety, affordability, and sufficiency of an energy supply to meet daily energy service needs.

The index results produced from the Rasch model appear to provide the most consistent and accurate representation of the experience of being energy insecure. These results utilize the individual severity parameter estimates for all seven questions to generate energy insecurity scores for each household in our sample. The energy insecurity scores are used to rank households along the energy insecurity scale, such that households ranked higher on the scale are more energy insecure. The Rasch model index results intuitively make sense, as households who are energy secure are likely not going to respond affirmatively to any of the questions or only a few of the questions in Table 3.1.

These results are positively correlated with the other index measures we consider and correlated with covariates of interest as expected theoretically indicating they possess construct and convergent validity. Because this is the third iteration of the RECS to include these types of questions, we anticipate the questions from the 2015 RECS have more content validity than similar questions from prior iterations of the RECS. However, only by having an expert panel rate the "relevance" and "representativeness" of the new questions can we reach such a conclusion. We suggest future research should explore the idea of having an expert panel rate and review each question used to measure energy insecurity. Because of the evidence indicating that the Rasch model results provide a valid representation of a household's ability to maintain access to adequate energy services needed by the household to feel energy secure, we suggest the Rasch model be used to measure household energy insecurity in the United States.

This study had some limitations. First, despite the clear advantage of using multiple questions to construct the energy insecurity index, the questions included in the 2015 RECS all measure energy insecurity at the household level, and not at the individual level. This limits our ability to measure any variation in energy insecurity among individuals living in the same household. In addition, although the RECS is conducted once every few years by the EIA, the data produced from each iteration can only be treated as a single cross-section. The subset of questions used to measure household energy insecurity have also changed between each iteration. As a result, comparison of our results to the results of prior iterations is problematic. Nevertheless, this study is an important contribution toward obtaining an accurate measure of household energy insecurity in the United States. We suggest future research be aimed at examining how questions from these different iterations of the RECS can either be combined together or extracted from to create a consistent set of questions that can be used from year to year to measure household energy insecurity.

CHAPTER 4

ESSAY 3: EXAMINING THE THEORETICAL AND EMPIRICAL RELATIONSHIPS BETWEEN HOUSEHOLD ENERGY EFFICIENCY AND ENERGY SECURITY*

* Harker Steele, Amanda. To be submitted to the *Journal of Environmental Economics and Management.*

This chapter has benefited from comments received from the 24th Annual Camp Resources, the 6th World Congress of Environmental and Natural Resource Economics (WCERE) and the 2018 AAEA Annual Meetings.

ABSTRACT

Interest in energy efficiency investments has increased substantially over the past two decades as public policy and decision makers attempt to find low cost solutions to reducing global greenhouse gas emissions. Moreover, given their ability to reduce household energy expenses, energy efficiency investments have been suggested as a way to improve household energy security in the developed world. Such connections however, have largely been ignored in the applied economics literature. The aim of this paper is to develop a theoretical model and empirical procedure which can be used to examine the relationship between making an in-home energy efficiency improvement and a household's energy security status. To do this, we rely on the theory of household production to first capture a household's demand for, and production of, energy services. We then utilize a stochastic production frontier approach to explain why households who are inefficient in their production of energy services might choose to invest in an energy efficiency improvement or upgrade. We then explain how the return to such an investment could lead to higher levels of household energy security. To empirically test our hypothesis, we use responses from the 2015 Residential Energy Consumption Survey conducted by the U.S. Energy Information Agency. We examine factors that influence subjective feelings of household energy insecurity, focusing specifically on in-home energy efficiency improvements made in the past year.

Keywords: Household Energy Security, Energy Efficiency Improvements, Household Production Theory, Stochastic Production Frontier

4.1 INTRODUCTION

Over the past decade investments in demand-side energy management programs, particularly those focused on end-use energy efficiency, have skyrocketed as a result of policies aimed at reducing energy consumption and mitigating the harmful effects of global climate change. From taking advantage of a utility-sponsored in-home energy audit to replacing home appliances with more energy efficient models, the opportunities to become a more energy efficient household are more prevalent now than ever before. Moreover, given their ability to reduce household energy expenses, investments in energy efficiency have also been suggested as one possible strategy for improving a household's energy security status in the developed world (Reames 2016).

Despite this suggestion, few theoretical or empirical investigations exist that have specifically analyzed the relationship between making an in-home energy efficiency improvement and a household's energy security status. The overall objective of this paper is to help fill this gap in the literature by addressing the research question: "Do energy efficiency investments have a significant, positive effect on a household's energy security status?" Recall from Chapter 3 (Essay 2) that we assume to be energy secure a household must have consistent physical and economic access to a sufficient, safe, and affordable energy supply that meets each household members' most basic daily energy service needs. [62] In other words, in order to be energy secure, a household has consistent access to adequate energy services.

Because energy efficiency improvements are designed to reduce the amount of energy necessary to provide household energy services, in theory they should make the provision of household energy services more affordable and thus help to alleviate the presence of household

[62] The definition provided for household energy security is based on the definition of food security as defined by the United States Department of Agriculture (USDA). More specifically, the USDA defines household Food Security as having access at all times to enough food for an active, healthy life.

energy insecurity (i.e., a lack of household energy security). [63] However, because in-home energy efficiency improvements lower the overall price of providing household energy services, making such investments (e.g., purchasing a more energy efficient home appliance) could lead to an increase in the demand for household energy services by members of the household (Gillingham, Rapson, and Wagner 2016). This phenomenon is known as the "rebound effect" and has been the subject of much debate surrounding the implementation of in-home energy efficiency improvements (Gillingham et al. 2013; Gillingham et al. 2016).

While this study does not necessarily focus on the rebound effect, it does make two main contributions to the literature on energy efficiency and household energy security. First, to the best of our knowledge, this study is one of the first to follow-up on the suggestion by Hernandez (2013) to theoretically and empirically examine whether or not making in-home energy efficiency improvements helps to alleviate the presence of household energy insecurity. While there has been some recent work by Fowlie et al. (2018) that examines whether or not investing in household energy efficiency leads to actual cost savings, a direct connection to a household's energy security status is not the main motivation of the paper. [64]

Instead Fowlie et al. (2018)'s work is primarily motivated by the presence of the "energy efficiency gap," a phenomenon used to define the difference between the cost-minimizing level of energy efficiency achievable and the level of energy efficiency actually achieved by a household. Furthermore, unlike Fowlie et al. 2018, our study does not utilize a field experiment, but rather takes a top-down approach to examining the outcome of investing in energy efficiency.

---

[63] For a complete description of what it means to be energy secure see Chapter 3.

[64] More specifically, Fowlie et al. (2018) analyzes the cost-savings from participating in the Weatherization Assistance Program (WAP), a federally funded program designed to make energy costs more affordable for low-income households by increasing the energy efficiency of their homes.

In addition to the work of Fowlie et al. (2018), there have been numerous other studies that have focused specifically on the outcomes of investing in demand-side energy efficiency. Examples include work by Levinson (2016) who estimated energy cost-savings from implementing building code standards in California; Allcott and Greenstone (2017) who examine whether or not imperfect information or behavioral biases impact consumers' decisions to take advantage of and make suggested changes from receiving an in-home energy audit; and Novan and Smith (2018) who consider the role of electricity rate structures on incentivizing investments in energy efficiency. However, these previous studies do not directly examine the connection between a household's energy security status and making energy efficiency improvements.

Second, given our definition of household energy security, our study provides a unique theoretical approach for examining the relationship between making an in-home energy efficiency investment and a household's energy security status. We begin by considering the household's decision to produce and consume household energy services as a two-stage optimization problem following the theory of household production (Becker 1965; Deaton and Muellbauer 1980). The solution to the household's problem yields a certain level of satisfaction (i.e., utility) for the household, which we assume can be used to represent the household's underlying latent energy security status (i.e., subjective feelings of energy security/ insecurity).

To examine how the decision to make an investment in energy efficiency will impact a household's energy security status, we rely on a stochastic production frontier approach. Under the stochastic production frontier approach, we assume one reason a household may identify as energy insecure is because they are technically inefficient in their production of household energy services. One cause of their technical inefficiency in production may be the current low technical efficiency ratings of the capital technology inputs they currently employ. By adopting more energy efficient capital technology or making an energy efficiency improvement in the home (e.g., by sealing air leaks or adding insulation) a household will be able to reduce technical inefficiency in its current production of household energy services. By reducing technical inefficiency, the household should therefore be able to produce more energy services and reach a higher level of satisfaction. This increase in satisfaction, in turn, results in greater subjective feelings of energy security (lower subjective feelings of energy insecurity).

It is difficult however, for a researcher to directly observe the production and consumption of energy services by members of the household. As a result, determining whether or not a household is energy secure or energy insecure based on its ability to produce and consume household energy services is problematic. Following the literature on food security, we determine a household's energy security status by combining household responses to a set of survey questions included in the 2015 Residential Energy Consumption Survey (RECS). The questions we utilize are designed specifically to capture a household's subjective feelings about their energy security/insecurity status.

Household responses to the questions are reflective of a household's latent, unobserved, level of energy security/insecurity. [65] In particular, the extent or severity of energy insecurity experienced by an individual household is reflected by the number of energy insecurity questions from the 2015 RECS to which a household provided an affirmative response. The energy insecurity index measures we use from Chapter 3 (Essay 2) are based on the Dichotomous Rasch model. The Dichotomous Rasch model results are used to place households along a latent energy insecurity scale. The energy insecurity scale ranges in value from the least severe indicator of energy

[65] For a complete explanation of the methods used to separate households into different energy security categories see Chapter 3 (Essay 2).

insecurity (a value of “0” which indicates zero affirmative responses to any questions) to the most severe indicator of energy insecurity (the value of the item calibration for the most severe energy insecurity question from the 2015 RECS).

For household’s who identify as being energy insecure, which may result from being technically inefficient in their production of energy services, the decision to make an energy efficiency investment is modeled using a random utility framework. Under the random utility framework, a household will make an energy efficiency investment or upgrade if the utility it receives from making the investment or upgrade exceeds the utility it would have received if it did not make the investment or upgrade. Using our Rasch model index, a household’s energy security/insecurity status can only take on a fixed, categorical set of values. Because of the categorical nature of our Rasch model index, we empirically estimate the random utility model using an order logit model econometric specification.

The remainder of this chapter (essay) is organized as follows. Section 4.2 provides some additional background information on the role energy efficiency investments or upgrades play in reducing household energy insecurity. Section 4.3 presents the theoretical framework we utilize to examine the connections between the household’s production and consumption of household energy services and a household’s latent energy insecurity status. In Section 4.4 we discuss our methodological approach including the data used for our empirical analysis, our econometric model estimation procedures, and how we address concerns of potential endogeneity. In Section 4.5 we present our estimation results and discuss policy implications and study limitations. We provide some conclusions in Section 4.6.

4.2 ENERGY EFFICIENCY MECHANISMS

Energy efficiency mechanisms, which have been around since the 1970s, are currently being utilized by electric utility companies across the United States to meet emissions reductions targets, reduce energy production costs, and to help lower their customers home energy burdens (i.e., energy expenses as a share of disposable income) in hopes of achieving higher levels of household energy security (Barbose et al. 2013). The latter has become exceedingly important as maintaining household energy security has become increasingly difficult for many families living in the United States (Hernandez 2016).

Those struggling the most to achieve energy security include low-income and other economically marginalized households, such as those with elderly and differently abled individuals living in the home (Wilkinson et al. 2001; Hernández 2013; Drehobl and Ross 2016; O'Mera 2016). From an economics point of view, low-income and other economically marginalized households may struggle to achieve energy security for three main reasons. First, due to credit constraints, many low-income and economically marginalized households are renters or temporary residents. Thus, they lack the financial incentives to invest in structural improvements (e.g., sealing air leaks, replacing windows) that could, over the long-run, increase the energy efficiency of their home and as a result increase their energy security status.

Second, even if these households are motivated to invest in energy efficiency improvements as a result of being energy insecure, many are unable to afford the high upfront costs associated with purchasing or financing such investments, due to other constraints on their overall budget. In other words, the high upfront costs associated with making an in-home energy efficiency investment or upgrade creates a barrier toa doption for these types of households (Kapur et al. 2011). Third, compared to the general population, low-income and other economically marginalized households are more likely to inhabit residences that are older and overall less energy efficient (Drehobl and Ross 2016). As a result, their homes require more energy to provide energy services such as heating and/or cooling (Penney and Kloer 2015). Consuming more energy can lead to an increase in one's home energy burden, resulting in greater feelings of energy insecurity.

One general reason households who are energy insecure fail to invest in energy efficiency upgrades is imperfect and inadequate information (Scott et al. 2008; Allcott and Greenstone 2012). For example, households may be unaware of how poorly insulated their home is and as a result, may not choose to invest in adding insulation or other types of weatherization. Furthermore, compared to other products, energy efficient products typically have higher upfront costs. For example, different housing units often have different levels of insulation, which directly contribute to how much energy the housing unit consumes to stay warm/cool.

If an individual is considering renting two different housing units, then without adequate knowledge of each housing unit's level of insulation, the individual may only evaluate the two alternatives based on their overall price and otherwise noticeable amenities (e.g., location). Failure to account for the efficiency gains from having adequate insulation can lead households who are energy insecure to underinvest in energy efficiency. Nevertheless, while barriers are believed to exist that prevent households from investing in energy efficiency, a theoretical and empirical examination of whether or not making in-home energy efficiency improvements actually helps to alleviate the presence of household energy insecurity has yet to be adequately addressed (Hernandez 2013).

4.3 THEORETICAL FOUNDATION

In the following section of this essay, we present the theoretical model used to examine how making energy efficiency improvements and/or upgrades in the home could impact a household's self-reported energy security status. We begin by modeling the household's decision to produce and consume household energy services as a two-stage optimization problem using the theory of household production (Becker 1965; Deaton and Muellbauer 1980). The solution to the household's two-stage optimization problem yields a certain level of satisfaction (i.e., utility) for the household, which we assume can be used to represent the household's underlying, latent energy security status.

To examine how having access to more "energy efficient" capital technology influences the outcome of the household's two-stage optimization problem, we rely on a stochastic production frontier approach. Under the stochastic production frontier approach, one reason households may not identify as energy secure is because they are inefficient in their production of household energy services due to the low technical efficiency of their current capital inputs. For example, households with air leaks, inadequate insulation, or high energy consuming capital technology equipment, often consume more energy/fuel than necessary to produce an adequate level of household energy services. Therefore, by adopting more energy efficient capital technology, or making energy efficiency improvements in the home, a household should be able to reduce the inefficiency with which it currently produces household energy services resulting in higher felt levels of household energy security and overall satisfaction.

HOUSEHOLD CONSUMPTION AND PRODUCTION OF ENERGY SERVICES

To develop a theoretical framework that allows us to address the determinants of household energy security, we first establish a set of assumptions about the structure of and operations within the household. Following the work of Filippini (1999), Boogen, Datta, and Filippini (2014), and more recently Burnett and Madariaga (2017) we assume a household's demand for energy (i.e., the fuel resources it utilizes including electricity, natural gas, heating oil, or propane) is derived from its demand for energy services. That is, a household does not necessarily demand energy by itself, but rather demands the energy services (e.g., a warm house, lit room, or cooked food) provided by having access to energy resource inputs.

Continuing to follow the work of Filippini (1999) and Boogen et al. (2014) we assume a household's demand for energy service inputs can be specified using the basic framework of household production theory. According to the theory of household production, households act as both consumers and producers, who in order to produce goods and services use a set of inputs purchased from the market (Becker 1965; Deaton and Muellbauer 1980). The goods and services produced enter the household's utility function directly. That is, outputs which are produced by the household are consumed by members of the household and not sold in the market (Becker 1965). [66]

In our specific case, households combine energy resource inputs (i.e., primary fuel inputs) with energy-using capital technology equipment inside the household to produce household energy services (Filippini 1999; Thompson 2002; Sanstad 2011; Boogen et al. 2014). More specifically, households combine energy resource inputs E (e.g., the primary fuel resources used) with capital technology equipment inputs K to produce a set of household energy services, labeled here as $ESERV$.[67] The production function for household energy services can generally be expressed according to the following quasi-concave household production function:

(1) $ESERV = ESERV[K, E],$

[66] Examples of commodities produced by households might include recreation experiences (Deaton and Muellbauer 1980), leisure (Varian 1992), and meals for consumption (Hamermesh 2007).

[67] We ignore the idea of time dedicated to the production of energy services in our analysis for simplicity.

where, as stated previously, E is used to represent the primary energy resource inputs used by the household (e.g., electricity, natural gas, propane or heating oil) to operate its specific capital technology inputs K (e.g. heating and/or air conditioning unit, refrigerator, stove, lamp, washer, and/or dryer).

Access to more "energy efficient" [68] capital technology impacts the production of energy services within the household since the total amount of energy input consumed by the household to produce household energy services depends on the corresponding efficiency level/rating of the current capital technology the household operates. That is, the output of energy services produced is dependent upon the capital technology's corresponding efficiency level/rating, such that higher efficiency levels/ratings lead to more efficient production of household energy services. Therefore, the capital technology equipment employed by the household can be expressed by the following function:

(2) $K = K(\gamma)$,

where γ is used to represent the specific capital technology's corresponding energy efficiency level/rating.

In (2), an increase in the value of γ leads to an increase in the amount of output per unit of energy resource input consumed (i.e., an increase in efficiency) or, equivalently, a reduced energy input requirement per unit of energy service output produced (Sanstad 2011). To accommodate changes in the efficiency level/rating of the capital technology available for purchase on the market, we assume the value of γ is continuous and strictly positive, such that it exists within the

[68] More energy efficient capital technology requires less energy to perform the same function as non-energy efficient capital technology. For example, an energy efficient light bulb (i.e., LED) uses less energy to produce the same amount of light as a traditional, compact fluorescent light (CFL) light bulb.

interval $[0, \gamma_{max}]$. Here γ_{max} is used to represent the maximum level of energy efficiency obtainable for the stock of capital technology available for purchase on the market by the household during the time period under consideration (Sanstad 2011). [69]

For the time being, we treat the household's primary residence and the stock of capital technology it operates, $K,$ as its fixed factors of production, while the corresponding fuel input used/consumed, E, is treated as a variable input. In this specific case the efficiency level/rating of capital technology employed by the household is assumed to be fixed such that, $\gamma = \bar{\gamma}$. Later in this paper we relax this assumption when we specifically examine how the decision to adopt more energy efficient capital technology is influenced by the energy security status of the household. In this section, however, we discuss just the household's decision-making processes as it pertains to energy services, assuming the household operates as both a producing and consuming unit of energy services.

Following the theory of household production, we assume households obtain utility from the consumption of only two commodities, the energy-related services discussed previously, and a composite commodity, AOS used to represent all other non-energy related goods and/or services from which the household derives utility (Becker 1981; Li 2011). The household's utility function can be generically summarized by a strictly concave utility function of the form

(3) $$U = U(AOS, ESERV(K(\bar{\gamma}), E); HC, DC),$$

where the term HC represents a vector of structural and spatial household characteristics such as square footage, number of bedrooms, number of bathrooms, local weather, and geographical

[69] We allow the interval to be continuous to reflect the continuous changes in the efficiency ratings of household appliances. For example, because technology changes from year to year, an appliance purchased by a household five years ago is likely to be less efficient, in terms of the amount of energy it consumes, than an appliance purchased six months ago.

location (i.e., region) while the term DC represents a vector of household socio-demographic characteristics such as the number and age of people living in the household that could potentially influence a household's demand for energy services, as well as its preferences for other non-energy related services (Li 2011).

In the present application of this basic framework, the objective of each household i is to maximize its utility, as represented by equation (1) subject to the following household budget constraint,

(4) $$M = P_S * ESERV + AOS.$$

were M is used to represent money income and P_S is used to represent the price of providing energy services. All other goods and services (AOS) is assumed to be numeraire and therefore its price is normalized to 1 (Filippini 1999; Boogen et al. 2014). The solution to the utility maximization problem above can be examined as a two-stage optimization problem (Feleke, Kilmer, and Gladwin 2005; Filippini 1999; Deaton and Muellbauer 1980, Muellbaur 1974). [70]

In the first stage, the household acts as a firm whose objective is to minimize the cost of producing household energy services ($ESERV$). To produce household energy services, it is assumed that the household faces two separate prices. The first being the price paid per unit of fuel consumed (e.g., price per kWh of electricity), which is also known as the cost to operate their capital technology inputs K. We label this price simply as P_E and assume it is constant over a given time period (Thompson 2002). The second price households pay to produce household energy services is the price paid to purchase their capital technology inputs (K) from the market.

[70] It is important to note, within our theoretical framework, we assume that the household possesses perfect information on these goods, services, and prices, and solves the choice problem under conditions of certainty.

We label the purchase price of capital technology inputs K as P_K. Therefore, the full price of producing household energy services is equal to the following function:

(5) $[P_E + P_K]$.

Following Sanstad (2011), we assume more efficient capital technology equipment is initially more expensive to purchase but less expensive to operate. Therefore, the full price of producing energy services initially increases if the household makes the decision to purchase new capital technology inputs with corresponding efficiency level $\gamma > \bar{\gamma}$, but decreases over time as the household consumes less fuel to produce energy services as a result of the decision to make an energy efficiency upgrade (Sanstad 2011).

Using the above information, we can write the first stage optimization problem of the household who operates capital technology inputs with corresponding efficiency level $\bar{\gamma}$ as follows:

(6) $Min\ C = (P_E * E) + (P_K * K(\bar{\gamma}))$,

subject to the following production function

(7) $ESERV = ESERVE[K(\bar{\gamma}), E]$.

The Lagrangian for this problem is given by,

(8) $\min\limits_{E,K,\phi} \Psi = (P_E * E) + (P_K * K(\bar{\gamma})) + \phi(ESERV - ESERVE[K(\bar{\gamma}), E])$.

The first order conditions (F.O.C) from the cost cost-minimization problem above assuming an interior solution are as follows:

(8a.) $E\colon\ P_E - \phi \frac{\partial ESERV}{\partial E} = 0$

(8b.) $K\colon\ P_K - \phi \frac{\partial ESERV}{\partial K} = 0$

(8c.) $\phi\colon\ ESERV^* - ESERVE[K(\bar{\gamma}), E] = 0$, which implies the following optimality conditions:

(8d.) $P_E = \phi \frac{\partial ESERV}{\partial E}$

(8e.) $P_K = \phi \frac{\partial ESERV}{\partial K}$

(8f.) $ESERV^* = ESERVE[K(\bar{\gamma}), E]$.

Using the optimality conditions above, we can solve for the following derived conditional input demand functions for capital technology inputs K and energy resource inputs E as follows:

(9) $ESERV = ESERV[K, E]$,

(10) $K(\bar{\gamma})^* = K(P_E, P_K, ESERV)$

(11) $E^* = E(P_E, P_K, ESERV)$.

Using these conditional input demand functions as solutions to the household's first stage optimization problem generates the following minimum cost function,

(12) $C^* = C(P_E, P_K, ESERV)$,

which is assumed to be homogenous of degree one in prices, increasing in $ESERV$ and non-decreasing and concave in prices (Varian 1992). [71]

Because we assume the household is both a producer and consumer of energy services and the least cost combination of inputs necessary to produce an adequate level of energy services has been identified, what remains is for the household to choose the combination of services (both energy-related and non-energy related) that maximize its utility. The choice of services is represented below as the household's second-stage optimization problem. In the second stage of the optimization problem the household solves the following utility maximization problem,

---

[71] To recover the derived input demand functions for energy and capital inputs we could also apply Shephard's Lemma to the cost function.

(13) $\max U = U(ESERV(K(\bar{\gamma}), E), AOS; HC, DC)$

subject to the household budget constraint

(14) $M = AOS + C(P_E, P_K, ESERV)$.

In (14), we have simply replaced the price of producing household energy services with the associated minimum cost function. The corresponding Lagrangian function for the household's second stage optimization problem is given by:

(15) $\max_{\substack{ESERV, \\ AOS, \lambda}} \mathcal{L} = U(ESERV(K(\bar{\gamma}), E), AOS; HC, DC) + \lambda(M - C(P_E, P_K, ESERV) - AOS)$,

where, as stated before the price of all other goods and services (AOS) is normalized to unity. Assuming an interior solution, the F.O.C. can be written as follows: [72]

(15a.) $ESERV\colon\ U_{ESERV} - \lambda \frac{\partial C(P_E, P_K, ESERV)}{\partial ESERV} = 0$

(15b.) $AOS\colon U_{AOS} - \lambda = 0$

(15c.) $\lambda\colon M - C(P_E, P_K, ESERV) - AOS = 0$,

which imply the following optimality conditions:

(15d.) $U_{ESERV} = \lambda \frac{\partial C(P_E, P_K, ESERV)}{\partial ESERV}$

(15e.) $U_{AOS} = \lambda$

(15f.) $M = C(P_E, P_K, ESERV) + AOS$.

Here U_{ESERV} is the marginal utility received from consuming household energy services with $U_{ESERV} > 0$, $\frac{\partial U_{ESERV}}{\partial ESERV} < 0$; U_{AOS} is the marginal utility received from consuming all other services with $U_{AOS} > 0$; $\frac{\partial U_{AOS}}{\partial AO} < 0$; and λ is the marginal utility of income. Additionally, it is assumed that

[72] A corner solution makes little sense in this context, because it would indicate households do not produce any energy services or produce nothing but energy services and prefer to consume nothing else (Thompson 2002).

utility function in (13) and (15) exhibits diminishing marginal rates of substitution of energy services for all other services.

Rearranging optimality conditions (15d. through 15f.) we can obtain the following demand functions for household energy services and the composite commodity used to represent all other services as follows,

(16) $ESERV^* = ESERV(P_E, P_K, M, DC, HC)$

(17) $AOS^* = AOS(P_E, P_K, M, DC, HC)$

The household's conditional input demand functions for capital technology and energy resource can be found substituting equation (16) into equations (9) and (10) as follows,

(18) $E^* = E[P_E, P_K, ESERV^*(P_E, P_K, M, DC, HC)]$

(19) $K(\bar{\gamma})^* = K[P_E, P_K, ESERV^*(P_E, P_K, M, DC, HC)]$,

or equivalently

(20) $E^* = E[P_E, P_K, \ M, DC, HC]$

(21) $K(\bar{\gamma})^* = K[P_E, P_K, M, DC, HC]$.

Equations (20) and (21) along with the equation in (16) reflect the equilibrium consumption amounts for each household i given its current level of fixed capital technology with corresponding efficiency rating $\bar{\gamma}$. The solution to the households two-state optimization problem, assuming all other services $\widetilde{AOS}$ remain constant is illustrated graphically below by Figure 4.1.

In Figure 4.1, the solution to the household's two-stage optimization problem above is represented as A^*, where the highest attainable indifference curve, represented by $U^* = U(ESERV^*, \widetilde{AOS}; HC, DC)$ is tangent with the household's production possibilities frontier (PPF) (i.e., the largest combination of energy services $ESERV$ and all other services AOS that can be produced by the household given the current budget and resource constraints it faces. The minimum cost function $C^* = C(P_E, P_K, ESERV)$ is imbedded within the household's PPF in Figure 4.1 (Deaton and Muellbaur 1980).

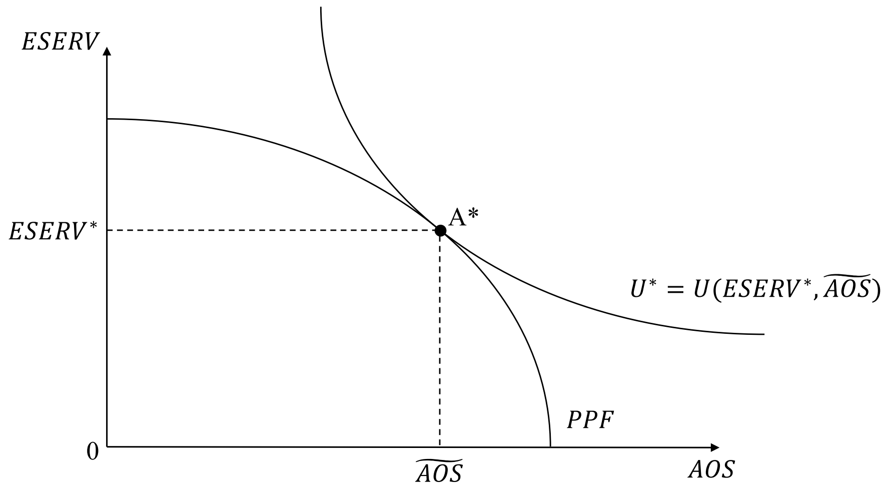


Figure 4.1 Solution to the Household's Two-Stage Optimization Problem

The solution to the model above is static in that it assumes an instantaneous adjustment to a new equilibrium will occur if the prices faced by the household or the household's income change. If the price of energy/fuel inputs P_E changes, more specifically if it increases, then we can expect two types of responses from the household. First, in the short-run the household might choose to lower the rate with which it utilizes its current stock of capital technology with a given efficiency level $\bar{\gamma}$ (Filippini 1999). For example, the household could decide to wash fewer loads of laundry and/or lower/raise the temperature of their thermostat. In the long-run, however, since changes in P_E can result in changes in the relative prices of inputs, the household might choose to alter the input mix it utilizes all together according to equations (16) and (17) (Filippini 1999).

According to our model and (Filippini 1999), when faced with higher energy/fuel prices (P_E) we expect households will replace their existing capital technology with more energy efficient capital technology (i.e., appliances with higher ratings/values of efficiency, γ). In addition to higher inputs prices, a household may choose to adopt more energy efficient capital technology for other reasons including environmental awareness, an unanticipated broken appliance, and/or taking advantage of discounted sale prices.

In this study, we focus on cases where the household may be technically inefficient in its production of household energy services resulting in feelings of energy insecurity. For example, if the household's primary residence is older, it may be more prone to experiencing air leaks, resulting in more fuel being consumed than is actually necessary to produce a given level of energy services. One remedial measure may be to increase the energy efficiency of its current capital technology.

By adopting more energy more energy efficient capital technology (i.e., an increase in the value of γ), a household can increase the amount of energy service outputs produced per unit of energy input consumed. As a result, in the case of technically inefficient production, the level of energy services produced will get closer to the optimal level of energy services, which in turn will allow the household to reach a higher level of utility.

INEFFICIENT PRODUCTION OF HOUSEHOLD ENERGY SERVICES

From an economics perspective, a firm's production process is considered to be technically efficient if there is no additional amount of output feasible that would require the use of fewer inputs (i.e. input-oriented efficiency), or if the firm is producing the maximum amount of output possible given its fixed set of inputs (i.e., output-oriented efficiency). In the case of the household, who is assumed to be a producer of energy services, the production of energy services is considered to be technically efficient if the household is able to produce the maximum possible output of energy services achievable given its inputs. In this situation, the household is said to be operating along its production possibilities frontier (PPF) for energy services.[73]

Conversely, if the actual output of energy services produced by the household falls short of the maximum possible output of energy services achievable, then the producer (i.e., the household) is said to be operating below its production possibilities frontier for energy services and therefore, is experiencing technical inefficiency in production (production inefficiency). To account for the inefficiency in producing energy services by the household, we modify the production function in (1) using a stochastic production frontier approach, proposed originally but independently in the literature by Aigner, Lovell, and Schmidt (1977) and Meeusen and Van den Broeck (1977).

To specify the stochastic production frontier, we first substitute (2) into (1) to obtain,

(22) $ESERV^* = ESERV[K(\gamma), E]$.

The stochastic production frontier is then specified by which we can model the production of energy services by each household i as,

(23) $ESERV_i = ESERV[\ (K(\gamma), E)\ ] - \eta_i$,

where the error term η_i represents "inefficiency" in the production of household energy services. Graphically, production plans that are considered to be technically inefficient in production are those located below the frontier represented by $ESERV^*$. Figure 4.2 below provides an example of production inefficiency.

---

73 The production possibilities frontier (PPF) is used to represents the maximum possible output achievable for a given firm given its fixed set of inputs and the production technology it has available. If one is operating along the PPF, then it is assumed all available resources are being fully utilized and operated efficiently (Bergstrom and Randall, Chapter 5). In our case $ESERV^*$ is used to represent the PPF for energy services.

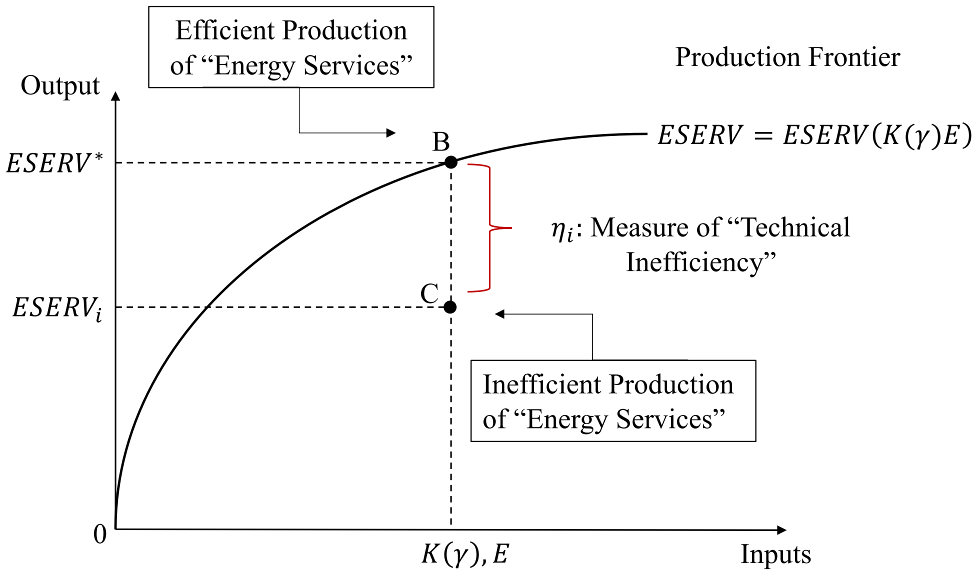


Figure 4.2 Production Inefficiency Household Energy Services

In Figure 4.2, the extent of production inefficiency experienced by each household can be measured by the distance between point *B*, the efficient production point, and point *C*, the inefficient production point.

A limitation in (23) of using only η_i to represent production inefficiency is that it ignores unforeseen or uncontrollable factors that can and do affect the production of energy services within a household. An example of a random factor might include a severe weather event, such as a hurricane, tornado, thunderstorm, or ice/snow storm that prevents access to fuel inputs, hindering the production of cooked meals, indoor heating or cooling, and/or indoor lighting by the household. To accommodate such random factors/occurrences, we decompose the error term η_i into two parts such that,

(24) $\eta_i = v_i - \mu_i,$

where the term v_i is used to capture random factors outside the control of the producer (household) that might impact its ability to produce energy services, while μ_i captures the inefficiency or the shortfall from maximal output dictated by the household's production function. [74] One can think of μ_i as factors within the producers (households) control that prevent it from operating along its production frontier for household energy services.

We can now specify the production of energy services by members of the household who operate capital technology inputs K with corresponding efficiency rating $\bar{\gamma}$ as the following stochastic relationship:

(25) $$ESERV^0 = ESERV[K(\bar{\gamma}), E] - \eta_i = ESERV[K(\bar{\gamma}), E] + v_i - \mu_i$$

where $ESERV^0$ represents the actual (current) output of energy services produced by the household; $ESERV[K(\bar{\gamma}), E]$ represents the deterministic production frontier for household energy services, v_i represents the white noise random error component which is assumed to independently and identically distributed; and μ_i is a one-sided error term representing inefficiency in the production of household energy services resulting from a household's failure to be technically efficient in the way it utilizes its inputs.. [75]

To examine how the production of household energy services can be influenced by increasing the energy efficiency level of the household's capital stock, we begin by assuming the household operates as technically inefficient in its production of energy services. We assume further that the household is a price-taker and is allocating its other inputs (i.e., those not used in production of household energy services) efficiently. [76] If the household is found to be inefficient

[74] It is important to note, due to randomness, not every household can produce the maximum possible output of energy services achievable at all times, even if inputs are same across households (Parmeter and Kumbhakar 2014). The inclusion of the error term v_i allows us to account for situations where operation along the production possibilities frontier for energy services by at least some households is not possible.

[75] Given that u_i leads directly to a shortfall in output, it only reduces output and as such it is assumed to stems from a one-sided distribution

[76] A household is said to be allocating its inputs efficiently if it is deploying or utilizing its resource inputs in the most efficient manner, considering its own preferences and the respective costs of the inputs. In this study, we assume households are efficient in their allocation of inputs.

in its production of household energy services, it is said to be producing energy services at above minimum cost.

Under the stochastic production function framework outlined above, the first-stage cost-minimization problem of the household now becomes,

(26) $Min\ C = (P_E * E) + (P_K * K(\bar{\gamma}))$

subject to the following production function

(27) $ESERV^0 = ESERV[K(\bar{\gamma}), E] - \eta_i = ESERV[K(\bar{\gamma}), E] + v_i - \mu_i.$

In equation (27), $ESERV^0$ is used to represent the energy services produced by the household, assuming it is operating as technically inefficient. The result of the first stage optimization problem above produces the following cost function,

(28) $C^0 = C[P_E, P_K, (ESERV[K(\bar{\gamma}), E] + v_i - \mu_i)].$

Under the assumption of technical inefficiency in the production of household energy services, the second stage optimization problem of the household becomes:

(29) $\max U = U[(ESERV[K(\bar{\gamma}), E] + v_i - \mu_i), AOS; HC, DC],$

subject to the following household budget constraint

(30) $M = AOS + C[P_E, P_K, (ESERV[K(\bar{\gamma}), E] + v_i - \mu_i)].$

In (30), as before, we have simply replaced the price of producing household energy services with the cost function associated with producing household energy services.

Under the assumption of technical inefficiency in production, the solution to the household's second stage optimization problem, represented by equations (29) and (30), produces the following demand functions for household energy services and the composite commodity used to represent all other services as follows,

(31) $ESERV^0 = ESERV(P_E, P_K, M, DC, HC)$

(32) $AOS^0 = AOS(P_E, P_K, M, DC, HC).$

The solution to the above problem is illustrated graphically below in Figure 4.3.

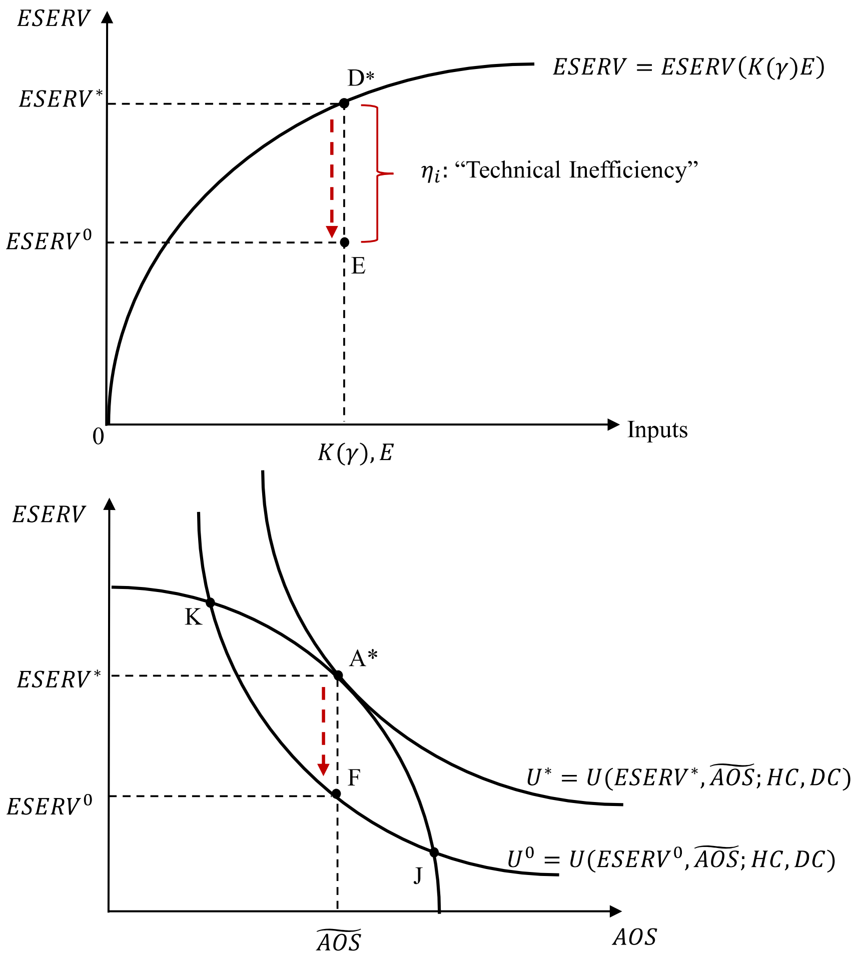


Figure 4.3 Solution to the Household's Problem Under Technical Inefficiency

In Figure 4.3 above, the household is assumed to be technically inefficient in its production of household energy services.[77] As a result, it is not able to operate along its stochastic production

[77] At point A* in Figure 4.3, the household is achieving both technical efficiency and allocative efficiency in production. All points along the PPF are considered to be technically efficient. However, just because a household (firm) is operating along its production possibility frontier for services does not mean the household (firm) is operating as allocativley efficient. Allocative efficiency refers to a state where inputs have been allocated in such a way that they represent the "best" combination of inputs to produce a desired output. The "best" combination of inputs is the least-cost combination of inputs that can be used to produce a desired output. If the household were operating at a point such as K, then the household would be technically efficient in its production of services but not allocativley efficient. At point K the household is producing too many energy services and not enough of all other services. A similar issue occurs at point J where the household is producing too many of all other services and not enough energy services. At

function or frontier and achieve the level of energy services associated with point $D*$. [78] Instead the household is said to be operating below the production frontier; for example, at point E in Figure 4.3. The household is also operating inside of its production possibilities frontier (PPF); for example, point "F" in Figure 4.3. For simplicity, in Figure 4.3 and all other figures used subsequently, we allow the composite error term η_i to represent the total technical "inefficiency" in the production of household energy services.[79]

Assuming a household is a price-taker, technical inefficiency t in its production of energy services implies the household is producing energy services at above minimum cost; that is, $C^0 > C^*$. As before, after the cost function is determined (even an inefficient one), what remains is for the household to choose the combination of services (both energy-related and non-energy related) that maximize its utility. This is represented by the solution to the second stage optimization problem of the household where in Figure 4.3, the household is able to achieve utility level U^0at point F.

Because of the technical inefficiency in production, the household is unable to consume the optimal amount of energy services. Therefore, in addition to being unable to operate along its PPF, the household is unable to reach its optimal, constrained level of satisfaction (i.e., the level of utility it would have achieved in a technically efficient situation). As a result, the utility level achieved by the household in the case of technical inefficiency, U^0 is lower in value (in absolute

both point J and K the household is said to be allocativley inefficient in production. Households who are allocativley inefficient are not producing services at minimum cost (Kumbhakar, Wang, and Horncastle 2015).

[78] The point $D*$ is used to represent the level of energy services *ESERV** in equation (15). It is the amount of energy service outputs produced by the household under the case of technical efficiency.

[79] See Figure 4.1 B. in the Chapter 4 Appendix for an explanation of why an increase in the efficiency rating of the capital technology does not represent a movement in the PPF for the household.

terms) than the utility the household would have been able to be achieve under the case of full efficiency represented by U^* in Figure 4.3.

In other words, given its fixed inputs, including the fixed energy efficiency rating of the capital technology it operates ($\bar{\gamma}$), under the case of technical inefficiency the household is unable to achieve its optimal level of utility. Therefore, the household is unable to experience the sense of energy security it would like to be able to in an unconstrained situation (e.g., in a rental apartment where the landlord has little incentive to update appliances to more energy efficient models). As a result, the household feels “energy insecure.”

DETERMINING THE EXTENT OF TECHNICAL INEFFICIENCY

To measure the extent of technical inefficiency in production (considering both controllable and uncontrollable factors), we follow suggestions from Battese and Coelli (1995) who proposed that the technical inefficiency of a specific production unit (e.g., firm or household) at any given point in time can be estimated as follows:

(33) $\eta_i = \sum y_k \Gamma_k,$

where Γ_k represents parameters to be estimated and y_k represents a vector of observable factors which are thought to influence the level of technical inefficiency experienced by the firm or household. For our analysis, the observable factors in y_k include specific information about the household and its members, including the number of children living inside the household, the household’s income and overall level of education, the geographic location of household, the year the household was built, the household demographics, as well as the type and efficiency rating of the capital technology operated by the household.

In the estimated version of equation (33), a statistically insignificant parameter estimate ($\hat{\Gamma}_k$) implies that the observable factor associated with that parameter estimate does not influence the level of technical inefficiency experienced by the household. Based on the literature, households with more children, households which were built more than 50 years ago, and non-white households are expected to experience higher levels of technical inefficiency (i.e., higher level of energy insecurity) (Drehol and Ross 2016). For this study, we are specifically interested in how the efficiency rating of the capital technology the household chooses to operate influences the level of technical inefficiency it experiences.

THE EFFECTS OF ENERGY EFFICIENCY

Energy efficiency improvements, by design, reduce the amount of energy/fuel input required to produce household energy services. Therefore, under the framework specified above, one way a household could become more efficient in its production of household energy services (i.e., produce more energy services given its fixed set of inputs) is by increasing the energy efficiency level of the capital technology it operates, assuming other inputs are held constant. For example, a household could become more efficient in its production of energy services by purchasing an *Energy Star*® certified appliance, or making improvements such as installing energy efficient windows, sealing air leaks, or adding insulation.

Making such investments is expected to reduce the amount of energy input consumed per unit of energy service output produced by the household, resulting in more efficient production of household energy services. To examine this impact theoretically, assume a household is technically inefficient in its production of household energy services as described earlier. Under the assumption of technically inefficient production, the household is said to be operating below its technically efficient production function or frontier at point such as E in the top panel of Figure 4.3, and therefore is only able to achieve a utility level of U^0 in the bottom panel of Figure 4.3.

Assume the current capital technology inputs operated by the household have a corresponding energy efficiency level/rating of $\bar{\gamma}$. Furthermore, assume we are now operating in the long-run. In the long-run, the household's residence, as well as the stock of capital technology inputs it chooses to operate (K) and the corresponding efficiency level of the capital technology inputs (γ) chosen are free to vary. That is, the household no longer faces fixed capital inputs (K) and fixed technical efficiency of K $(\bar{\gamma})$.

For simplicity assume the household can choose from only one alternative $K(\gamma^*)$ such that $\gamma^* > \bar{\gamma}$. Both $\gamma^*$and $\bar{\gamma}$ are still assumed to exist within the interval $[0, \gamma_{max}]$ such that values of γ closer to γ_{max} are considered to be more "efficient" than values of γ farther from γ_{max}.[80] Because $\gamma^* > \bar{\gamma}$, it is inferred that γ^* is closer to γ_{max} and therefore, the energy efficiency rating/level of the alternative choice is higher than the energy efficiency rating/level of the capital technology employed currently by the household.

Recall that an increase in the efficiency rating of a household's capital technology inputs (i.e., an increase in the value of γ) leads to an increase in the amount of energy service outputs produced per unit of fuel/energy input consumed by the household. Therefore, by operating capital technology inputs K with a corresponding efficiency level of γ^*, (i.e., $K(\gamma^*)$) the household is able to produce more energy services while consuming fewer units of fuel/energy than when it operated as technically inefficient while using capital technology inputs K with a corresponding efficiency level of $\bar{\gamma}$.

Therefore, households who adopt more energy efficient capital technology will become more technically efficient in their production of household energy services, and as a result operate

[80] Here, as before, γ_{max} is used to represent the maximum level of energy efficiency obtainable for the stock of capital technology available for purchase by the household.

closer to their production frontier for energy services (e.g., at point H in the top panel of Figure 4.3) and in some cases actually along their technically efficient production frontier for energy services. [81] As a result, the household would also be able to produce energy services at costs lower as compared to C^0.

The household's production function for household energy services under the assumption of technical inefficiency, but with more efficient capital technology inputs can be written as follows,

(34) $ESERV^H = ESERV[K(\gamma^*), E] - \eta_H.$

Here $ESERV^H$ represents the amount of energy services the household is able to produce if it is technically inefficient in its production of household energy services but has chosen to adopt more energy efficient capital technology inputs, such that $K(\gamma) = K(\gamma^*)$.

The extent of technical inefficiency experienced is labeled as η_H, which is assumed to be less than (in absolute value) the extent of technical inefficiency experienced when the household chose to operate capital technology inputs with corresponding efficiency level/rating $\bar{\gamma}$. We label the extent of technical inefficiency experienced by the household, when it operates capital technology inputs with corresponding efficiency level/rating $\bar{\gamma}$ as η_o. It is important to note, that while technical inefficiency has decreased as a result of the household's decision to invest in energy efficiency, it has not been completely eliminated altogether.

Technical inefficiency in production can only be completely eliminated by investing in energy efficiency if the only source of technical inefficacy is the efficiency rating of the capital

[81] In order for the efficiency improvement from adopting more energy efficient capital technology to result in a case where the household is able again operate along its production frontier, capital technology inputs could be the only contributor to technical inefficiency.

technology inputs the household chose to operate *a prior*. Under this case, the extent of technical inefficiency in production must be exactly equal to the efficiency gains from investing in more energy efficient capital technology inputs. For this study, we consider the case where a gain in efficiency from the adoption of more energy efficient capital technology inputs decreases the extent of technical inefficiency in production, as well as the case where a gain in efficiency from the adoption of more energy efficient capital technology inputs eliminates technical inefficiency in production. The overall increase in efficiency resulting from the household's decision to adopt energy efficient capital technology inputs (i.e., and energy efficient appliance) is represented below in Figure 4.4.

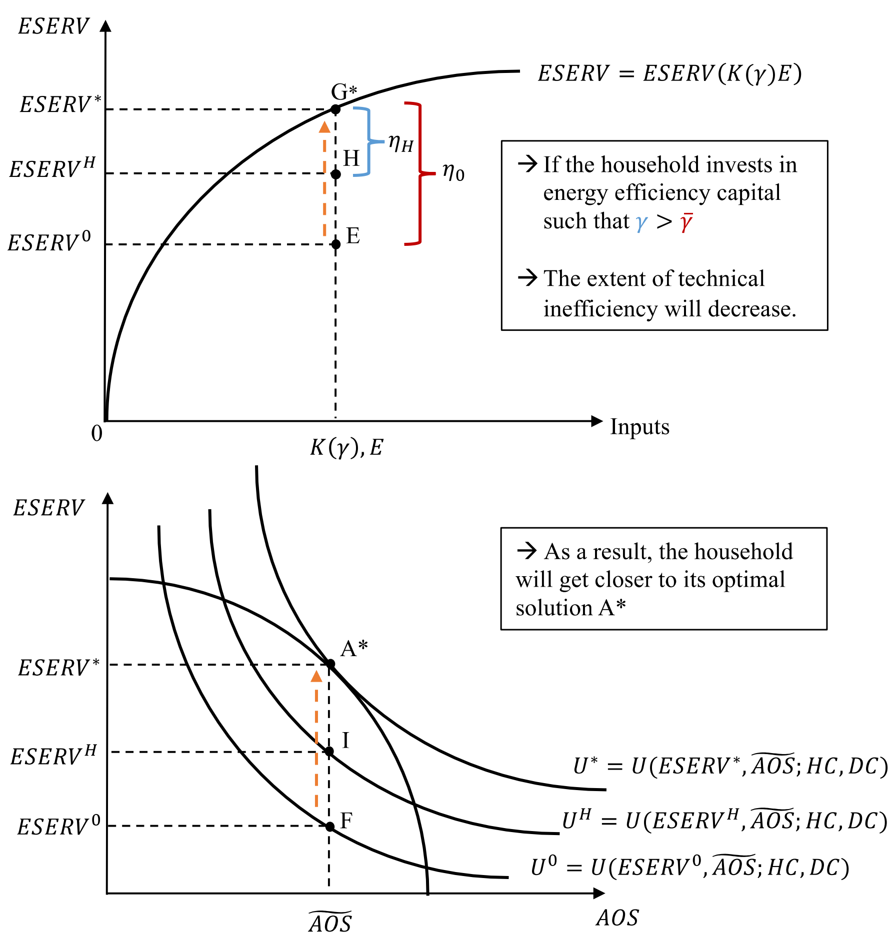


Figure 4.4 Energy Efficiency's Impact on the Technically Inefficient Production

In Figure 4.4, a move from point E to point H represents an increase in the amount of energy service outputs a household would be able to produce in the case where it has chosen to adopt more energy efficient capital technology inputs (i.e., capital technology inputs with corresponding efficiency level [γ^*]) but the household is still considered to be technically inefficient in its production of household energy services. One can interpret point H as a case where the efficiency improvement from purchasing a more energy efficient appliance or making an energy efficiency upgrade has a significant negative effect on the level of technical inefficiency being experienced by the household, but unfortunately is not enough to completely eliminate it. A move from point E to point H in the top panel of Figure 4.4 corresponds to move from point F to point I in the lower panel of Figure 4.4 where the household is operating inside of its PPF.

A move from point H (or point E) point G* in the top panel of Figure 4.4 is used to represent the case where the technical inefficiency experienced by the household has been completely eliminated by its choice to adopt an energy efficient appliance or make an energy efficiency upgrade. Thus, point G* represents an output of energy services that is technically efficient. A move from point H (or point E) Point G* in the top panel of Figure 4.4 corresponds to a move from point I (or point F) to point A* in the bottom panel of Figure 4.4. As illustrated in the bottom panel of Figure 4.4, households who produce energy services more efficiently as the result of adopting more energy efficient capital technology end up on a higher indifference curve (i.e., a higher level of utility) than households who do not, assuming all other inputs remain fixed.

In Figure 4.4, the optimal, but constrained, utility level associated with point G* is U^*, the utility level associated point H is U^H, and the utility associated with point E is U^0. Points F and E in Figure 4.4 are used to illustrate the solution to the household's problem as described under the assumption of technical inefficiency in production from choosing to operate capital technology inputs with corresponding efficiency level/rating $\bar{\gamma}$. Points *H* and *I* are used to illustrate the solution to the household's problem, assuming the extent of technical inefficiency has been reduced as a result of the decisions household to make an energy efficiency investment/upgrade. Points *G** and *A** represent the case where technical efficiency has been eliminated and the household operating along its stochastic production frontier and PPF.

HOUSEHOLD ENERGY SECURITY

In economics, it is difficult to measure the amount of utility an individual receives from consuming goods and services, because as a concept utility is inherently subjective. The term "utility" itself, is used to represent the level of *satisfaction* one receives from consuming goods and/or services. In this study, we assume the utility one receives from consuming energy services can be interpreted as the feeling of being energy secure, such that higher levels of security imply more utility is being received from the household's production and consumption of energy services.

In its most basic format, the term "security" is used to refer to an emotional state individuals experience when they have or believe they have sufficient resources to fulfill their needs (Bericat 2014). As stated earlier, a state of being "energy secure" occurs when a household and its members feel they have adequate access to sufficient, safe, and affordable energy inputs that meet each household members' most basic daily energy service needs. That is, the household production of energy services is meeting the household's minimum "felt needs" with respect to energy services. Within our theoretical framework discussed above, the necessary and sufficient conditions for a household to feel energy secure are as follows.

The first necessary condition is that households are achieving technical efficiency in the production of household energy services; that is, households are operating along their production function or frontier and PPF for energy services. The second necessary condition is that they are meeting allocative efficiency by achieving the *optimal* or highest level of utility possible given the constraints they face. The highest level of utility possible is represented by indifference curve U^* in Figure 4.4. Given the constraints they face, households are able achieve utility level U^* at point A^* in Figure 4.4. At point A* the indifference curve U^* is tangent with the household's PPF.

Because being energy secure is a subjective feeling, a sufficient condition for household to be energy secure is that the optimal level of household utility derived from the consumption of energy services from an efficiency standpoint (U^* in Figures 4.4) is greater than or equal to the minimum level of utility generated by the consumption of energy services which makes a household feel energy secure by meeting their most basic daily energy service needs. Let this minimum this level of utility be denoted by U' which is generated by a corresponding consumption level of energy services, $ESERV'$. If $U^* \geq U'$ and the household is operating along its PPF, then we consider this a sufficient condition for household energy security.

An implication of these necessary and sufficient conditions is that the likelihood of a household feeling energy secure (insecure) increases (decreases) as the household is able to achieve higher levels of utility (as represented by the different indifference curves in Figure 4.4) from its production and consumption of energy services. It could be the case that the constrained optimal level of energy services $ESERV^*$ is equal to the level of energy services which just meet a household's felt needs $ESERV'$. For simplicity, in this study we assume this is the case. Therefore, achieving utility level U^*results in feelings of being energy secure.

4.4 METHODOLOGICAL APPROACH

This section provides an overview of our methodological approach, including the data we use for our analysis, estimation procedures we use to test our hypothesis, and how we address endogeneity.

DATA

Data for our analysis comes from the 2015 Residential Energy Consumption Survey (RECS). As mentioned in Chapter 3 (Essay 2), the RECS is a national multi-phase survey administered by the U.S. Energy Information Administration (EIA) about once every three years. The RECS solicits information on energy consumption, expenditures, and use patterns from households across the United States. [82] Data for the 2015 RECS, in particular, was collected over three separate phases: the household survey phase, the energy use collection phase, and the end-use consumption and expenditure estimates phase. It was administered using a combination of in-person computer-assisted personal interviews, mailed paper surveys, and web-based questionnaires.

A total of 5,686 responses were collected from the 2015 RECS, indicating a response rate of about 44%. [83] In this study, we focus only on single-family detached homes.[84] We specifically use the 2015 RECS survey because it asks participants to reveal information about the energy use patterns within their home. [85] Additionally, the 2015 RECS collects information about the physical characteristics of each individual's home including the age of the primary dwelling, the number of bedrooms, the number of bathrooms, the number of square feet, the number of and type of appliances used, as well as the type of heating and ventilation equipment used regularly by the household.

---

[82] While data collection for the RECS dates back to 1980, each survey administered targets a different set of households. Also, the questions asked vary from year to year. Thus, each set of observations should be treated a s single a cross-section.

[83] To create a general, representative sample of all U.S. households, the EIA used a multistage area probability sample design. This design begins by dividing the United States into different geographical areas by randomly selecting public micro data areas (PUMAs). Each PUMA was then divided into several different census block groups (CBGs), resulting in a total of 800 total CBGs, four per each PUMA. In the third and final stage randomization, households are randomly chosen from an overall list of households in each of the selected CBGs. In most of the CBGs, the list of households is created from the United States Postal Service Delivery Sequence File (DSF).

[84] Single-family detached homes include stand-alone houses. These types of home do not include multi-family residential dwellings such as townhomes, apartments, or duplexes.

[85] These energy use patterns include information such as whether or not the household has received an in-home energy audit within the past year, made any energy efficiency upgrades, or had been the recipient of funding from an energy assistance program.

Lastly, we utilize data from the 2015 RECS because the survey specifically asked households a subset of questions related to any challenges they may have faced over the past twelve months in paying their energy bills or maintaining heating or cooling inside their home. We interpret these questions as indicators a household was unable to produce an adequate level of energy services and therefore is unable to achieve energy security (e.g., the household is unable to reach U^* in Figures 4.1 and 4.4). Household responses to these questions are used to construct the Rasch model-based energy insecurity index in Chapter 3 (Essay 2). Descriptive statistics for 2015 RECS are listed below in Table 4.1

Table 4.1 Summary Statistics 2015 Residential Energy Consumption Survey (RECS)

Variable	**Description**	**Mean**	**St. Error**	**Min**	**Max**
Income	Median Household Income ($)	62,370.74	42,640.72	10,000	140,000
Employed	= 1 if Respondent for Household is Employed	0.59	0.49	0	1
Education	= 1 if Respondent for Household has College Degree	0.69	0.46	0	1
Age	Respondent for Household's Age (Years)	52.30	17.02	18	85
Gender	= 1 if Respondent for Household is Male	0.44	0.50	0	1
Children	Number of Children Aged 16 and Younger	0.61	1.04	0	10
CDD65	Cooling Degree Days	1,719.21	1,193.56	0	6,607
HDD65	Heating Degree Days	3,707.85	2,149.27	0	9,843
House Age	= 1 if Dwelling Built Before 1959 (+60 years old)	0.25	0.43	0	1
Hispanic	= 1 if Respondent for Household is Hispanic	0.13	0.33	0	1
African American	= 1 if Respondent for Household is African American	0.10	0.31	0	1
Audit	= 1 if Household Received an in home Energy Audit	0.08	0.27	0	1
Energy Star	Number of *Energy Star* ® Appliances	2.37	2.23	0	7
Windows	= 1 if Household has Triple/Double Pane Windows	0.61	0.49	0	1
Insulation	= 1 if Household has "Adequate" Insulation	0.83	0.38	0	1
Reduce	= 1 if Reduced or Forgone Expenditures	0.14	0.35	0	1
Unsafe	= 1 if Kept Household at Unsafe Temperature	0.06	0.24	0	1
Notice	= 1 if Household Received a Disconnection Notice	0.06	0.24	0	1
No Fuel	= 1 if Household Couldn't Afford Elec., Nat. Gas, or Propane	0.03	0.16	0	1
HVAC	= 1 if Household's HVAC is Broken and Can't Afford Repair	0.06	0.24	0	1
Medical	= 1 if Household Sought Medical Attention	0.01	0.12	0	1
Days	= 1 if Days without Heat/AC Exceeds 36 days	0.02	0.13	0	1
Dehumidifier	Number of Months Dehumidifier in Use	0.76	2.24	0	12
Free Audit	= 1 if Household Received a Free In-Home Energy Audit	0.02	0.14	0	1
Appliance Rebate	= 1 if Utility Offered an Appliance Rebate	0.04	0.20	0	1
Ownership	= 1 if Household is an Owner	0.69	0.46	0	1
Bedrooms	Number of Bedrooms	2.83	1.11	0	10
Bathrooms	Number of Complete Bathrooms	1.75	0.75	0	6
Observations	**5,686**				

In addition to collecting information on the ability of households to maintain access to energy services, the 2015 RECS also asked households to report information on whether or not they have participated in an "energy program" over the past 12-months or made any energy efficiency upgrades to their home. For example, household respondents were asked to report on whether or not any of the appliances in their home (e.g., refrigerators, freezers, dishwashers, water heaters etc.) are *Energy Star®* certified.

The survey also asks households whether or not they have received an in-home energy audit sometime over the last 12-months. Additional questions include whether or not the household has made any energy efficiency upgrades over the past 12-months including whether or not they have installed triple or double pane windows or felt they had adequate insulation. We use the information collected in this section of the survey to distinguish between households who have made energy efficiency upgrades and those that have not.

EMPIRICAL MODEL SPECIFICATION

We now present the empirical model we used to examine how making energy efficiency improvements in the home impacts the presence of household energy insecurity. The general formulation is given in terms of the household's true, latent level of energy insecurity,

(35) $EINSECURE_i^* = \sum_{i=1}^{n=26} \beta_k X_i + \varepsilon_i$,

where $EINSECURE_i^*$ is used to represent the true, but unobserved level of energy insecurity for each household i; the β_k's are parameters to be estimated; X_i represents a vector of covariates believed to influence a household's energy insecurity status; and ε_i is a random disturbance term.

Specific covariates in X_i are defined as follows: X_1 is a continuous variable equal to median household income ($); X_2 is a binary variable equal to 1 if the respondent for the household is employed either part-time or full time and 0 otherwise; X_3 is a variable equal to the age of the respondent for the household (years); X_4 is a binary variable equal to 1 if the household has an additional degree beyond a high school diploma and 0 otherwise; X_5 is a continuous variable equal to the number of children aged 16 and under living in the home; X_6 is a variable equal to one if the respondent of the household is male and zero otherwise; X_7 and X_8 are continuous variables corresponding to the average number of cooling degree days (CDD) and heating degree days (HDD), respectively; X_9 is a binary variable equal to 1if the household's primary dwelling was built before 1959 and 0 otherwise; X_{10} is a binary variable equal to 1 if the head of the household identifies as Hispanic/Latino and 0 otherwise; X_{11} is a binary variable equal to 1 if the household identifies as African- American and 0 otherwise; X_{12-20} are indicator variables set equal to 1 for the U.S. Census region where a household is located and 0 otherwise where Census regions are defined respectively as, New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain North, Mountain South, and Pacific; X_{21} is a continuous variable equal to the number of bedrooms the home of the respondent has; X_{22} is a continuous variable equal to the number of complete bathrooms the home of the respondent has; X_{23} is an indicator variable equal to 1 if the household received an in-home energy audit sometime during the past 12-months and 0 otherwise; X_{24} is a binary variable equal to 1 if the household has energy efficient windows (e.g., triple and/or double pane windows) and 0 otherwise; X_{25} is an indicator variable equal to 1 if the household has adequate insulation [86] and 0 otherwise; and X_{26} is continuous variable equal to the number of *Energy Star®* appliances operated by the household. *Energy Star®* appliances considered include: washers, dryers, dishwashers, refrigerators, freezers, light bulbs, and hot water heaters.

[86] Households whose responses to the 2015 RECS indicated that their home was well insulated or adequately insulated were assumed to have an adequate level of insulation, while those who indicated their household was poorly insulated were considered to have inadequate insulation.

We employ two empirical/statistical methods for estimating Equation 35 (explained in more detail below), both of which are based on assuming there exists an underlying random utility function that explains the overall utility a household receives from consuming energy services specified as:

(36) $U^i = h\left(ESERV, \widetilde{AOS}; HC, DC\right) + e_i$,

where U^i is the level of utility received by a household from the production and consumption of energy services and a given level of all other services $\widetilde{AOS}$; $h(\cdot)$ is a non-differentiable function representing the deterministic component of U^i and e_i is a random error term representing stochastic factors affecting U^i. Furthermore, because U^i is assumed to be strictly increasing in its arguments, an increase in energy services leads to a higher level of utility being achieved by the household.

Recall from Section 4.3 the three separate utility levels, U^*, U^H, and U^0 are considered achievable by the household, as illustrated in Figure 4.4. These different utility levels represent the different levels of satisfaction the household receives by having access to energy services in the amounts of $ESERV^*$, $ESERV^H$, and $ESERV^0$, respectively That is,

(37) $U^* = h\left(ESERV^*, \widetilde{AOS}; HC, DC\right) + e_*$

(38) $U^H = h\left(ESERV^H, \widetilde{AOS}; HC, DC\right) + e_H$

(39) $U^0 = hU\left(ESERV^0, \widetilde{AOS}; HC, DC\right) + e_0$,

such that $U^* > U^H > U^0$ and each level of U^i represents a different level of subjective energy security felt by a household holding "all other services" constant.

The difference in utility a household receives from consuming different levels of energy services, holding “all other services” constant, can be modeled as follows for the choice between energy services $ESERV^*$and $ESERV^0$: [87]

(40) $\Delta U = \left[h\left(ESERV^*, \widetilde{AOS}; HC, DC\right) + e_*\right] - \left[h\left(ESERV^0, \widetilde{AOS}; HC, DC\right) + e_0\right]$

or equivalently

(41) $\Delta \mathrm{U} = \left[h\left(ESERV^*, \widetilde{AOS}; HC, DC\right) - h\left(ESERV^0, \widetilde{AOS}; HC, DC\right)\right] + [e_* - e_0]$,

and as follows for the choice between energy services $ESERV^H$ and $ESERV^0$

(42) $\Delta U = \left[h\left(ESERV^H, \widetilde{AOS}; HC, DC\right) + e_H\right] - \left[h\left(ESERV^0, \widetilde{AOS}; HC, DC\right) + e_0\right]$

or equivalently

(43) $\Delta \mathrm{U} = \left[h\left(ESERV^H, \widetilde{AOS}; HC, DC\right) - h\left(ESERV^0, \widetilde{AOS}; HC, DC\right)\right] + [e_H - e_0]$.

The first part of equation (40) $\left[h\left(ESERV^*, \widetilde{AOS}; HC, DC\right) - h\left(ESERV^0, \widetilde{AOS}; HC, DC\right)\right]$ and equation (42) $\left[h\left(ESERV^H, \widetilde{AOS}; HC, DC\right) - h\left(ESERV^0, \widetilde{AOS}; HC, DC\right)\right]$ represents the deterministic components of ΔU, while $[e_* - e_0]$ and $[e_H - e_0]$ represent the random components of ΔU (Hanemann 1984).

The random utility model presented above is useful for predicting consumer choices. For example, when presented with the choice to make an energy efficiency upgrade, a household will make the upgrade if the utility it receives from making the upgrade exceeds the utility it would have received if it did not make the upgrade, assuming all other inputs remain fixed. As discussed in Section 4.3, under the assumption of technical inefficiency (e.g., the household is operating at a point such as “*F*” in Figure 4.4) when household decides to adopt more energy efficient capital

[87] See Figure 4.3 for more clarification.

technology inputs, the household can reduce its technical inefficiency in production and produce more energy services, assuming all other inputs remain fixed.

Under this scenario, the household can operate at a point such as I in Figure 4.4 and therefore is able to reach a higher level of utility. The difference between the optimal level of utility (U* achieved at point $A*$ in Figure 4.4) and the household's current sub-optimal level of utility (U^0 achieved at point F in Figure 4.4 where it is assumed the household is operating as technically inefficient and operating capital technology inputs with corresponding efficiency rating/level $\bar{\gamma}$) will decrease if the household chooses to make and energy efficiency upgrade. If the household chooses to make an energy efficiency upgrade, then it is able to reach utility level U^H, which is illustrated by a move from point "*F*" in Figure 4.4 to point "*I*" in Figure 4.4. The decrease in the difference of utility can be expressed mathematically as follows,

(44) $$U^* - U^H = \left[h\left(ESERV^*, \widetilde{AOS}; HC, DC\right) + e_*\right] - \left[h\left(ESERV^H, \widetilde{AOS}; HC, DC\right) + e_H\right] <$$

$$U^* - U^0 = \left[h\left(ESERV^*, \widetilde{AOS}; HC, DC\right) + e_*\right] - \left[h\left(ESERV^0, \widetilde{AOS}; HC, DC\right) + e_0\right].$$

By adopting more energy efficient capital technology the household is able to produce energy services that are closer to the amount of energy services along their PPF, thereby decreasing ΔU between the optimal utility level and a sub-optimal utility level(s).

In the 2015 RECS, respondents were asked a subset of questions related to challenges they may have faced over the past twelve months maintaining access to household energy services. Questions were worded in the negative (e.g., "In the last year, did anyone in your household need medical attention because your home was too cold/hot?") and therefore, affirmative responses were interpreted as an "inability" to maintain access to energy services. Households are assumed to respond to these questions, according to their latent energy insecurity status, such that the more energy insecure the household, the larger the probability the household will give a positive response.

Thus, related to the utility differences illustrated in Figure 4.4 and equations (41) and (43), we expect that as the ΔU between the optimal level of utility (generated by the optimal level of energy services) and the sub-optimal levels of utility (generated by the sub-optimal levels of energy services) increases, so too will the probability that a household answers "yes" or responds affirmatively to the RECS energy insecurity questions. The reasoning behind this expectation is that as a household moves farther away from the optimum level of utility achievable, it is producing fewer household energy services, and therefore is likely feeling more energy insecure as they are also likely to be moving farther and farther away from the minimum level of utility and energy services necessary to satisfy their most basic daily energy service needs (which would be the case, for example, if we assume $U^* = U'$- see related discussion in Section 4.3).

We do not observe each household's true level of energy insecurity. We only observe their responses to questions included in the 2015 RECS. Therefore, we construct an energy insecurity index, using household responses to the RECS questions and a Dichotomous Rasch model (see Chapter 3 for more information). The energy insecurity index value assigned to each household i ($EISINDEX_i$) depends on the number of "yes" responses by the household to the questions included in the RECS. Households who respond "no" to all questions are considered "Energy Secure," and are assigned an index value $EISINDEX_i = 0$. Households who respond affirmatively to any questions are assumed to be "Energy Insecure" are assigned an energy insecurity index value $EISINDEX_i > 0$. The farther the energy insecurity index value from zero, the greater the extent of energy insecurity being experienced by the household.

Following the random utility framework above, using the Rasch model results our observation for the first index we construct can be modeled as follows,

(45) $EISINDEX_i = 1$ if $EINSECURE_i^* > 0$

and

(46) $EISINDEX_i = 0$ if $EINSECURE_i^* \leq 0$.

Here the energy insecurity index can take on one of two values: $EISINDEX_i = 1$ if the household responds affirmatively to at least one questions of interest from the 2015 RECS, and $EISINDEX_i = 0$ if the household responds affirmatively to none of the questions. If the household responds affirmatively to none of the questions, it seems we can safely assume the household is feeling energy secure at least in the sense that it is able to maintain access to energy services (i.e., meeting its most basic daily energy service needs).

Using the first index, the energy insecurity model in equation (35) can be cast as a binary response model of the form:

(47) $P(EISINDEX = 1|\boldsymbol{x}) = G(\boldsymbol{x\beta}) \equiv p(\boldsymbol{x})$,

where $\boldsymbol{x}$ is $1 \times K$, $\boldsymbol{\beta}$ is $K \times 1$, and the first value of $\boldsymbol{x}$ is equal to unity. The model in equation (47) is referred to as an index function model because it restricts the way in which the response probability depends on $\boldsymbol{x}$ (Wooldridge 2010). The $p(\boldsymbol{x})$ is a function of $\boldsymbol{x}$ only through the index $\boldsymbol{x\beta}$ (Wooldridge 2010).

In most applications, $G(\boldsymbol{x\beta})$ is a cumulative distribution function (cdf) whose specific form is derived from the theoretical framework underlying the economic model used for the analysis (Wooldridge 2010). In our case, the index model we construct to measure a household's energy insecurity status is derived from an underlying random utility model (see equations 37 through 44). The first index $EISINDEX_i$ is modeled as a binary indicator, set equal to one if the household identifies as energy insecure and zero otherwise. Using only the results of the first index, the energy insecurity model in equation (35) can be cast as a qualitative response model of the form,

(48) $\Lambda_i = Prob(EISINDEX_i = 1|X_i) = Prob\left(\sum_{i=1}^{j=k} \beta_j X_i + \varepsilon_i\right)$

where Λ_i is the probability that a household is not energy secure (responds affirmatively to at least one question, and is classified as energy insecure according to the index).

The standard parametric approach used to analyze the data would be a logistic regression. The logistic model of household energy insecurity can be specified as:

(49) $\Lambda_i = Prob(EISINDEX_i = 1|X_i) = \frac{\exp(\hat{\beta}_0+\sum_{i=1}^{n=22} \hat{\beta}_k X_i)}{1+\exp(\hat{\beta}_0+\sum_{i=1}^{n=22} \hat{\beta}_k X_i)}$,

where Λ_i is the conditional probability of the household i being not energy secure; the β_k's are the parameters to be estimated, and X_i is the set of covariates described earlier. The marginal effects for individual variable considered to have impact on household energy insecurity can be calculated as follows:

(50) $\frac{d\Lambda(x_i\beta_k)}{d(x_i\beta_k)} = \frac{\exp(x_i\beta_k)}{[\exp(x_i\beta_k)^2} = \Lambda(x_i\beta_k)[1-\Lambda(x_i\beta_k)]$. [88]

Marginal effects for any categorical variables should be interpreted as a change in the probability that a household is energy insecure as the categorical variable changes from 0 to 1, holding all other variables constant.[89]

Following the random utility framework, the probability a household will identify as energy insecure (receives an energy insecurity index value not equal to zero) increases as the distance between U^* and some other sub-optimal utility level (e.g., U'). Following equations (41)

---

[88] The loglikelihood function for the logistic model is listed in the Chapter 4 Appendix.

[89] The discrete choice modeling procedure described above in the case of a binary dependent variable is similar to discrete choice contingent valuation models where survey respondents are asked to give a "Yes or No" response to a willingness-to-pay question such as, "Would you support a program to increase the size of a local public park by X acres if the cost your household was $Y per year? (Hanemann 1984).

and (43) if the change in utility is greater than zero, then the probability the household will identify as energy insecure (providing at least one "yes" response to the questions in the 2015 RECS) is expected to increase. Following the random utility framework, the probability that a household identifies as energy insecure theoretically depends on changes in the amount of energy services the household is able to produce and consume, the efficiency rating of the capital technology the household chooses to operate, energy/fuel consumption of the household, structural and spatial characteristics of the household, and the household's socio-demographic characteristics.

Table 4.2 lists all of the variables included in our empirical model specification believed to influence the probability that a household identifies as energy insecure. To provide a clear connection to our theoretical framework (e.g., equations 1-34 and related discussion), we also include a label for each variables theoretical counterpart and the hypothesized sign of their respective regression coefficients.

Table 4.2 Empirical Variables used to Measure a Household Energy Insecurity

Empirical Variable	**Label**	**Theoretical Counterparts**	**Expected Sign of Estimated Coefficient**
Median Household Income ($)	Income (X_1)	HC	Negative
= 1 if Respondent for Household is Employed	Employed (X_2)	HC	Negative
Respondent for Household's Age (Years)	Age (X_3)	HC	Positive
= 1 if Respondent for Household has Education Beyond High school	Education (X_4)	HC	Negative
Number of Children Aged 16 and Younger	Children (X_5)	HC	Positive
= 1 if Respondent for Household is Male	Gender (X_6)	HC	Negative
Cooling Degree Days (No. of Days Temperatures are above 65°F)	CDD65 (X_7)	E	Positive
Heating Degree Days (No. of Days Temperatures are below 65°F)	HDD65(X_8)	E	Positive

= 1 if Dwelling Built Before 1959 (+60 years old)	House Age(X_9)	DC	Positive
= 1 if Respondent for Household is Hispanic	Hispanic (X_{10})	HC	Positive
= 1 if Respondent for Household is African American	Af. American (X_{11})	HC	Positive
Census Dummies	Census (X_{12-20})	DC	
	New England		Positive
	Middle Atlantic		Negative
	East North Central		Positive
	West North Central		Indeterminate
	South Atlantic		Indeterminate
	East South Central		Positive
	West South Central		Positive
	Mountain North		Positive
	Mountain South		Negative
	Pacific		Negative
Number of Bedrooms	Bedrooms (X_{21})	DC	Positive
Number of Complete Bathrooms	Bathrooms (X_{22})	DC	Positive
= 1 if Household Received an in home Energy Audit	Audit (X_{23})	$K(\gamma)$	Negative
Number of *Energy Star* ® Appliances	*Energy Star*(X_{24})	$K(\gamma)$	Negative
= 1 if Household has Triple/Double Pane Windows	Windows (X_{25})	$K(\gamma)$	Negative
= 1 if Household has “Adequate” Insulation	Insulation (X_{26})	$K(\gamma)$	Negative

In addition to the first method of partitioning, we also create three additional energy insecurity indices using results from the Dichotomous Rasch model. Rather than dividing households into two groups, all of the additional energy insecurity indices we create, allow us to examine the extent of energy insecurity being experienced by each individual household. The extent of energy insecurity experienced, depends on the difference between the household's placement along an energy insecurity scale and the chosen value of the threshold τ which varies across the different indices.

The additional index measure we create are similar to index measures used in the subjective well-being (SWB) literature. SWB surveys are designed to elicit information on people's feelings, including the sense of satisfaction or "happiness" they feel in their overall life. Thus, in SWB models, the dependent variable is a strictly positive variable, which is assumed to measure utility (i.e., satisfaction, happiness). Following the SWB approach, we interpret the second energy security indices produced from the application of the Dichotomous Rasch model as a self-reported, direct measure of the level of dissatisfaction households feel regarding their energy circumstances based on their subjective feelings about their level of energy insecurity.

For the first additional index, the threshold value of τ is set equal to zero. Households who respond affirmatively to zero questions are placed at the bottom of the energy insecurity scale (see Figure 3.3). These households are considered fully energy secure (i.e., not energy insecure) and therefore, are assigned an energy insecurity index $EISINDEX = 0$. Other households who respond affirmatively to one or more questions are assigned energy insecurity index values equal to the difference between the value of the energy insecurity scale they receive and the threshold value of $\tau = 0$. The most severe condition of energy insecurity, in this case is represented by affirmative responses to all seven questions from 2015 RECS, which corresponds to an energy insecurity index value, $EISINDEX = 6.34$.

In the second additional index, the threshold value of τ is set equal to 3.08, the severity parameter estimate associated with the first question on the energy insecurity scale (Question 2 Reduce). Households who respond affirmatively to either zero or only one question are considered "Energy Secure." As a result, these households receive an energy insecurity index value $EISINDEX_i = 0$. Other households, who respond affirmatively to two or more questions are considered "Energy Insecure." The extent of energy insecurity, just as before, is determined by the difference between the value on the energy insecurity scale where the household has been placed, and the threshold value of $\tau = 3.08$. The most severe condition of energy insecurity, in this case is still represented by affirmative responses to all seven questions. The energy insecurity index value, however is equal to the difference between 6.34 (the energy insecurity scale value associated with responses to seven questions) and the threshold value of $\tau = 3.08$, such that $EISINDEX_i = 4.26$.

Because the severity parameter estimates for some of the questions used to create the energy insecurity scale are similar to one another (e.g., Question 5 [HVAC], Question 2 [Unsafe], and Question 3 [Notice]) for the third additional energy insecurity index we create, the energy insecurity scale is divided into four separate categories. Each category represents a range along the scale. Households who have scores within a certain range are all assigned the energy insecurity index value that is associated with that range.

Based on the severity parameter estimates produced from the Dichotomous Rasch model (see Table 3.25) that are used to create the scale, we consider households who respond affirmatively to either only one or no questions to be "High Energy Secure." These households are assigned an energy insecurity index value = 1 ($EISINDEX_i = 1$). Household who respond affirmatively to two, three, or four questions are considered "Marginally Energy Secure," and are assigned an energy insecurity index value = 2 ($EISINDEX_i = 2$). Households who respond affirmably to five questions are considered to be "Low Energy Secure," and are assigned an energy insecurity index value = 3 ($EISINDEX_i = 3$). Households who respond affirmatively to six or seven questions are considered to be "Very Low Energy Secure." They are assigned an energy insecurity index value = 4 ($EISINDEX_i = 4$).

Each of the additional index measures produced from the Dichotomous Rasch model results, generate a different set of numerical values for each of the levels of energy insecurity able to be experienced along the scale by each individual household i. Based on their responses to the questions of interest and the index chosen, each household is assigned a value within a set. The set of values changes depending on the threshold value of τ chosen. For the first additional index measure we create, where the value of the threshold $\tau = 0$, eight separate energy insecurity index values are identified: $EISINDEX_i \in \{0, 3.08, 4.38, 4.44, 4.48, 5.65, 6.20, 6.34\}$. For the second additional index measure we create, where the value of the threshold $\tau = 3.08$, seven separate energy insecurity index values are identified: $EISINDEX_i \in \{0, 1.30, 1.36, 1.40, 2.57, 3.12, 3.2\}$.

Across both methods of partitioning, 0 is the lowest (least severe) level of energy insecurity able to be experienced by a household. In the first 6.34 is the highest (most severe) while in the second 3.26 is the highest (most severe) level of energy insecurity able to be experienced. It is important to note that the fact that 5.65 (2.57) is worse than 4.48 (1.40) in the sets conveys important information, but nothing is lost if a different set of numbers are used, as long as the ordinal magnitudes between values is consistent with the original set (Wooldridge 2010). For the third additional index measure we create, where the value of the threshold $\tau = 3.08$ and the energy insecurity scale is divided into four groups, four separate energy insecurity index values are identified: $EISINDEX_i \in \{1, 2, 3, 4\}$. Here 1 is the lowest (least severe) level of energy insecurity able to be experienced by a household and 4 is the most severe.

Following the second and third method of partitioning in which responses to none or only one question indicates the household is still "Energy Secure" is consistent with the methods used in food security literature (see Hamilton et al. 1997, Bickel et al. 2000, and Nord 2003). Again, nothing is lost if a different set of numbers is used for the energy insecurity index values, as long as the ordinal magnitude between each value is consistent. [90]

Using these additional energy insecurity indices, the energy insecurity model in equation (35) can be cast as an ordered response model, more specifically an ordered logit model (Greene 2012). For each household i we hypothesize that there is a continuously varying true level of energy insecurity being experienced that underlies the household's response pattern to the seven questions included in Table 3.1 (i.e., questions from Section L on the 2015 RECS). The true level of energy insecurity is labeled in equation (35) as $EINSECURE_i^*$ and is assumed to exist over the interval $-\infty < EINSECURE_i^* < +\infty$.

Although the household's true level of energy insecurity likely varies continuously in the space of individual utility, the experience of being energy insecure in this analysis is provided by a discrete outcome on a scale (Greene 2012). Households receive an energy insecurity index value ($EISINDEX_i$) based on the number of questions to which they respond affirmatively and the chosen value of the threshold. Logically, then, the translation from the household's true underlying energy insecurity status and its position along the energy insecurity scale, which determines the

[90] The food insecurity scale is typically transformed to range in value from 0 to 10. Therefore, the scale values included are linear transformations of the original item calibrations produced for the 18 questions from the Rasch model (Hamilton et al. 1997; Bickel et al. 2000).

energy insecurity index value the household receives, could be viewed as a censoring of the true underlying level of energy insecurity experienced by the household (Greene 2012).

We do not observe a household's true energy insecurity status. What we do observe how their pattern of responses determines their place along the scale and how this placement determines the energy insecurity index value they receive. Under the first additional index (setting the threshold value $\tau = 0$) our observations can be recorded and ordered as follows,

(51a.) $EISINDEX_i = 0$ if $EINSECURE_i^* \leq 0$

(51b.) $EISINDEX_i = 3.08$ if $0 < EINSECURE_i^* \leq \omega_{3.08}$

(51c.) $EISINDEX_i = 4.38$ if $\omega_{3.08} < EINSECURE_i^* \leq \omega_{4.38}$

(51d.) $EISINDEX_i = 4.44$ if $\omega_{4.38} < EINSECURE_i^* \leq \omega_{4.44.}$

(51e.) $EISINDEX_i = 4.48$ if $\omega_{4.44} < EINSECURE_i^* \leq \omega_{4.48}$

(51f.) $EISINDEX_i = 5.65$ if $\omega_{4.48} < EINSECURE_i^* \leq \omega_{5.65}$

(51g.) $EISINDEX_i = 6.20$ if $\omega_{5.65} < EINSECURE_i^* \leq \omega_{6.20}$

(51h.) $EISINDEX_i = 6.34$ if $\omega_{6.20} < EINSECURE_i^*$,

where ω_i for $i = 3.08, 4.38, 4.44, 4.48, 5.65, 6.20$ represent the different thresholds along the underlying continuum of a household's true level of energy insecurity. In this case, the thresholds represent the cut points between each calibration included on the scale. They are parameters to be estimated (Greene 2012).

We can derive the conditional probability of each household i receiving an energy insecurity index value $EISINDEX_i$ as follows:

(52a.) $P(EISINDEX_i = 0|X_i) = P(EINSECURE_i^* \leq 0|X_i)$

(52b.) $P(EISINDEX_i = 3.08|X_i) = P(0 < EINSECURE_i^* \leq \omega_{3.08}|X_i)$

(52c.) $P(EISINDEX_i = 4.38|X_i) = P(\omega_{3.08} < EINSECURE_i^* \leq \omega_{4.38}|X_i)$

(52d.) $P(EISINDEX_i = 4.44|X_i) = P(\omega_{4.38} < EINSECURE_i^* \leq \omega_{4.44}|X_i)$

(52e.) $P(EISINDEX_i = 4.48|X_i) = P(\omega_{4.44} < EINSECURE_i^* \leq \omega_{4.48}|X_i)$

(52f.) $P(EISINDEX_i = 5.65|X_i) = P(\omega_{4.48} < EINSECURE_i^* \leq \omega_{5.65}|X_i)$

(52g.) $P(EISINDEX_i = 6.20|X_i) = P(\omega_{5.65} < EINSECURE_i^* \leq \omega_{6.20}|X_i)$

(52h.) $P(EISINDEX_i = 6.34|X_i) = P(\omega_{6.20} < EINSECURE_i^*|X_i)$

Replacing equation (35) for $EINSECURE_i^*$ and assuming a logistic model specification, the probabilities listed in equations (52a.) through (52h.) can be rewritten as follows

(53a.) $P(EISINDEX_i = 0|X_i) = 1 - \Lambda(\sum_{i=1}^{n=22} \beta_k X_i)$

(53b.) $P(EISINDEX_i = 3.08|X_i) = \Lambda(\omega_{3.08} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(-\sum_{i=1}^{n=22} \beta_k X_i)$

(53c.) $P(EISINDEX_i = 4.38|X_i) = \Lambda(\omega_{4.38} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{3.08} - \sum_{i=1}^{n=22} \beta_k X_i)$

(53d.) $P(EISINDEX_i = 4.44|X_i) = \Lambda(\omega_{4.44} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{4.38} - \sum_{i=1}^{n=22} \beta_k X_i)$

(53e.) $P(EISINDEX_i = 4.48|X_i) = \Lambda(\omega_{4.48} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{4.44} - \sum_{i=1}^{n=22} \beta_k X_i)$

(53f.) $P(EISINDEX_i = 5.65|X_i) = \Lambda(\omega_{5.65} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{4.48} - \sum_{i=1}^{n=22} \beta_k X_i)$

(53g.) $P(EISINDEX_i = 6.20|X_i) = \Lambda(\omega_{6.20} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{5.65} - \sum_{i=1}^{n=22} \beta_k X_i)$

(53h.) $P(EISINDEX_i = 6.34|X_i) = 1 - \Lambda(\omega_{6.20} - \sum_{i=1}^{n=22} \beta_k X_i)$.

For the probabilities represented in equations (53a.) through (53h) to be positive, it must be the case that $0 < \omega_{3.08} < \omega_{4.38} < \omega_{4.44} < \omega_{4.48} < \omega_{5.65} < \omega_{6.20}$ (Greene 2012). [91]

Using the second energy insecurity index (setting the threshold value $\tau = 3.08$) households responses to the set of questions listed in Table 3.1 still determines their place along the scale, which determines the energy insecurity index value they receive. The energy insecurity index values are still ordered from least severe (most secure, least insecure) to most severe (least secure,

[91] The loglikelihood function for the ordered logit model is listed in the Chapter 4 Appendix.

most insecure). Our observations under the second method of partitioning can be recorded and ordered as follows,

(54a.) $EISINDEX_i = 0$ if $EINSECURE_i^* \leq \omega_{3.08}$

(54b.) $EISINDEX_i = 1.30$ if $\omega_{3.08} < EINSECURE_i^* \leq \omega_{4.38}$

(54c.) $EISINDEX_i = 1.36$ if $\omega_{4.38} < EINSECURE_i^* \leq \omega_{4.44.}$

(54d.) $EISINDEX_i = 1.40$ if $\omega_{4.44} < EINSECURE_i^* \leq \omega_{4.48}$

(54e.) $EISINDEX_i = 2.57$ if $\omega_{4.48} < EINSECURE_i^* \leq \omega_{5.65}$

(54f.) $EISINDEX_i = 3.12$ if $\omega_{5.65} < EINSECURE_i^* \leq \omega_{6.20}$

(54g.) $EISINDEX_i = 3.26$ if $\omega_{6.20} < EINSECURE_i^*$,

The conditional probability of each household i receiving an energy insecurity index value $EISINDEX_i$ can be derived as follows:

(55a.) $P(EISINDEX_i = 0|X_i) = P(EINSECURE_i^* \leq \omega_{3.08}|X_i)$

(55b.) $P(EISINDEX_i = 1.30|X_i) = P(\omega_{3.08} < EINSECURE_i^* \leq \omega_{4.38}|X_i)$

(55c.) $P(EISINDEX_i = 1.36|X_i) = P(\omega_{4.38} < EINSECURE_i^* \leq \omega_{4.44}|X_i)$

(55d.) $P(EISINDEX_i = 1.40|X_i) = P(\omega_{4.44} < EINSECURE_i^* \leq \omega_{4.48}|X_i)$

(55e.) $P(EISINDEX_i = 2.57|X_i) = P(\omega_{4.48} < EINSECURE_i^* \leq \omega_{5.65}|X_i)$

(55f.) $P(EISINDEX_i = 3.12|X_i) = P(\omega_{5.65} < EINSECURE_i^* \leq \omega_{6.20}|X_i)$

(55g.) $P(EISINDEX_i = 3.26|X_i) = P(\omega_{6.20} < EINSECURE_i^*|X_i)$

Again, replacing equation (35) for $EINSECURE_i^*$ and assuming a logistic model specification, the probabilities listed in equations (55a.) through (55g.) can be rewritten as follows

(56a.) $P(EISINDEX_i = 0|X_i) = 1 - \Lambda(\sum_{i=1}^{n=22} \beta_k X_i)$

(56b.) $P(EISINDEX_i = 1.30|X_i) = \Lambda(\omega_{4.38} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(-\sum_{i=1}^{n=22} \beta_k X_i)$

(56c.) $P(EISINDEX_i = 1.36|X_i) = \Lambda(\omega_{4.44} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{4.38} - \sum_{i=1}^{n=22} \beta_k X_i)$

(56d.) $P(EISINDEX_i = 1.40|X_i) = \Lambda(\omega_{4.48} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{4.44} - \sum_{i=1}^{n=22} \beta_k X_i)$

(56e.) $P(EISINDEX_i = 2.57|X_i) = \Lambda(\omega_{5.65} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{4.48} - \sum_{i=1}^{n=22} \beta_k X_i)$

(56f.) $P(EISINDEX_i = 3.12|X_i) = \Lambda(\omega_{6.20} - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_{5.65} - \sum_{i=1}^{n=22} \beta_k X_i)$

(56g.) $P(EISINDEX_i = 3.26|X_i) = 1 - \Lambda(\omega_{6.20} - \sum_{i=1}^{n=22} \beta_k X_i)$.

Because the energy insecurity index values a household can receive, following the third additional index can take on one of four values {1,2,3,4}, the energy insecurity index can still be considered an ordered response. Individual household observations under this method of partitioning can be recorded and ordered as follows,

(57a.) $EISINDEX_i = 1$ if $EINSECURE_i^* \leq \omega_2$

(57b.) $EISINDEX_i = 2$ if $\omega_2 < EINSECURE_i^* \leq \omega_3$

(57c.) $EISINDEX_i = 3$ if $\omega_3 < EINSECURE_i^* \leq \omega_{4.}$

(57d.) $EISINDEX_i = 4$ if $\omega_4 < EINSECURE_i^*$

where ω_i for $i = 1, 2, 3$ represent the different thresholds. The conditional distribution of $EINSECURE_i^*$, assuming a logistic model specification can be written as follows,

(58a.) $P(EISINDEX_i = 1|X_i) = P(EISINSECURE_i^* \leq \omega_2)$

(58b.) $P(EISINDEX_i = 2|X_i) = P(\omega_2 < EISINSECURE_i^* \leq \omega_3)$

(58c.) $P(EISINDEX_i = 3|X_i) = P(\omega_3 < EISINSECURE_i^* \leq \omega_4)$

(58d.) $P(EISINDEX_i = 4|X_i) = P(\omega_4 < EISINSECURE_i^*)$.

Again, replacing equation (35) for $EINSECURE_i^*$ the preceding probabilities can be rewritten as follows,

(59a.) $P(EISINDEX_i = 1|X_i) = 1 - \Lambda(\sum_{i=1}^{n=22} \beta_k X_i)$

(59b.) $P(EISINDEX_i = 2|X_i) = \Lambda(\omega_3 - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(- \sum_{i=1}^{n=22} \beta_k X_i)$

(59c.) $P(EISINDEX_i = 3|X_i) = \Lambda(\omega_4 - \sum_{i=1}^{n=22} \beta_k X_i) - \Lambda(\omega_3 - \sum_{i=1}^{n=22} \beta_k X_i)$

(59d.) $P(EISINDEX_i = 4|X_i) = 1 - \Lambda(\omega_4 - \sum_{i=1}^{n=22} \beta_k X_i)$.

Under this third method of partitioning, a household's energy insecurity status is still ordered from least severe (most secure, least insecure), which corresponds to an energy insecurity index value $EISINDEX_i = 1$ to most severe (least secure, most insecure), which corresponds to an energy insecurity index value $EISINDEX_i = 4$.

Of interest to us are the partial effects of changes in the individual regressors on the probability of receiving a specific energy insecurity index value $EISINDEX_i$ from the list of ordered responses. Positive parameter estimates should be interpreted as an increase in the log-odds of reaching a higher level of energy insecurity (i.e., becoming more energy insecure). Negative parameter estimates should be interpreted as a decrease in the log-odds of reaching a higher level of energy insecurity (i.e., becoming more energy secure).

ENDOGENITY ISSUE

Energy efficiency improvements are designed to decrease the amount of energy necessary to provide household energy services. Therefore, the experience of being energy insecure (i.e., not being able to provide an adequate level of energy services) may prompt households to make energy efficiency upgrades. Conversely, making an energy efficiency upgrade in the home may improve a household's energy security status. Clearly, in our analysis, the determination of causality as opposed to simple correlation will be quite difficult given this present simultaneous relationship between the two variables.

To confront this issue, we implement an instrumental variables approach to estimate the model in equation (35). To motivate the instrumental variables approach, consider first the simple representative equation for a household's energy insecurity status, represented by the energy insecurity index value $EISINDEX$ the household receives, with observation subscript i suppressed for convenience,

(60) $EISINDEX = \delta EE + \boldsymbol{x}'\boldsymbol{\theta}_1 + \boldsymbol{w}'\boldsymbol{\theta}_2 + \boldsymbol{r}'\boldsymbol{\theta}_3 + \epsilon,$

where the energy insecurity index $EISINDEX$ assigned to the household depends on which of the two index approaches described above is used (therefore, $EISINDEX$ can be either binary or censored/ordered in nature); EE is a vector of three binary variables and one continuous variable that together represent whether or not the household has made an energy efficiency upgrade sometime over the past twelve months; $\boldsymbol{x}$ and $\boldsymbol{w}$ are vectors of observable characteristics such that $\boldsymbol{x}$ affects both the energy insecurity status of the household and its decision of whether or not to make an energy efficiency upgrade; $\boldsymbol{w}$ is believed to only influence whether a household is energy insecure or energy insecure; $\boldsymbol{r}$ is a vector of unobservable characteristics, that similar to $\boldsymbol{x}$ affects both the energy insecurity status of the household and its decision to make an energy efficiency upgrade; δ is a scalar parameter; $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$, and $\boldsymbol{\theta}_3$ represent parameters to be estimated; and ϵ is the random error term.

In a single equation framework, one would estimate the following equation:

(61) $EISINDEX = \delta EE + \boldsymbol{x}'\boldsymbol{\theta}_1 + \boldsymbol{w}'\boldsymbol{\theta}_2 + \epsilon^*$

where $\epsilon^* = \boldsymbol{r}'\boldsymbol{\theta}_3 + \epsilon$. Ordinary least-squares (OLS) estimation of equation (61) however, produces biased estimates of δ because $E(\epsilon^*|ES, \boldsymbol{x}, \boldsymbol{w}) \neq 0$. One remedial measure is to estimate the energy insecurity status equation in (60) using an instrumental variables (IV) approach. Before we apply the IV approach, we attempt to account for the censored nature of the energy insecurity index values and binary nature of three of the four measures that represent the household's decision to make an energy efficiency upgrade/improvement. Thus, we replace the variables $EISINDEX$ and EE with their latent counterparts $EINSECURE^*$ and EE^*,

(62) $EE^* = \boldsymbol{x}'\boldsymbol{\alpha}_1 + \boldsymbol{z}'\boldsymbol{\alpha}_2 + v_1$

(63) $EINSECURE^* = \delta EE^* + \boldsymbol{x}'\boldsymbol{\theta}_1 + \boldsymbol{w}'\boldsymbol{\theta}_2 + v_2$

Here $\boldsymbol{z}$ is a vector of observable characteristics, also known as instruments, believed to influence the household's decision to make an energy efficiency upgrade but not its true energy insecurity status, which is represented by $EINSECURE^*$; $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are parameters to be estimated; and the error terms v_1 and v_2 represent unobserved factors believed to influence a household's decision to make an energy efficiency upgrade and its energy security status respectively.

The reduced form equation system constitutes equation (62) and

(64) $EINSECURE^* = \boldsymbol{x}'(\delta\boldsymbol{\alpha}_1 + \boldsymbol{\theta}_1) + \boldsymbol{z}'(\delta\boldsymbol{\alpha}_2) + \boldsymbol{w}'\boldsymbol{\theta}_2 + v_2^*$

where $v_2^* = \delta v_1 + v_2$. On the basis of the reduced form equations (62) and (64), binary measures for energy efficiency upgrades (the binary variables in the vector EE) and the binary and censored energy insecurity index value of each household i can be characterized by the following relationship:

(65) $EE = 1(EE^* > 0)$ and $EISINDEX = 1\ (EINSECURE^* > 0)$ or

(66) $EE = 1(EE^* > 0)$ and $EISINDEX = \max(0, EINSECURE^*)$.

Here $1(\circ)$ denotes an indicator function, taking a value of 1 if event $(EE^* > 0)$ holds, and 0 otherwise. In our analysis, the event $(EE^* > 0)$corresponds with whether or not the household made an energy efficiency upgrade or improvement over the past twelve months.

Recall from equation (35) that in our analysis the first endogenous variable identified, EE, is represented by the inclusion of four different variables: (1) X_{23} a binary variable equal to one if the household received an in-home energy audit; (2) X_{24} a binary variable equal to one if the household installed triple and/or double pane windows; (3) X_{25} a binary variable equal to one if the household has adequate insulation; and (4) X_{26} a continuous variable equal to the number of *Energy Star®* certified appliances operated by the household.

Given we have four possible endogenous variables, in order to have a just identified system we need at least four separate instrumental variables. Within the 2015 RECS, we identified the following four variables as potential instruments for our analysis: Z_1 a binary variable equal to one if the household received a free in-home energy audit sometime over the past twelve months; Z_2 an indicator variable equal to one if the head of the household self-reports as an owner of the home and 0 otherwise (i.e., reports as a renter or temporary resident); Z_3 a binary variable equal to one if the household received an energy efficiency appliance rebate from their utility company to upgrade their capital stock of appliances; Z_4 a continuous variable equal to the number of months the household used a dehumidifier.

In our analysis Z_1 is a potential instrument for X_{23}, Z_2 is a potential instrument for X_{24}, Z_3 is a potential instrument for X_{25}, and Z_4 is a potential instrument for X_{26}. These variables were identified as potential instruments because they are believed to influence a household's decision to make an energy efficiency upgrade, but not whether or not the household identifies as energy insecure. For variable Z_1 through Z_4 to be considered as valid instruments, they must meet the following two conditions: (1) be uncorrelated with the error term v_2, or in other words be excludable in the sense that they have no direct effect on a household's energy insecurity status (i.e., they are exogenous); and (2) be correlated with the endogenous variables identified (X_{23} through X_{26}) for which they serve as instrument.

Because the first condition involves examining the covariance between the identified instruments Z_i and the unobserved error term v_2 we cannot test whether or not it holds. By contrast, the condition that the instruments Z_i are correlated with the endogenous variables of interest, after controlling for the other exogenous variables of interest can be tested. The easiest way to test whether or not condition (2) holds is to estimate the following equation using a simple linear regression of the form:

(67) $X_i = \theta_0 + \theta_1 X_1 + \cdots + \theta_{18} X_{22} + \pi_k Z_i + \zeta_i$

Here each endogenous variable identified is projected linearly onto all of the exogenous variables in the model (i.e., the exogenous variables in $\boldsymbol{w}$ and the proposed instrument Z_i). The key assumption from this linear regression is that π_k (the coefficient on the candidate instrumental variable identified) is non-zero. More specifically, if condition (2) holds we should be able to conduct a simple *t*-test which will lead to the rejection of the null hypothesis $H_0: \pi_k = 0$ against the two-sided alternative $H_A: \pi_k \neq 0$. If this is the case, then we can be fairly confident that Z_i is partially correlated with X_i, once the other exogenous variables have been netted out. To test if condition (2) holds, we estimate the linear model in equation (67) for each endogenous variable identified X_i for $i = 23, .., 26$ and the chosen instrument. The results are listed in Table 4.3.

Table 4.3 Test of Exclusion Restriction Results

Instrument	X_{23}: Audit	X_{24}:Windows	X_{25}:Insulation	X_{26}: *Energy Star®*
Z_1:Free Audit	0.9143*** (0.0220)	-	-	-
Z_2: Ownership	-	0.1443*** (0.0184)	-	-
Z_3: Dehumidifier	-	-	0.0031* (0.0018)	-
Z_4: App. Rebate	-	-	-	0.5759*** (0.1337)
Observations	**5,686**			

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

As stated earlier, the other endogenous variable, a household's energy insecurity status is constructed from each household i's responses to a set of questions related to challenges they may have faced over the past twelve months in maintaining adequate access to a sufficient, safe, and affordable energy supply to meet their basic daily energy service needs. Responses to the questions of interest were combined into a scaled measure, which was used to create two index measures for household energy insecurity. In the first index measure we created, a household's energy insecurity status was treated a binary variable, set equal one if the scale score produced was positive (household identifies as being energy insecure) and zero otherwise (household identifies as being energy secure because it did not respond affirmatively to any of the questions).

In the second case, the energy insecurity index measure created is positive and strictly greater than or equal to zero. The index values produced represent an ordered response outcome variable. The values assigned to each outcome indicate the extent of "energy insecurity" being experienced by each household i. While the values are arbitrary, the distance between them conveys important information about what circumstances contribute to household energy insecurity. Using the Dichotomous Rasch model, we created two different ordered response energy insecurity indices. In the first index we created, there are only eight possible values for the energy insecurity index measure. In the second index, there are seven possible values. In the third, there are four.

Based on both of the methods used to create the household energy security index (i.e., the binary index measure and the two ordered response index measures), the model for household energy insecurity in equation (35) is assumed to be non-linear. Therefore, in addition to having an identification issue in our analysis, an additional problem arises given the structure of our endogenous variables and the type of empirical model specifications required to account for the binary or ordered and censored nature of the index value chosen to be evaluated. More specifically, three of the four endogenous variables we identified are binary in nature. Consequently, the conditional expectation function associated with the first stage regression model outlined above, is likely to be non-linear (Angrist and Pischke 2008).

As a result, applying the usual OLS procedure in the first stage only serves as an approximation to the underlying true conditional expectation function of the energy efficiency indicator variables of interest. One way to confront this issue would be to estimate a non-linear, first stage regression model and then as before take the predicted values from that regression (i.e., the reduced form model) and plug them back into the second stage regression. However, this results in a forbidden regression because only OLS estimation of the first-stage is guaranteed to produce first-stage residuals that are uncorrelated with the fitted values and covariates used in the second stage (Angrist and Pischke 2009). An alternative would be to use the non-linear fitted values for each X_i produced in the first stage as instruments for themselves (Agrist and Pischke 2009). In other words, we could use $\hat{X}_i$ as an instrument for X_i and apply the same two-stage procedure described before. However, using non-linear first stage estimates as instruments, implicitly assumes non-linearity in the first-stage as a source of identification.

Another alternative suggested, in the case when both the endogenous variable and dependent variable are binary in nature (as is the case for the first energy insecurity index we created) is to estimate Bivariate Probit model (Angrist and Pischke 2009; Wooldridge 2010). The Bivariate Probit model can be specified for two binary response variables as follows,

(68) $EE^* = 1[\boldsymbol{x}'\boldsymbol{\alpha}_1 + \boldsymbol{z}'\boldsymbol{\alpha}_2 > v_1]$

(69) $EINSECURE^* = 1[\delta EE^* + \boldsymbol{x}'\boldsymbol{\theta}_1 + \boldsymbol{w}'\boldsymbol{\theta}_2 > v_2]$

where, as before $\boldsymbol{x}$ is a vector of exogenous variables believed to influence both the household's decision to make an energy efficiency investment and whether or not the household identifies as energy insecure; $\boldsymbol{w}$ is a vector of observable characteristics, believed to influence whether or not a household identifies as energy insecure; and $\boldsymbol{z}$ are valid instrumental variables. [92]The source of endogeneity in the Bivariate Probit set-up presented by equations (68) and (69) is the correlation between v_1 and v_2. That is, unmeasured random factors that influence a household's decision to make an energy efficiency upgrade are likely correlated with unmeasured random determinants of whether or not the household is energy insecure (Angrist and Pischke 2009).

Assuming the error terms, v_1 and v_2 have a joint bivariate normal distribution, we estimate the system of equations in (68) and (69) using maximum likelihood estimation (MLE).[93] MLE make no assumptions on the structure of the endogenous variables. Instead MLE allows the endogenous variables to be discrete, limited (i.e., categorical), or continuous variables. However, it places strong assumptions on the joint distribution of the error terms and is only applicable in the case of the first index measure we construct (when $EISINDEX = 1$ or 0), when the endogenous variable is binary, as in the case with X_{23}, X_{24}, X_{25}. When continuous endogenous variables are included, the Bivariate Probit is no longer applicable.

Instead, a control function (CF) approach is suggested. A CF approach can also be used in the case when second index measure (i.e., the ordered response energy insecurity index values) is used. Similar to the IV (2SLS) estimation procedure, the CF approach uses extra regressors in an attempt to break up the correlation between the unobserved effects and the included endogenous variables (Wooldridge 2010). However, rather than using the predicted values from the first stage regression as the additional regressors in the structural model, the CF approach uses the residuals from the first stage regression.

[92] The loglikelihood function for the Bivariate Probit model is listed in the Chapter 4 Appendix.
[93] Stata is used to estimate the models presented in this Chapter (Essay).

The CF approach for the binary energy insecurity index measure can be outlined as follows. Consider the model for a household's energy insecurity status, represented as before, by the energy insecurity index value $EISINDEX$ it receives,

(70) $EISINDEX_i = \delta EE + \boldsymbol{x}'\boldsymbol{\theta}_1 + \boldsymbol{w}'\boldsymbol{\theta}_2 + v_2.$

As before, EE is a vector of endogenous variables that represent whether or not the household has made an energy efficiency upgrade; $\boldsymbol{x}$ is a vector of variables that influence whether or not a household is energy insecure, as well as the household's decision to make an energy efficiency upgrade; $\boldsymbol{w}$ is a vector of observable characteristics believed to only influence whether or not a household identifies as energy insecure; and v_2 is the error term. Here $EISINDEX_i$ can take on one of two values: 0 if the household identifies as "Energy Secure" and 1 if the household identifies as "Energy Insecure." To implement the CF approach, similar to the instrumental variables approach we estimate a reduced form model for EE as follows:

(71) $EE = \boldsymbol{x}'\boldsymbol{\alpha}_1 + \boldsymbol{z}'\boldsymbol{\alpha}_2 + v_1,$

where $\boldsymbol{z}$ is a vector of exogenous variables known as instrumental variables that influence a household's decision to make an energy efficiency upgrade but not whether or not the household identifies as energy insecure. The CF approach is supported by the idea that the structural error term v_2 and the reduced form error term v_1 can be captured using a linear relationship as follows

(72) $v_2 = \rho_1 v_1 + e_1.$

Because neither v_1 or v_2 are assumed to be correlated with the exogenous variables in z then e_1 is also assumed to not be correlated with z and therefore, e_1 is not correlated with EE (Wooldridge 2015b). Therefore, we can obtain consistent parameter estimates by plugging v_1 into the structural equation for $EISINDEX_i$. Because we do not observe $v_1$we estimate it from the results of the first stage regression in (71). It is important to note that the variables included in EE

can be binary or continuous because the reduced form equations for each of the endogenous variables are estimated as linear projections. As long the instruments are valid, then regardless of the nature of EE (i.e., whether the variable under consideration in EE is continuous, categorical, or binary) the linear reduced form of equation (71) can always be specified (Wooldridge 2015b).

Because the Bivariate Probit model is estimated for the case when the energy insecurity index is binary and the endogenous variables are binary, we estimate a CF approach for the case when the endogenous variable is continuous (as in the case with X_{21}) and the energy insecurity index is binary. To implement the control function approach, we take the following steps. First, we estimate the reduced form equation in (71) by OLS and obtain the residuals,

(73) $\hat{v}_1 = EE - \boldsymbol{x}'\widehat{\boldsymbol{\alpha}}_1 - \boldsymbol{z}'\widehat{\boldsymbol{\alpha}}_2.$

The residuals are included as explanatory variables in the structural equation and a logistic model specification is assumed (Wooldridge 2010). Bootstrapped standard errors are suggested (Murray and Topel 1985; Newey and McFadden 1994).

The CF can also be applied in the case of the second energy insecurity index we create, assuming the same general set-up as described before. Under this case, the first stage reduced form equation in (71) is still estimated by OLS for each of the endogenous variables, both for the endogenous variables that are binary and the endogenous variables that are continuous. The residuals are obtained from each first stage regression and used as regressors in the second stage regression (i.e., the structural equation) which are estimated by MLE. Again, bootstrapped standard errors are suggested to account for the variation introduced by the inclusion of estimated value for the error term in the second stage regression, v_1 (i.e., $\hat{v}_1$). Following this suggestion, we bootstrap the standard errors.

4.5 ESTIMATION RESULTS

The parameter estimates from estimating equation (35) via maximum likelihood estimation, controlling for endogeneity are presented below in Tables 4.4 and 4.5. Table 4.4 lists the Bivariate Probit Model results. The Bivariate Probit model provides estimates for the effects of energy efficient windows, adequate home insulation, having received an in-home energy audit, and other explanatory variables on the probability that a household will identify as energy insecure. Equations are estimated separately following the Bivariate Probit model estimation techniques outlined in Wooldridge (2010). Table 4.5 provides the maximum likelihood parameter estimates from using a CF approach to estimate a logistic model which examines the effects of having *Energy Star* ® certified appliances in the household and other explanatory variables on the probability that a household will identify as energy insecure.

Table 4.4 Bivariate Probit Model Results on Effects of Energy Audits, Insulation, Windows, and other Explanatory Variables on Household Energy Insecurity

	Bivariate Probit (Audit)		**Bivariate Probit (Insulation)**		**Bivariate Probit (Windows)**	
Explanatory Variables	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
Income	-0.0001***	6.87 E-07	-9.85 E-06***	7.00 E-07	-9.25 E-06***	7.75 E-07
Employment	0.026	0.048	0.013	0.049	0.023	0.047
Education	-0.143***	0.046	-0.135***	0.046	-0.132***	0.045
Age	-0.010***	0.001	-0.009***	0.002	-0.009***	0.002
Gender	-0.179***	0.042	-0.176***	0.042	-0.173***	0.041
Children	0.114***	0.021	0.107***	0.022	-0.110***	0.021
CDD65	6.41 E-06	4.06 E-05	8.08 E-06	4.07 E-05	-2.97 E-05	4.22 E-05
HDD65	-3.84 E-05	2.73 E-05	-3.44 E-05	2.77 E-05	-2.48 E-05	2.74 E-05
House Age	0.072	0.050	-0.013	0.067	0.001	0.056
Hispanic	0.185***	0.060	0.173***	0.061	0.153***	0.061
African American	0.402***	0.063	0.398***	0.064	0.358***	0.066
Bedrooms	0.021	0.027	0.024	0.028	0.023	0.027
Bathrooms	-0.074*	0.039	-0.026	0.047	-0.011	0.044
Census Dummies	Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes	
Audit	0.0216	0.150				
Insulation			-0.589***	0.324		
Windows					-0.595***	0.211
Constant	1.010***	0.261	1.367***	0.322	1.145***	0.260
Log-likelihood	-3,796.22		-4,872.38		-5,822.42	
$\rho(v_1 v_2)$	-0.023	0.083	0.065	0.186	0.285	0.126
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4.5 CF Logistic Model Results on Effects of *Energy Star* ® Appliances and other Explanatory Variables on Household Energy Insecurity

	Logit (Control Function) *Energy Star* ®			
	Change in the Log-Odds		**Change in the Odds**	
Explanatory Variables	**Coefficient Estimate**	**St. Error**	**Odds Ratio Estimate**	**St. Error**
Income	-1.49E-05***	1.76E-06	1.00***	1.54E-06
Employment	0.093	0.093	1.10	0.10
Education	-0.166***	0.074	0.85*	0.07
Age	-0.017***	0.002	0.98***	0.00
Gender	-0.310***	0.074	0.73***	0.05
Children	0.211***	0.033	1.23***	0.04
CDD65	1.80E-05	8.79E-05	1.00	0.00
HDD65	-5.14E-05	5.74E-05	1.00	0.00
House Age	0.066	0.074	1.07	0.09
Hispanic	0.177	0.118	1.19	0.14
African American	0.564***	0.117	1.76***	0.19
Bedrooms	0.058	0.052	1.06	0.06
Bathrooms	-0.045	0.076	0.96	0.07
Census Dummies	Yes	Yes	Yes	Yes
Housing Dummies	Yes	Yes	Yes	Yes
Energy Star ®	-0.320***	0.114	0.73	0.08
$\hat{v}_1$	0.307***	0.119	1.36	0.16
Constant	2.011	0.522	7.47	3.10
Log-Likelihood	-2,552.07			
Pseudo R^2	0.1457			
Observations	5,686			

$*\, p < 0.10,\ **\, p < 0.05,\ ***\, p < 0.01$

The Bivariate Probit model results showing the effects of energy efficiency investments on the probability that a household will identify as energy insecure ($EISINDEX = 1$) are listed in Table 4.6 below.

Table 4.6 Bivariate Probit Model Predicted Probabilities of Energy Audits, Windows, and Insulation on Household Energy Insecurity

Explanatory Variables	$P(EISINDEX = 1)$	$P(EISINDEX = 0)$
$Audit = 1$	0.017	0.064
$Audit = 0$	0.201	0.718
$Windows = 1$	0.142	0.472
$Windows = 0$	0.085	0.300
$Insulation = 1$	0.175	0.652

$Insulation = 0$	0.045	0.128

The CF logistic model results showing the effects of a household's number of *Energy Star* ® certified appliances on the expected probability that a household will identify as energy insecure ($EISINDEX = 1$) of, based on are presented in Table 4.7.

Table 4.7 CF Logistic Model Predicted Probabilities of *Energy Star* ® Appliances on Energy Insecurity

Explanatory Variables	$P(EISINDEX = 1)$
$EnergyStar = 0$	0.311
$EnergyStar = 1$	0.247
$EnergyStar = 2$	0.192
$EnergyStar = 3$	0.147
$EnergyStar = 4$	0.112
$EnergyStar = 5$	0.084
$EnergyStar = 6$	0.062
$EnergyStar = 7$	0.046

Because the dependent variable across all model specifications is a latent class variable used to measure household energy insecurity, we can interpret the estimated coefficients from the CF logistic model and Bivariate Probit model broadly as the likelihood that a household will identify as energy insecure (i.e. not energy secure). Consistent with our theoretical expectations, across both model specifications we find households with higher incomes are significantly less likely to identify as being energy insecure. In addition, consistent with the previous literature we find African American or Hispanic households are more likely to identify as energy insecure. Similar to Drehobl and Ross (2016), we find households with more children living in the home age 16 or younger are more likely to identify as being energy insecure.

Across both the CF logistic model and Bivariate Probit model specifications, we find males (Gender = 1) are less likely to identify as energy insecure as compared to females (Gender = 0). As the age of the respondent increases, the likelihood they will identify as energy insecure decreases. More than 20% of the individuals in our sample indicated they were past full retirement age (older than 66 years old). Because individuals who are retired typically live on a fixed income, we would expect these households to be more likely to identify as energy insecure. The average income for households who are past full retirement age in our sample is approximately $55,000/year. In addition, nearly 22% of these households indicated they are still employed, either part time or full time. While these households may be receiving retirement benefits (i.e., social security, Medicaid), they are also receiving additional income from working. Having additional income may prevent these households from experiencing energy-service related hardships that prevent them from maintaining consistent adequate access to energy services. [94]

The Bivariate Probit model results indicate that having adequate insulation (Insulation = 1) and triple or double pane windows (Windows = 1) decreases the likelihood that a household will identify as energy insecure. Following the CF logistic model, we find as the number of *Energy Star* ® certified appliances increase in the home, the likelihood that a household identifies as energy insecure decreases. However, according to the Bivariate Probit model results, we fail to reject our hypothesis that having received an in-home energy audit *(*Audit = 1) decreases the likelihood that a household will identify as energy insecure. Tables 4.8, 4.9, and 4.10 list the maximum-likelihood parameter estimates from estimating equation (35) using an Ordered Logit model under the three different methods of partitioning households into different energy insecurity groups.[95]

[94] It is also important to note that none of the individuals who responded to the survey live in a retirement home. While they may live in a retirement community, they must occupy their own individual residence to be considered eligible to participate in the RECS.
[95] The Ordered Logit results under all three methods of partitioning households into different energy insecurity groups, but not controlling for endogeneity, are presented in the Appendix of this chapter for the reader.

Table 4.8 Ordered Logit Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with Threshold Value $\tau = 0$

	Ordered Logit (Audit)		**Ordered Logit (Insulation)**		**Ordered Logit (Windows)**		**Ordered Logit (*Energy Star*®)**	
Explanatory Variables	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
Income	-1.82E-05***	1.27E-06	-1.75E-05***	1.35E-06	-1.61E-05***	1.38E-06	-1.52E-05***	1.93E-06
Employment	0.038	0.096	-0.012	0.101	0.029	0.090	0.083	0.088
Education	-0.220***	0.086	-0.201***	0.073	-0.196***	0.066	-0.152***	0.076
Age	-0.017***	0.002	-0.014***	0.002	-0.013***	0.003	-0.016***	0.002
Gender	-0.328***	0.074	-0.299***	0.069	-0.312***	0.070	-0.316***	0.072
Children	0.184***	0.032	0.156***	0.035	0.184***	0.034	0.200***	0.033
CDD65	3.03E-05	6.55E-05	2.96E-05	6.49E-05	-1.00E-04	9.39E-05	2.64E-05	8.14E-05
HDD65	-5.47E-05	4.94E-05	-3.42E-05	4.61E-05	-9.47E-06	5.02E-05	-4.10E-05	5.39E-05
House Age	0.124	0.084	-0.227	0.143	-0.124	0.132	0.064	0.092
Hispanic	0.280***	0.117	0.229***	0.099	0.193**	0.101	0.174	0.114
Af. American	0.625***	0.120	0.585***	0.090	0.493***	0.116	0.527***	0.119
Bedrooms	0.038	0.043	0.043	0.047	0.041	0.038	0.064	0.048
Bathrooms	-0.130***	0.064	0.069	0.087	0.079	0.102	-0.047	0.086
Census Dummies	Yes		Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes		Yes	
Audit	0.432	0.273	-	-	-	-	-	-
Insulation	-	-	-2.543***	0.870	-	-	-	-
Windows	-	-	-	-	-2.087***	0.687	-	-
Energy Star®	-	-	-	-	-	-	-0.316***	0.146
$\hat{\upsilon}_1$	-0.122	0.303	1.655*	0.907	1.895***	0.707	0.301***	0.148
$\widehat{\omega}_{3.08}$	-1.923	0.431	-3.362	0.677	-2.390	0.495	-2.109	0.463
$\widehat{\omega}_{4.38}$	-0.803	0.435	-2.212	0.670	-1.269	0.485	-0.990	0.464
$\widehat{\omega}_{4.44}$	0.091	0.446	-1.295	0.683	-0.373	0.495	-0.094	0.465

$\widehat{\omega}_{4.48}$	1.042	0.432	-0.330	0.668	0.578	0.501	0.856	0.497
$\widehat{\omega}_{5.65}$	2.453	0.438	1.088	0.670	1.988	0.556	2.266	0.552
$\widehat{\omega}_{6.20}$	4.334	0.645	2.973	0.983	3.870	0.769	4.148	0.839
Log likelihood	-3,994.02		-3,935.24		-3,989.63		-3,993.87	
Pseudo R^2	0.102		0.115		0.103		0.102	
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4.9 Ordered Logit Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with Threshold Value $\tau = 3.08$

	Ordered Logit (Audit)		**Ordered Logit (Insulation)**		**Ordered Logit (Windows)**		**Ordered Logit (*Energy Star*®)**	
Explanatory Variables	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
Income	-2.02E-05***	2.15E-06	-1.92E-05***	1.72E-06	-1.77E-05***	2.24E-06	-1.69E-05***	2.59E-06
Employment	0.010	0.107	-0.049	0.124	0.001	0.098	0.057	0.128
Education	-0.208***	0.089	-0.192*	0.118	-0.187*	0.103	-0.141	0.102
Age	-0.016***	0.003	-0.013***	0.003	-0.012***	0.004	-0.015***	0.003
Gender	-0.406***	0.121	-0.374***	0.104	-0.385***	0.097	-0.391***	0.084
Children	0.134***	0.046	0.097***	0.042	0.132***	0.046	0.149***	0.048
CDD65	4.19E-05	8.33E-05	3.18E-05	1.13E-04	-1.09E-04	1.05E-04	3.46E-05	8.68E-05
HDD65	-1.75E-05	6.58E-05	2.25E-06	6.66E-05	3.16E-05	6.73E-05	-3.91E-06	5.51E-05
House Age	0.134	0.124	-0.239	0.213	-0.140	0.185	0.073	0.110
Hispanic	0.222	0.169	0.194	0.137	0.131	0.150	0.118	0.175
Af. American	0.458***	0.150	0.429***	0.155	0.313*	0.161	0.359***	0.160
Bedrooms	0.046	0.058	0.056	0.060	0.053	0.054	0.076	0.056
Bathrooms	-0.087	0.079	0.129	0.109	0.146	0.160	-0.004	0.113

Census Dummies	Yes		Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes		Yes	
Audit	0.424	0.318	-	-	-	-	-	-
Insulation	-	-	-2.597***	1.079	-	-	-	-
Windows	-	-	-	-	-2.342***	1.153	-	-
Energy Star ®	-	-	-	-	-	-	-0.327*	0.187
$\hat{v}_1$	-0.005	0.380	1.502	1.097	2.133*	1.169	0.300	0.188
$\widehat{\omega}_{4.38}$	-0.849	0.522	-2.268	0.887	-1.356	0.622	-1.025	0.565
$\widehat{\omega}_{4.44}$	0.049	0.537	-1.342	0.888	-0.458	0.626	-0.128	0.567
$\widehat{\omega}_{4.48}$	1.001	0.547	-0.372	0.908	0.494	0.632	0.824	0.587
$\widehat{\omega}_{5.65}$	2.411	0.567	1.048	0.958	1.904	0.644	2.234	0.628
$\widehat{\omega}_{6.20}$	4.293	0.858	2.935	1.003	3.786	0.673	4.116	0.918
Log likelihood	-2,145.17		-2,093.68		-2,142.83		-2,145.68	
Pseudo R^2	0.1035		0.1250		0.1045			
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4.10 Ordered Logit Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with Four Energy Security Groups

	Ordered Logit (Audit)		**Ordered Logit (Insulation)**		**Ordered Logit (Windows)**		**Ordered Logit (*Energy Star*®)**	
Explanatory Variables	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
Income	-2.01E-05***	2.20E-06	-1.89E-05***	1.88E-06	-1.73E-05***	2.68E-06	-1.63E-05***	2.85E-06
Employment	0.013	0.099	-0.046	0.096	0.003	0.116	0.066	0.093
Education	-0.219*	0.120	-0.200***	0.085	-0.193*	0.102	-0.141	0.131

Age	-0.016***	0.003	-0.013***	0.003	-0.011***	0.004	-0.015***	0.003
Gender	-0.408***	0.097	-0.378***	0.099	-0.387***	0.100	-0.392***	0.115
Children	0.122***	0.046	0.087*	0.050	0.121***	0.046	0.140***	0.052
CDD65	4.81E-05	1.04E-04	4.42E-05	1.02E-04	-1.19E-04	1.09E-04	3.89E-05	1.04E-04
HDD65	-1.05E-05	5.95E-05	1.19E-05	6.95E-05	4.26E-05	6.64E-05	3.54E-06	7.61E-05
House Age	0.123	0.132	-0.287	0.223	-0.180	0.177	0.054	0.142
Hispanic	0.239*	0.144	0.195	0.179	0.136	0.147	0.121	0.154
Af. American	0.415***	0.114	0.375***	0.135	0.256*	0.152	0.304	0.204
Bedrooms	0.041	0.063	0.050	0.061	0.048	0.067	0.075	0.061
Bathrooms	-0.086	0.097	0.151	0.130	0.172	0.148	0.010	0.139
Census Dummies	Yes		Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes		Yes	
Audit	0.430	0.516	-	-	-	-	-	-
Insulation	-	-	-2.902***	1.356	-	-	-	-
Windows	-	-	-	-	-2.591***	0.925	-	-
Energy Star ®	-	-	-	-	-	-	-0.372*	0.215
$\hat{\upsilon}_1$	-0.015	0.499	1.810	1.364	2.389***	0.912	0.347*	0.210
$\widehat{\omega}_1$	-0.811	0.585	-2.388	1.049	-1.384	0.534	-1.022	0.697
$\widehat{\omega}_2$	2.441	0.591	0.926	1.112	1.871	0.608	2.231	0.702
$\widehat{\omega}_3$	4.323	0.891	2.813	1.266	3.753	0.876	4.113	0.791
Log likelihood	-1,655.04		-1,603.99		-1,652.14		-1,654.98	
Pseudo R^2	0.128		0.155		0.130		0.128	
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

When the threshold value of τ = 0 (Table 4.8), the Ordered Logit model results indicate that household respondents who have a bachelor's degree or beyond (Education = 1) display an approximately a 0.2 decrease in the log-odds that the household will be in a higher energy insecurity category than they are in currently, assuming all else equal. In addition, results indicate that if Income increased by $10,000, the log-odds that a household will be in a higher energy insecurity category than they are currently in will decrease between 0.152 and 0.182. Additional results of the Ordered Logit model when τ = 0 (Table 4.8) are as follows. The log-odds of being placed in a higher energy insecurity category decrease as the age of the respondent of the household increases. For every additional child born, the log-odds of being in a more severe energy insecurity category increase between 0.156 and 0.184. Finally, the log-odds of being in a higher energy insecurity category are higher for households who identify as Hispanic or African American.

We found similar Order Logit results when the threshold value of τ = 3.08 (Table 4.9) and when four different energy insecurity groups are specified (Table 4.10). When the threshold value of τ = 3.08 and only seven energy insecurity groups are considered, fewer households in our sample overall identify as energy insecure ($EISINDEX > 0$). As with the previous model with τ = 0, the results with τ = 3.08 indicate that having a higher income, being older, having a college degree or beyond, and being a male decreases the log-odds of a household being in a higher energy insecurity category (i.e., a category other than $EISINDEX = 0$). Also, the results with τ = 3.08 still indicate that the log-odds of being in a higher energy insecurity category increases if the household identifies as African American. However, identifying as Hispanic does significantly affect the log-odds that a household will be in a more severe energy insecurity category.

In the Four Group model (see Table 4.10), the change in the log-odds of being in a more energy insecure category are consistent with the results when the threshold value of τ = 0 and τ

= 3.08. In the Four Group model, we found that having one more child increases the log-odds of being in a more energy insecure category by 0.10. The results listed in Tables 4.8, 4.9, and 4.10 can be interpreted as changes in the log-odds of being in a higher energy insecurity group (i.e., being more energy insecure). In addition to the above results we also estimate odds-ratios that correspond to each set of models and results. The odds-ratios for each of three different methods of partitioning are listed in Tables 4.11, 4.12, and 4.13. Predicted probabilities follow these results in Tables 4.14, 4.15, and 4.16.

Table 4.11 Odds Ratio Ordered Logit Results of Effects of Energy Audits, Insulation, Windows, *Energy Star®* Appliances, and other Explanatory Variables on Household Energy Insecurity with Threshold Value $\tau = 0$

	Ordered Logit (Audit)		**Ordered Logit (Insulation)**		**Ordered Logit (Windows)**		**Ordered Logit (*Energy Star®*)**	
Variables	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**
Income	1.000***	1.34E-06	1.000***	1.15E-06	1.000***	1.56E-06	1.000***	1.70E-06
Employment	1.039	0.106	0.988	0.068	1.030	0.067	1.086	0.098
Education	0.802***	0.070	0.818***	0.060	0.822***	0.055	0.859***	0.075
Age	0.983***	0.003	0.986***	0.002	0.987***	0.003	0.984***	0.002
Gender	0.720***	0.052	0.741***	0.044	0.732***	0.046	0.729***	0.042
Children	1.202***	0.044	1.169***	0.048	1.202***	0.043	1.222***	0.039
CDD65	1.000	6.83E-05	1.000	7.13E-05	1.000	8.38E-05	1.000	6.04E-05
HDD65	1.000	4.73E-05	1.000	4.60E-05	1.000	5.22E-05	1.000	4.39E-05
House Age	1.132	0.098	0.797	0.105	0.883	0.110	1.066	0.099
Hispanic	1.323***	0.130	1.258***	0.143	1.213	0.137	1.190	0.140
Af. American	1.869***	0.226	1.796***	0.169	1.638***	0.191	1.694***	0.182
Bedrooms	1.039	0.054	1.044	0.040	1.042	0.045	1.066	0.052
Bathrooms	0.878	0.074	1.071***	0.089	1.082	0.096	0.954	0.066
Census Dummies	Yes		Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes		Yes	
Audit	1.540	0.442	-	-	-	-	-	-
Insulation	-	-	0.079***	0.062	-	-	-	-
Windows	-	-	-	-	0.124***	0.086	-	-
Energy Star ®	-	-	-	-	-	-	0.729***	0.099
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4.12 Odds Ratio Ordered Logit Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with Threshold Value $\tau = 3.08$

	Ordered Logit (Audit)		**Ordered Logit (Insulation)**		**Ordered Logit (Windows)**		**Ordered Logit (*Energy Star*®)**	
Variables	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**
Income	1.000***	2.25E-06	1.000***	1.80E-06	1.000***	2.04E-06	1.000***	2.40E-06
Employment	1.010	0.111	0.952	0.102	1.001	0.115	1.059	0.110
Education	0.812*	0.101	0.825*	0.093	0.830	0.100	0.869	0.096
Age	0.984***	0.004	0.987***	0.003	0.989***	0.003	0.985***	0.003
Gender	0.666***	0.083	0.688***	0.071	0.680***	0.078	0.676***	0.074
Children	1.143***	0.058	1.102***	0.053	1.141***	0.052	1.160***	0.053
CDD65	1.000	9.19E-05	1.000	7.17E-05	1.000	1.02E-04	1.000	8.60E-05
HDD65	1.000	5.39E-05	1.000	5.02E-05	1.000	5.80E-05	1.000	5.51E-05
House Age	1.144	0.138	0.787	0.169	0.869	0.159	1.076	0.151
Hispanic	1.249	0.202	1.215	0.200	1.140	0.197	1.126	0.199
Af. American	1.582***	0.180	1.536***	0.248	1.368*	0.227	1.432***	0.237
Bedrooms	1.047	0.075	1.058	0.064	1.055	0.063	1.080	0.067
Bathrooms	0.917	0.089	1.138	0.137	1.157	0.151	0.996	0.109
Census Dummies	Yes		Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes		Yes	
Audit	1.528	0.646	-	-	-	-	-	-
Insulation	-	-	0.074***	0.075	-	-	-	-
Windows	-	-	-	-	0.096***	0.86	-	-
Energy Star ®	-	-	-	-	-	-	0.721*	0.125
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4.13 Odds Ratio Ordered Logit Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with Four Energy Security Groups

	Ordered Logit (Audit)		**Ordered Logit (Insulation)**		**Ordered Logit (Windows)**		**Ordered Logit (*Energy Star*®)**	
Explanatory Variables	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**	**Odds Ratio**	**St. Error**
Income	1.000***	2.10E-06	1.000***	2.40E-06	1.000***	2.00E-06	1.000***	2.05E-06
Employment	1.013	0.113	0.955	0.114	1.003	0.107	1.068	0.112
Education	0.804***	0.074	0.819	0.103	0.824***	0.073	0.868	0.110
Age	0.984***	0.003	0.988***	0.004	0.989***	0.004	0.985***	0.003
Gender	0.665***	0.082	0.686***	0.072	0.679***	0.074	0.675***	0.067
Children	1.130***	0.055	1.091***	0.044	1.129***	0.046	1.150***	0.044
CDD65	1.000	1.05E-04	1.000	6.82E-05	1.000	9.05E-05	1.000	1.11E-04
HDD65	1.000	6.59E-05	1.000	6.20E-05	1.000	6.13E-05	1.000	6.95E-05
House Age	1.131	0.128	0.750	0.165	0.835	0.134	1.056	0.129
Hispanic	1.269*	0.168	1.215	0.162	1.146	0.170	1.129	0.159
Af. American	1.515***	0.239	1.455***	0.209	1.292*	0.211	1.355*	0.217
Bedrooms	1.042	0.053	1.051	0.081	1.049	0.070	1.077	0.076
Bathrooms	0.918	0.103	1.164	0.132	1.188	0.158	1.010	0.102
Census Dummies	Yes		Yes		Yes		Yes	
Housing Dummies	Yes		Yes		Yes		Yes	
Audit	1.538	0.519	-	-	-	-	-	-
Insulation	-	-	0.055***	0.068	-	-	-	-
Windows	-	-	-	-	0.075	0.076	-	-
Energy Star ®	-	-	-	-	-	-	0.689	0.104
Observations	**N = 5,686**		**N = 5,686**		**N = 5,686**		**N = 5,686**	

$* p < 0.10,\ ** p < 0.05,\ *** p < 0.01$

Based on the results in Table 4.11, for a one-unit increase in *Age* (i.e., the respondent for the household is one year older) the odds of being in a higher energy insecurity category than the household is in currently are 0.98 times greater, assuming all else equal. Likewise, the odds of a household being in a more sever energy insecurity category than they are in currently are between 0.72 and 0.74 higher for respondents who are male. The odds of a household being in a higher energy insecurity category than they are in currently are 1.2 as the number of children inside the household increases by one. The odds of a household being in a higher energy insecurity category than the category they are in currently increase between 1.6 and 1.8 for households who identify as African American, and between 1.2 and 1.3 for households who identify as Hispanic. The results outlined above are consistent with the odds ratio results in Table 4.13 when the threshold value of $\tau =$ 3.08 and in Table 4.14 when only four energy security groups are identified (Four Group model).

To examine how improvements in energy efficiency affect household energy security, we focus on the estimated coefficients of four primary variables across the different ordered logit model results: Audit, *Energy Star* ®, Windows, and Insulation. Consistent with our theoretical expectations, our results in Tables 4.8, 4.9, and 4.10 all indicate that households who reported having adequate levels of insulation (Insulation = 1) were statistically less likely to identify as being more energy insecure, all else equal. Improving home insulation is one of the key ways to reduce heat and cooling loss through air leaks in walls, ceilings, and floors. By preventing waste heat, insulation helps to reduce home fuel consumption thereby making energy services more affordable.

Households who consume less fuel to produce energy services can produce energy services more efficiently (operate closer to their efficient production frontier) assuming all other inputs remain fixed and therefore, are able to achieve a higher level of energy security. Based on the ordered logit model results, we conclude having adequate insulation decreases log-odds (see Tables 4.8, 4.9, and 4.10) that a household will become more energy insecure. The odd ratio (see Tables 4.11, 4.12, and 4.13) of the household shifting to a higher energy insecurity category are estimated to be between 0.05 and 0.08.

A similar result was observed for the windows variable which showed that the log-odds that households, who had either double or triple pane windows (Windows = 1) were at a higher energy insecurity category than they are in currently decreased between 2.3 and 2.8. Similar to having adequate insulation, having triple or double pane windows helps to prevent heat and cooling loss. The space between the window panes in triple and double windows is typically filled with either argon or krypton gas (Energy Guard 2017). These gases help prevent condensation from building up outside of the windows, as well as help to reduce drafts during the cold winter and hot summer months (Energy Guard 2017).

Every household included in our sample had a least one *Energy Star®* certified appliance, and many had more than one. Estimation results for the continuous variable measuring the number of *Energy Star®* appliances in a household indicated that as the number of *Energy STAR ®* certified appliances increases in the household, the log-odds the household becomes more energy insecure (i.e., is at a higher energy insecurity category than they are currently) decreases significantly. *Energy Star ®*certified appliances consume fewer units of energy to produce energy service outputs (e.g., clean loads of laundry), leading to greater feelings of energy security. Inconsistent with our expectations, we fail to reject the null hypothesis that households who received an in home energy audits (Audit=1) would place in a lower energy insecurity category.

Table 4.14 Ordered Logit Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star*® Appliances on Household Energy Insecurity with Threshold Value of $\tau = 0$

	$\mathrm{P}(EISINDEX = 0)$	$\mathrm{P}(EISINDEX = 3.08)$	$\mathrm{P}(EISINDEX = 4.38)$	$\mathrm{P}(EISINDEX = 4.44)$
$Audit = 0$	0.830	0.107	0.036	0.016
$Audit = 1$	0.761	0.146	0.053	0.024
$Insulation = 0$	0.370	0.280	0.173	0.101
$Insulation = 1$	0.882	0.077	0.024	0.010
$Windows = 0$	0.567	0.234	0.107	0.054
$Windows = 1$	0.913	0.057	0.018	0.008
$EnergyStar = 0$	0.690	0.182	0.071	0.034
$EnergyStar = 1$	0.753	0.150	0.055	0.025
$EnergyStar = 2$	0.807	0.120	0.041	0.019
$EnergyStar = 3$	0.852	0.094	0.031	0.014
$EnergyStar = 4$	0.887	0.073	0.023	0.010
$EnergyStar = 5$	0.915	0.055	0.017	0.007
$EnergyStar = 6$	0.937	0.042	0.013	0.005
$EnergyStar = 7$	0.953	0.031	0.009	0.004
	$\mathrm{P}(EISINDEX = 4.48)$	$\mathrm{P}(EISINDEX = 5.65)$	$\mathrm{P}(EISINDEX = 6.20)$	$\mathrm{P}(EISINDEX = 6.34)$
$Audit = 0$	0.008	0.002	0.0004	-
$Audit = 1$	0.012	0.003	0.0006	-
$Insulation = 0$	0.056	0.017	0.003	-
$Insulation = 1$	0.005	0.001	0.0002	-
$Windows = 0$	0.028	0.008	0.001	-
$Windows = 1$	0.004	0.001	0.0001	-
$EnergyStar = 0$	0.017	0.005	0.001	-
$EnergyStar = 1$	0.012	0.003	0.001	-
$EnergyStar = 2$	0.009	0.003	0.000	-
$EnergyStar = 3$	0.007	0.002	0.000	-
$EnergyStar = 4$	0.005	0.001	0.000	-
$EnergyStar = 5$	0.004	0.001	0.000	-
$EnergyStar = 6$	0.003	0.001	0.000	-

$EnergyStar = 7$	0.002	0.001	0.000	-

Table 4.15 Ordered Logit Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star*® Appliances on Household Energy Insecurity with Threshold Value of $\tau = 3.08$

	P($EISINDEX = 0$)	P($EISINDEX = 1.2$)	P($EISINDEX = 1.36$)
$Audit = 0$	0.939	0.035	0.015
$Audit = 1$	0.910	0.051	0.023
$Insulation = 0$	0.650	0.174	0.101
$Insulation = 1$	0.961	0.023	0.010
$Windows = 0$	0.780	0.117	0.061
$Windows = 1$	0.974	0.015	0.007
$EnergyStar = 0$	0.873	0.071	0.034
$EnergyStar = 1$	0.905	0.054	0.025
$EnergyStar = 2$	0.930	0.040	0.018
$EnergyStar = 3$	0.948	0.030	0.013
$EnergyStar = 4$	0.962	0.022	0.010
$EnergyStar = 5$	0.972	0.016	0.007
$EnergyStar = 6$	0.980	0.012	0.005
$EnergyStar = 7$	0.985	0.009	0.004
	P($EISINDEX = 1.40$)	P($EISINDEX = 2.57$)	P($EISINDEX = 3.12$)
$Audit = 0$	0.008	0.002	0.0004
$Audit = 1$	0.011	0.003	0.0006
$Insulation = 0$	0.056	0.016	0.003
$Insulation = 1$	0.005	0.001	0.0002
$Windows = 0$	0.032	0.009	0.002
$Windows = 1$	0.003	0.001	0.0002
$EnergyStar = 0$	0.017	0.005	0.0009
$EnergyStar = 1$	0.012	0.003	0.0006
$EnergyStar = 2$	0.009	0.002	0.0004
$EnergyStar = 3$	0.006	0.002	0.0003
$EnergyStar = 4$	0.005	0.001	0.0002

$EnergyStar = 5$	0.003	0.001	0.0002
$EnergyStar = 6$	0.002	0.001	0.0001
$EnergyStar = 7$	0.002	0.005	0.0001

Table 4.16 Ordered Logit Model Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star*® Appliances on Household Energy Insecurity with Four Energy Security Groups

	$EISINDEX = 1$	$EISINDEX = 2$	$EISINDEX = 3$	$EISINDEX = 4$
$Audit = 0$	0.939	0.058	0.002	0.0004
$Audit = 1$	0.910	0.086	0.003	0.0006
$Windows = 0$	0.753	0.235	0.011	0.002
$Windows = 1$	0.976	0.023	0.001	0.0001
$Insulation = 0$	0.592	0.384	0.021	0.004
$Insulation = 1$	0.963	0.035	0.001	0.0002
$EnergyStar = 0$	0.861	0.133	0.005	0.0010
$EnergyStar = 1$	0.900	0.096	0.004	0.0007
$EnergyStar = 2$	0.929	0.068	0.003	0.0005
$EnergyStar = 3$	0.950	0.048	0.002	0.0003
$EnergyStar = 4$	0.965	0.034	0.001	0.0002
$EnergyStar = 5$	0.975	0.024	0.001	0.0001
$EnergyStar = 6$	0.983	0.016	0.001	0.0001
$EnergyStar = 7$	0.988	0.011	0.0004	0.0001

Tables 4.14, 4.15, and 4.16 above provide the predicted probability that a household will be in each of the different energy insecurity categories we construct using the three methods of partitioning, depending upon whether or not a household has made an energy efficiency upgrade or not.

Because none of the households in our sample responded affirmatively to all seven questions included to create the index, the predicted probabilities for $EISINDEX = 6.34$ and $EISINDEX = 3.26$ are not able to be calculated, as the averages for the other variables considered do not exist in the dataset.

When the threshold value of τ is set equal to zero and seven different energy insecurity categories are considered, the predicted probability that a household who has adequate insulation identifies as energy secure ($EISINDEX = 0$) is 0.882. The predicted probability of being in a higher energy insecurity category decreases if households have adequate insulation (Insulation=1). While the predicted probability of identifying as be more energy insecure also decreases for households who have inadequate insulation (Insulation=0) the predicted probability of being energy insecure (i.e., receiving an energy insecurity index value not equal to zero) is higher across all categories for households who have inadequate insulation as opposed to adequate insulation.

As the number of *Energy Star* ® certified appliances increases, so too does the probability that a household will identify as energy secure. Conversely, having more *Energy Star* ® certified appliances (e.g., having seven *Energy Star* ® certified appliances instead of only one) decreases the predicted probability that a household will be in a more severe energy insecurity category. The effects of *Energy Star* ® certified appliances are consistent across the seven different energy insecurity categories. Overall, we find the predicted probability that a household identifies as being energy insecure to some extent, which is represented by the index value it receives, decreases between 10-20% for each additional *Energy Star*® certified appliance added.

The predicted probability of being Energy Secure is higher for households who have triple or double pane windows than households who do not. Consistent with our results for *Energy Star*® certified appliances and having adequate insulation, using only seven energy insecurity index values, we found the predicted probability that a household identifies as energy insecure is 0.97 if the household has triple or double pane windows and 0.78 if the household does not. The predicted probability of receiving an energy insecurity index value of 1.2 is only 0.015 if the household has triple or double pane windows. The predicted probability of receiving an energy insecurity index value of 1.2 is close to 0.12 if the household does not have triple or double pane windows. Similar results for predicted probabilities are found for households who have adequate insulation and operate *Energy Star* ® certified appliances in the home.

When four energy insecurity categories are considered, we find sharp decreases in the predicted probability that household will identify as energy insecure if the household is more energy efficient. For example, the predicted probability that a household will identify as "High Energy Secure" ($EISINDEX = 1$) in the case when only four energy insecurity groups are identified increases if the household has more *Energy Star*® certified appliances. The predicted probability that a household identifies as "Marginally Energy Insecure" ($EISINDEX = 2$) in the case when only four energy insecurity groups are identified) is less if the household has adequate insulation.

4.6 CONCLUSIONS

In this study, we examined the theoretical and empirical relationships between household energy efficiency and energy insecurity. Our theoretical model, which is based on the theory of household production, depicts households as both consumers and producers of energy services. Households produce energy services such as hot, cooked foods by combining fuel inputs (e.g., electricity) with capital technology inputs (e.g., a stove). The capital technology employed by the household is assumed to have a corresponding efficiency level rating, such that higher efficiency level ratings are associated with more energy service outputs being produced for a given fuel input being consumed.

Assuming the household is operating as technically inefficient in its production of household energy services, we theoretically model the household's decision to adopt an energy efficient appliance or make an energy efficiency upgrade. We examine the impact of this decision as it relates to the household's overall level of utility, which is assumed to reflect its subjective "felt level" of energy security. To measure energy security, we use three different index-measures created using results from the application of a Dichotomous Rasch model. The Dichotomous Rasch model allows us to assign households to different energy insecurity groups, based on the number of questions the household responds affirmatively to related to energy-service related hardships they may have faced over the past twelve months. Results from the Dichotomous Rasch model indicate that about 22% of the households living in the United States identified as being energy insecure to some extent in 2015.

Based on our theoretical model, and the results of our index, we empirically explore the relationship between making energy efficiency upgrades in the home and household energy security. Our main empirical analysis results are based on two approaches: 1) a CF logistic regression model and corresponding Bivariate Probit model and 2) an Ordered Logistic model specification. In these models, a household's current stock of energy efficient capital is represented by four primary variables: (1) having adequate insulation (2) installation of double- or triple-pane windows, (3) the number of *Energy Star*® rated appliances in the household, and (4) receiving an in-home energy audit.

Our estimation results indicate that three of these four variables have a negative and statistically significant relationship with a household's self-reported level of energy insecurity.

The implication of these results, consistent with our theoretical model, is that energy efficient capital inputs enable household to "produce" more energy services resulting in higher "felt levels" of energy security, *ceteris paribus*. Across all model results, we find households with higher incomes are statistically less likely to self-identify as energy insecure.

A limitation of our analysis is incorporation of household energy prices. Electricity prices vary greatly by state, with states in the Northeast United States paying almost three times what those in the Southeast pay. We were not able to explore the impact of prices on energy security due to the lack information in the RECS data set on energy prices faced by households. Future research should attempt to incorporate energy prices into energy security/insecurity models. In addition, future research could explore the impacts of the different fuel sources used to produce energy services and determine how each influences household energy security.

Also, it would be interesting to compare energy insecurity with other measures of overall insecurity or general poverty reported by the household. For example, future research could explore whether or not SNAP beneficiaries are more likely to identify as being energy insecure. Future research could also explore the impact of energy efficiency on home energy security specifically in low-income populations. Lastly, as in the food security literature, all of our energy insecurity indices are categorical in nature. Future research should explore developing more continuous measures of household energy security/insecurity.

APPENDIX CHAPTER 4

Table 4. B1 CF Logistic Model Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity

	Logit			
	Change in the Log-Odds		**Change in the Odds**	
	Coefficient Estimate	**St. Error**	**Odds Ratio Estimate**	**St. Error**
Income	-1.77E-05***	1.26E-06	1.00***	0.086
Employment	0.030	0.084	1.030	0.064
Education	-0.229***	0.080	0.796***	0.003
Age	-0.017***	0.003	0.984***	0.054
Gender	-0.315***	0.074	0.730***	0.044
Children	0.188***	0.037	1.207***	0.000
CDD65	1.92E-05	7.18E-05	1.000	0.000
HDD65	-5.56E-05	4.86E-05	1.000	0.089
House Age	-0.004	0.089	0.996	0.137
Hispanic	0.262***	0.105	1.299***	0.203
African American	0.645***	0.107	1.907***	0.050
Bedrooms	0.034	0.048	1.035	0.067
Bathrooms	-0.044	0.070	0.957	0.086
Census Dummies	Yes	Yes	Yes	Yes
Housing Dummies	Yes	Yes	Yes	Yes
Audit	0.334**	0.128	1.396***	0.179
Energy Star ®	-0.007	0.018	0.993	0.018
Windows	-0.096	0.077	0.908	0.070
Insulation	-0.815***	0.086	0.442	0.038
Constant	9.889***	4.579		
Log-Likelihood	-2504.50			
Pseudo R^2	0.1616			
Observations	5,686			

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4. B2 CF Logistic Model Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star*® Appliances on Household Energy Insecurity

	$\boldsymbol{Pr(EISINDEX = 1)}$	**St. Error**
$Audit = 1$	0.168	0.006
$Audit = 0$	0.219	0.021
$Windows = 1$	0.180	0.009
$Windows = 0$	0.166	0.007
$Insulation = 1$	0.289	0.016
$Insulation = 0$	0.152	0.006
$EnergyStar = 0$	0.174	0.009

$EnergyStar = 1$	0.173	0.007
$EnergyStar = 2$	0.172	0.006
$EnergyStar = 3$	0.171	0.006
$EnergyStar = 4$	0.170	0.007
$EnergyStar = 5$	0.169	0.009
$EnergyStar = 6$	0.168	0.011
$EnergyStar = 7$	0.166	0.013

Table 4. B3 Ordered Logit Model Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with $\tau = 0$

	Ordered Logit $\boldsymbol{\tau = 0}$			
	Log-Odds	**Standard Error**	**Odds Ratio**	**Standard Error**
Income	-1.79E-05	1.25E-06	1.000	1.25E-06
Employment	0.012	0.080	1.012	0.081
Education	-0.218	0.077	0.804	0.062
Age	-0.016	0.003	0.984	0.002
Gender	-0.318	0.072	0.728	0.053
Children	0.176	0.034	1.192	0.041
CDD65	2.68E-05	6.91E-05	1.000	6.91E-05
HDD65	-4.37E-05	4.69E-05	1.000	4.69E-05
House Age	-0.015	0.086	0.985	0.085
Hispanic	0.255	0.101	1.290	0.130
African American	0.609	0.101	1.838	0.186
Bedrooms	0.042	0.046	1.043	0.048
Bathrooms	-0.033	0.069	0.968	0.067
Census Dummies	Yes		Yes	
Housing Dummies	Yes		Yes	
Audit	0.351	0.124	1.420	0.176
Energy Star ®	-0.008	0.018	0.992	0.018
Windows	-0.089	0.075	0.915	0.068
Insulation	-0.888	0.082	0.412	0.034
Thresholds				
$\widehat{\omega}_{3.08}$	-2.441	0.444		
$\widehat{\omega}_{4.38}$	-1.292	0.444		
$\widehat{\omega}_{4.44}$	-0.374	0.445		
$\widehat{\omega}_{4.48}$	0.591	0.450		
$\widehat{\omega}_{5.65}$	2.009	0.481		
$\widehat{\omega}_{6.20}$	3.894	0.665		
Log-Likelihood	-3.932.73			
Pseudo R^2	0.1160			
Observations	5,686			

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4. B4 Ordered Logit Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star*® Appliances on Household Energy Insecurity with $\tau = 0$

	$EISINDEX = 0$	$EISINDEX = 3.08$	$EISINDEX = 4.38$	$EISINDEX = 4.44$
$Audit = 0$	0.832	0.108	0.035	0.015
$Audit = 1$	0.777	0.140	0.048	0.021
$Windows = 0$	0.820	0.115	0.038	0.017
$Windows = 1$	0.833	0.107	0.035	0.015
$Insulation = 0$	0.698	0.181	0.069	0.031
$Insulation = 1$	0.849	0.097	0.031	0.014
$EnergyStar = 0$	0.8253	0.112	0.037	0.016
$EnergyStar = 1$	0.8265	0.111	0.036	0.016
$EnergyStar = 2$	0.8277	0.110	0.036	0.016
$EnergyStar = 3$	0.8289	0.110	0.036	0.016
$EnergyStar = 4$	0.8301	0.109	0.036	0.015
$EnergyStar = 5$	0.8313	0.108	0.035	0.015
$EnergyStar = 6$	0.8325	0.108	0.035	0.015
$EnergyStar = 7$	0.8337	0.107	0.035	0.015
	$EISINDEX = 4.48$	$EISINDEX = 5.65$	$EISINDEX = 6.20$	$EISINDEX = 6.34$
$Audit = 0$	0.007	0.002	0.0001	-
$Audit = 1$	0.010	0.003	0.001	-
$Windows = 0$	0.008	0.002	0.0004	-
$Windows = 1$	0.007	0.002	0.0004	-
$Insulation = 0$	0.015	0.004	0.0007	-
$Insulation = 1$	0.006	0.002	0.0003	-
$EnergyStar = 0$	0.008	0.002	0.0004	-
$EnergyStar = 1$	0.008	0.002	0.0004	-
$EnergyStar = 2$	0.008	0.002	0.0004	-
$EnergyStar = 3$	0.007	0.002	0.0004	-
$EnergyStar = 4$	0.007	0.002	0.0004	-
$EnergyStar = 5$	0.007	0.002	0.0004	-
$EnergyStar = 6$	0.007	0.002	0.0004	-
$EnergyStar = 7$	0.007	0.002	0.0004	-

Table 4. B5 Ordered Logit Model Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with $\tau = 3.08$

	Ordered Logit $\tau = 3.08$			
	Log-Odds	**Standard Error**	**Odds Ratio**	**Standard Error**
Income	-1.95E-05	1.91E-06	1.000	1.91E-06
Employment	-0.029	0.110	0.971	0.107
Education	-0.207	0.107	0.813	0.087
Age	-0.015	0.003	0.985	0.003
Gender	-0.392	0.105	0.675	0.071
Children	0.117	0.046	1.124	0.052
CDD65	3.50E-05	9.69E-05	1.000	9.69E-05
HDD65	-6.30E-06	6.48E-05	1.000	6.48E-05
House Age	-0.043	0.122	0.958	0.117
Hispanic	0.205	0.140	1.228	0.172
African American	0.441	0.137	1.555	0.214
Bedrooms	0.055	0.065	1.056	0.068
Bathrooms	0.037	0.098	1.038	0.102
Census Dummies	Yes		Yes	
Housing Dummies	Yes		Yes	
Audit	0.438	0.168	1.550	0.260
Energy Star ®	-0.017	0.026	0.983	0.025
Windows	-0.054	0.105	0.947	0.099
Insulation	-1.098	0.107	0.334	0.036
Thresholds				
$\widehat{\omega}_{4.38}$	-1.429	0.606		
$\widehat{\omega}_{4.44}$	-0.502	0.607		
$\widehat{\omega}_{4.48}$	0.468	0.610		
$\widehat{\omega}_{5.65}$	1.888	0.633		
$\widehat{\omega}_{6.20}$	3.774	0.782		
Log-Likelihood	-2,091.17			
Pseudo R^2	0.126			
Observations	5,686			

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4. B6 Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star*® Appliances on Household Energy Insecurity with Energy Insecurity Index with $\tau = 3.08$

	$EISINDEX = 0$	$EISINDEX = 1.2$	$EISINDEX = 1.36$
$Audit = 0$	0.943	0.034	0.014
$Audit = 1$	0.915	0.050	0.022
$Windows = 0$	0.939	0.036	0.015
$Windows = 1$	0.942	0.034	0.015
$Insulation = 0$	0.866	0.076	0.035
$Insulation = 1$	0.951	0.029	0.012
$EnergyStar = 0$	0.939	0.036	0.015
$EnergyStar = 1$	0.940	0.035	0.015
$EnergyStar = 2$	0.941	0.035	0.015
$EnergyStar = 3$	0.942	0.034	0.015
$EnergyStar = 4$	0.943	0.034	0.014
$EnergyStar = 5$	0.944	0.033	0.014
$EnergyStar = 6$	0.945	0.033	0.014
$EnergyStar = 7$	0.946	0.032	0.014
	$EISINDEX = 1.40$	$EISINDEX = 2.57$	$EISINDEX = 3.12$
$Audit = 0$	0.007	0.002	0.0003
$Audit = 1$	0.010	0.003	0.001
$Windows = 0$	0.007	0.002	0.0004
$Windows = 1$	0.007	0.002	0.0003
$Insulation = 0$	0.017	0.005	0.0008
$Insulation = 1$	0.001	0.002	0.0002
$EnergyStar = 0$	0.007	0.002	0.0004
$EnergyStar = 1$	0.007	0.002	0.0004
$EnergyStar = 2$	0.007	0.002	0.0003
$EnergyStar = 3$	0.007	0.002	0.0003
$EnergyStar = 4$	0.007	0.002	0.0003
$EnergyStar = 5$	0.007	0.002	0.0003
$EnergyStar = 6$	0.007	0.002	0.0003
$EnergyStar = 7$	0.006	0.002	0.0003

Table 4. B7 Ordered Logit Model Results of Effects of Energy Audits, Insulation, Windows, *Energy Star*® Appliances, and other Explanatory Variables on Household Energy Insecurity with Four Energy Insecurity Groups

	Ordered Logit			
	Log-Odds	**Standard Error**	**Odds Ratio**	**Standard Error**
Income	-1.94E-05	1.92E-06	1.000	1.92E-06
Employment	-0.020	0.112	0.981	0.109
Education	-0.217	0.108	0.805	0.087

Age	-0.015	0.004	0.985	0.003
Gender	-0.399	0.105	0.671	0.071
Children	0.111	0.047	1.117	0.053
CDD65	4.87E-05	9.84E-05	1.000	9.84E-05
HDD65	1.56E-06	6.60E-05	1.000	6.60E-05
House Age	-0.052	0.123	0.949	0.117
Hispanic	0.212	0.141	1.236	0.174
African American	0.392	0.139	1.480	0.206
Bedrooms	0.048	0.065	1.049	0.068
Bathrooms	0.037	0.099	1.038	0.102
Census Dummies	Yes		Yes	
Housing Dummies	Yes		Yes	
Audit	0.4477	0.1693	1.565	0.265
Energy Star ®	-0.0164	0.0261	0.984	0.026
Windows	-0.0475	0.1058	0.954	0.101
Insulation	-1.1007	0.1079	0.333	0.036
Thresholds				
$\widehat{\omega}_1$	-1.377	0.615		
$\widehat{\omega}_2$	1.937	0.641		
$\widehat{\omega}_3$	3.824	0.789		
Log-Likelihood	-1,601.84			
Observations	5,686			

$* \, p < 0.10, \; ** \, p < 0.05, \; *** \, p < 0.01$

Table 4. B8 Ordered Logit Predicted Probabilities of Energy Audits, Insulation, Windows, *Energy Star®* Appliances on Household Energy Insecurity with Four Energy Insecurity Groups

	$EISINDEX = 1$	$EISINDEX = 2$	$EISINDEX = 3$	$EISINDEX = 4$
$Audit = 0$	0.943	0.054	0.002	0.0003
$Audit = 1$	0.914	0.082	0.002	0.0005
$Windows = 0$	0.940	0.058	0.002	0.0004
$Windows = 1$	0.942	0.055	0.002	0.0003
$Insulation = 0$	0.866	0.128	0.005	0.001
$Insulation = 1$	0.951	0.047	0.002	0.0001
$EnergyStar = 0$	0.939	0.058	0.002	0.0004
$EnergyStar = 1$	0.940	0.058	0.002	0.0004
$EnergyStar = 2$	0.941	0.057	0.002	0.0003
$EnergyStar = 3$	0.942	0.056	0.002	0.0003
$EnergyStar = 4$	0.943	0.055	0.002	0.0003
$EnergyStar = 5$	0.944	0.054	0.002	0.0003
$EnergyStar = 6$	0.945	0.053	0.002	0.0003
$EnergyStar = 7$	0.945	0.052	0.002	0.0003

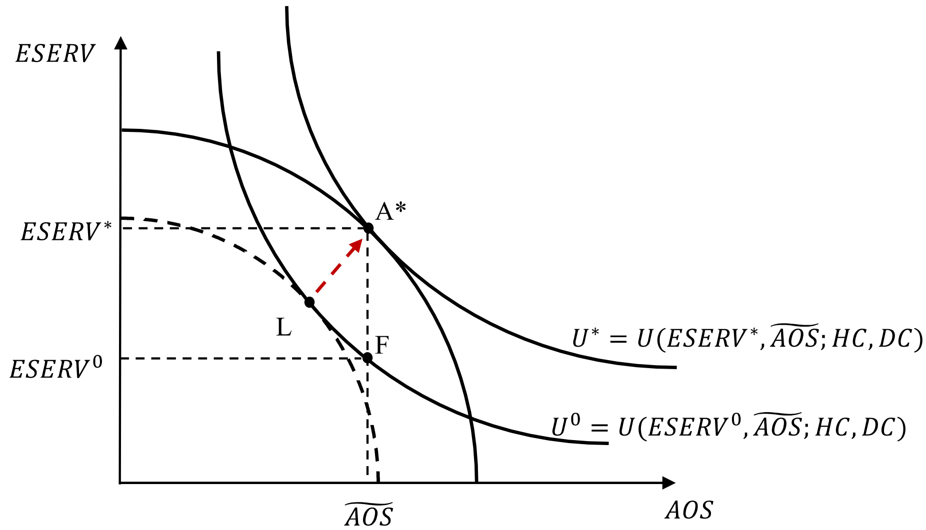
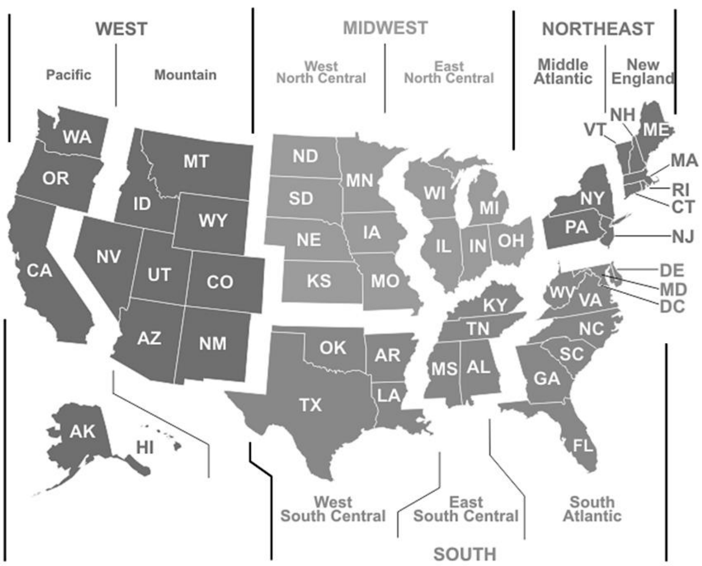


Figure 4.1 B. The Effect of Adopting Energy Efficient Capital Technology Inputs

For our analysis, we assume one reason a household is inefficient in its production of energy services is because the efficiency rating of the capital technology inputs it employs is too low. Following the stochastic production frontier approach, the extent of inefficiency experienced by the household (η_i) is assumed to be a function of things within the control of the household and things outside of the control of the household. The choice of capital technology inputs is within the control of the household and therefore can directly contribute to the extent of inefficiency experienced (Battese and Coelli 1995). By operating capital technology inputs with lower efficiency ratings, relative to the cost-minimizing household (i.e., the household who operates capital technology inputs with an efficiency rating $= \gamma^*$) the household is inefficient in its production of energy services and as a result the household is not able to operate along its PPF.

In presentations of this work, questions have been raised about why the decision to adopt capital technology inputs with higher efficiency ratings does not shift the households PPF for services outward, which is represented graphically in Figure 4.1 B. as a move from point L to point A*. Recall that the PPF represents the maximum possible output (i.e., combinations of two goods or services) one can achieve when all the resources one has available are fully and efficiently employed. A shift in the PPF would represent the adoption of more or additional capital technology inputs which would allow the household to produce more services all together. For example, if the household chose to purchase a second washer and dryer, it would be able to produce more loads of clean laundry.

Conversely, by replacing a less efficient washer with a more efficient model (e.g., an *Energy Star* ® certified model) the household is using the same number of capital inputs but producing energy services more efficiently by decreasing the amount of energy necessary to produce energy services. This increase in efficiency is represented graphically by a move toward the PPF rather than a shift in the PPF, because relative to the cost-minimizing household the inefficient household should be able to produce more energy services than it is able to when it is inefficient in its production.



Source: https://www.eia.gov/analysis/requests/powerplants/cleanplan/

Figure 4.2 B. United States Census Divisions

Log-Likelihood Functions for Models in Chapter 4 (Essay 3)

Logit Model:

(B.1) $\log L(\beta) = \sum \left[EISINDEX_i \log\left(\frac{\exp(x_i\beta_k)}{1+\exp(x_i\beta_k)}\right) + (1 - EISINDEX_i)\log\left(1 - \left(\frac{\exp(x_i\beta_k)}{1+\exp(x_i\beta_k)}\right)\right)\right]$

Bivariate Probit Model:

(B.2) $\log L(\beta) = \sum \log \Phi(t_{1i}, t_{2i}, \rho_i^*)$

Here for $i = 1, .., N$, $t_{1i} = (2EISINDEX_i - 1)(x_i\beta_k + EE_i\delta)$, $t_{1i} = (2EE_i - 1)(x_i\beta_k + z_i\alpha)$,

and $\rho_i^* = (2EISINDEX_i - 1)(2EE_i - 1)\rho$.

Ordered Logit Model:

(B.3) $\log L(\omega, \beta) = 1[\,EISINDEX_i = 0] \log\left(\omega_{3.08} - \frac{\exp(x_i\beta_k)}{1+\exp(x_i\beta_k)}\right) + 1\,[EISINDEX_i =$

$3.08] \log\left(\omega_{4.34} - \frac{\exp(x_i\beta_k)}{1+\exp(x_i\beta_k)}\right) + \cdots + 1[EISINDEX_i = 6.34] \log\left(1 - \omega_{6.34} - \frac{\exp(x_i\beta_k)}{1+\exp(x_i\beta_k)}\right)$

Note: The thresholds ω and index values $EISINDEX_i$ change depending on which index is being modeled with the ordered logit

CHAPTER 5

GENERAL CONCLUSIONS

5.1 ARE THE ENERGY CHANGES WE EXAMINE EFFICIENT AND EQUITABLE?

This dissertation examines how changing the way we produce and consume energy across the United States has impacted both consumers and producers of energy resources. We focus on two specific changes in the way energy is now produced and consumed. First, we consider how changing the energy resource mix used to generate electricity to include a greater share of intermittent renewable resources can impact the efficiency with which energy is produced and the reliability of energy delivery to end-consumers. Second, we consider how investments in energy efficiency can influence how "secure" a household feels about their ability to maintain adequate access to household energy services.

Intermittent renewable resources and home energy efficiency investments are potential solutions to ensuring an ample supply of energy remains available for consumption both now and in the future, energy remains affordable for both firms and consumers, and the environmental impacts associated with energy production and consumption are kept to a minimum (Yacoucci 2016). However, despite providing solutions to pressing contemporary problems and issues related to energy production and consumption, the relationships between intermittent renewable resources, home energy efficiency investments, and the economic efficiency and distributional equity of energy production and consumption have yet to be adequately addressed in the applied economics literature.

Modifying the energy resource mix to include a greater share of intermittent renewable resources provides an opportunity for the United States to become self-sustaining in terms of supplying its own energy (Flavin and Dunn 1999; Johansson 2013). However, evidence from Chapter 2 (Essay 1) indicates that as the capacity of electricity generated by intermittent renewables increases, consumers can expect to experience longer power outages. Predicted power outages over time are expected to be longer as the percentage of the capacity of electricity generated by intermittent renewable resources increases.

These power outages may cause losses in economic efficiency to both power generating utilities and other firms that rely on reliable power for efficient production of outputs. In addition, our results suggest scaling-up effects may result in relatively large, non-marginal increases in the economic costs of predicted outages. These increased costs including increased power costs faced by both firms and consumers have distributional equity implications such as reduction in the affordability of energy to lower-income and other economically vulnerable households.

In particular, it is important to consider the fact that the effects of power system outages are felt well beyond just the lights going out. When a household or a business goes without electricity they face both real out-of-pocket costs and opportunity costs. For the household these costs may include losses in well-being (welfare) due to an inability to access the inputs necessary to produce cooked meals or properly heat and/or cool the home, purchasing more away-from-home meals during outages, and alternative lodging costs (e.g., hotel costs). For a business these costs may include costs of back-up generators, costs of alternative heating sources in the winter (e.g., propane or kerosene heaters), and opportunity costs in the form of lost revenues when production operations are shut-down. Furthermore, power system outages can result in serious consequences for companies in charge of critical infrastructures such as financial services, water supplies, telecommunication services, hospitals, and other emergency medical centers (Byrd and Mattewmen 2014). From a distributional equity perspective, lower-income, elderly, and other vulnerable households may especially suffer from loss of critical infrastructure such as medical facilities when the power goes out.

The potential issues and problems discussed above beg the question: Do the environmental benefits (i.e., reduced CO_2 emissions) and sufficiency of supply gains from increased production from intermittent renewable resources lead to outcomes that are both economically efficient and distributionally equitable? Unless the gainers from increased production from intermittent renewable resources could compensate the losers (i.e., those affected by and who face costs from experiencing power system outages) and still be better off, policies that call for increased production from intermittent renewables cannot be considered a Pareto Improvement (or Potential Pareto Improvement if compensation is not actually paid). Given that some individuals are being made worse off (in terms of facing costs when the power goes out) as a result of policies such as Renewable Energy Standards, the equity implications of transitioning our electricity sector to depend more on intermittent renewable resources should receive more attention and research.

In addition to increasing the capacity of electricity generated by intermittent renewables, investments in home energy efficiency have also been identified as a potential solution to meet the three main energy policy goals of the United States as outlined by Yacoucci (2016). As stated in Chapter 4, home energy efficiency investments are designed to increase the efficiency with which a household is able to produce and consume household energy services. By using fewer energy/fuel inputs to provide energy services, these investments lower home energy costs. In addition, they reduce the GHG emissions associated with energy production by lowering overall energy consumption (Fowlie et al. 2018).

Another major potential benefit of home energy efficiency investments is the positive effect such investments can have on the energy security of households. To gauge how energy secure/insecure a household feels, in Chapter 3 we compare and contrast five different empirical measures of household energy insecurity (i.e., a lack of household energy security). Our validity testing results suggest using the Dichotomous Rasch model to analyze household responses to the 2015 RECS provides an accurate representation of what it means for a household to feel energy insecure. We therefore use the Rasch model to propose a unique energy insecurity index that can be applied to accurately place households into different energy insecurity categories similar to the categories used by the U.S. Department of Agriculture for food insecurity.

To estimate how investments in home energy efficiency influence a household's energy insecurity status, in Chapter 4 we employ our unique energy insecurity index and empirically examine the relationship between a household's latent level energy insecurity and specific home energy efficiency investments or upgrades including having adequate insulation, energy efficient windows, receiving an in-home energy audit, and *Energy Star* ® certified appliances. Overall, our results suggest that these investments or upgrades, with the exception of home energy audits, lead to decreases in the probability that a household will identify as being energy insecure or be placed in a more severe energy insecurity category than they are in currently.

From an economic efficiency perspective, the theoretical model presented in Chapter 4 suggests that investments in home energy efficiency can enable households to more efficiency "produce" energy services, and as a result become more energy secure leading to increases in overall household utility or satisfaction. The empirical results reported in Chapter 4 supports our hypothesis that home energy efficiency investments lead to greater self-reported, felt-levels of household energy security. Therefore, programs and initiatives that promote such investments should be considered by public policy and decision makers who seek to meet the three objectives of the U.S. energy policy agenda discussed in Chapter 1.

Of the home energy efficiency investments, we examined in Chapter 4, adequate home insulation has the largest impact on the probability that a household will fall into the category of "Energy Secure" rather than "Energy Insecure." However, as discussed in Chapter 1, only 2% of the budget of the LIHEAP is dedicated to helping households add insulation to their homes through weatherization assistance (U.S. Department of Health and Human Services 2018). The remaining funds are spent to provide one-time energy assistance to households who are at risk of having their power shut-off. As pointed out by Salvador (2018), receiving energy assistance to help pay for energy services is only a temporary solution to the overall problem of household energy insecurity Thus, future research should consider whether funds from the LIHEAP are being used and distributed in an economically efficient manner. Furthermore, given the LIHEAP program is the largest federally funded energy assistance program in the United States (U.S. Department of Health and Human Services 2018), our results also suggest more attention and research should also be devoted to the distributional equity implications of providing more funding for one-time home energy assistance versus providing more funds for weatherization assistance to homes.

Finally, we conclude with a brief summary of the major, unique contributions of this dissertation research. First, in addition to being one of the first studies to examine how intermittent renewable resources impact power system reliability, we are the first to utilize a state-contingent production function approach to theoretically model power-system outages. The state-contingent production framework is unique in that it allows us to incorporate the uncertainty that power system operators face when determining which energy resources to bring online to meet demand.

Drawing from the engineering literature on power grid management, we provide an avenue for economists to explore new ways to model operations within the electric utility industry.

We also provide empirical evidence that increasing the capacity of intermittent renewable resources has a positive, statistically significant relationship with power system outages. Results of our forecasting procedure provide evidence that if technology were to remain the same and more and more intermittent renewable resources are brought online, then customers across the United States can expect to experience longer power system outages. Therefore, as public policy initiatives such as RES continue to be proposed calling for increased capacity from renewables, our results strongly suggest the need for power companies and those who regulate power companies to engage in and support research and development aimed at ensuring our power system is reliable.

Second, while previous studies have introduced alternative measures for household energy insecurity, we are the first to provide an in-depth explanation of the cautions and caveats associated with each of the different index measures. We also propose a unique, conceptually and empirical robust measure of energy insecurity based on a Rasch model. Our hope is that by outlining the benefits and drawbacks of each of the different index measures, we can motivate public policy and decision makers to create a set of standard procedures for measuring the extent and depth of energy insecurity across the United States. Furthermore, by showcasing how each index measure identifies energy insecure households differently, we open the door for further research related to refining the way we currently measure household energy insecurity.

Third, we are the first to use the results of an energy insecurity index to examine how investments in energy efficiency impact household energy insecurity. While numerous other studies have focused on why consumers invest in energy efficiency, our study provides empirical evidence on what happens after customers decide to invest. We also provide a detailed theoretical model that justifies why a household might decide to invest in energy efficiency, without assuming the household faces different prices, an avenue that has yet to be adequately addressed in the literature. By focusing specifically on the outcomes of household energy insecurity, we provide a different dimension to the discussion on why energy efficiency investments should be considered by public policy and decision makers who are interested in using energy efficiency to combat energy affordability and household energy insecurity.

REFERENCES

Abeyasekera, S. (2003). Chapter 18: Multivariate methods for index construction. *Household surveys in developing and transition countries: design, implementation and analysis.* Retrieved from https://unstats.un.org/unsd/hhsurveys/finalpublication/ch18fin3.pdf

Abrell, J. and Kunz, F. (2015). Integrating intermittent renewable wind generation-a stochastic multi-market electricity model for the European electricity market. *Networks and Spatial Economics*, 15, 17-147.

Abrell, J. and Rausch, S. (2016). Cross-country electricity trade, renewable energy and European transmission infrastructure policy. *Journal of Environmental Economics and Management*, 79, 87-113.

Allcott, H. and Greenstone, M. (2012). Is there an energy efficiency gap? *Journal of Economic Perspectives, 26*(1), 3-28.

Allcott, H. and Greenstone, M. (2017). Measuring the Welfare Effects of Residential Energy Efficiency Programs, Mimeo.

Ames, A. J., and Penfield, R. D. (2015). An NCME Instructional Module on Item-Fit Statistics for Item Response Theory Models. *Educational Measurement: Issues and Practice, 34*(3), 39-48.

Angrist, J., D., and Pischke, J. (2009). *Mostly harmless econometrics. An empiricist's companion*. Princeton University Press, Princeton, N.J.

Apogee Interactive. (2013). Generation dispatch. Retrieved from http://c03.apogee.net/contentplayer/?coursetype=foe&utilityid=rmec&id=4551

Aravamudhan, N. R., and Krishnaveni, R. (2015). Establishing and reporting content validity evidence of new training and development capacity building scale (TDCBS). *Management, 20*(1), 131–158.

Averch, H. and Johnson, L. L. (1962). Behavior of the firm under regulatory constraint. *The American Economic Review*, *52*(5), 1052-1069.

Bagozzi, R., Yi, Y., and Phillips, L. (1991). Assessing Construct Validity in Organizational Research. *Administrative Science Quarterly*, *36*(3), 421-458.

Balistreri K. S. (2016). A decade of change: Measuring the extent, depth and severity of food insecurity. *Journal of Family and Economic Issues, 37*(3), 373–382.

Battese, G. E., and Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*(2), 325-332.

Beck, C. T., and Gable, R. K. (2001). Further validation of the postpartum depression screening scale. *Nursing Research*, *50*(3), 155-164.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, *75*(229), 493-517.

Becker, G.S. (1981). *Treatise on the family*. Cambridge, MA: Harvard University Press.

Bergstrom, J. C. and Randall, A. (2016). *Resource Economics: An Economic Approach to Natural Resource and Environmental Policy*. 4th ed. Edward Elgar Publishing: Northampton, MA

Bhattacharyya, S. C. (1995). Power sector privatization in developing countries: Will it solve all problems? *Energy Sources*, *17*, 373-389.

Bickel, G., Nord, M., Price, C., Hamilton, W., and Cook, J. (2000). Guide to measuring household food security. Retrieved from https://fns-prod.azureedge.net/sites/default/files/FSGuide.pdf

Bird, L., Chapman, C., Logan, J., Summer, J., and Short, W. (2011). Evaluating renewable portfolio standards and carbon cap scenarios in the U.S. electric sector. *Energy Policy, 39*, 2573-2285.

Bishop, R. C., and Boyle, K. J. (2019). Reliability and validity in nonmarket valuation. *Environmental & Resource Economics, 72*(2), 559-582.

Blumsack, S. (2018). The dispatch of power plants by an electric utility. (Lesson 4: Economic Dispatch of Power Plants, Penn State, College of Earth and mineral Resources). Retrieved from https://www.e-education.psu.edu/ebf483/node/534

Boardman, B. (1991). *Fuel Poverty: From Cold Homes to Affordable Warmth.* Belhaven Press, London.

Bockstael, N. E. and McConnell, K. E. (1983). Welfare measurement in the household production framework. *The American Economic Review, 73*(4), 806-814.

Boogen, N., Datta, S., and Filippini, M. (2014). Going beyond tradition: Estimating residential electricity demand using an appliance index and energy services. CER-ETH Economics working paper series 14/200, CER-ETH - Center of Economic Research (CER-ETH) at ETH Zurich

Burnett, J.W. and Madariaga, J. F. (2017). A top-down economic efficiency analysis of U.S. household energy consumption. *The Energy Journal, forthcoming*.

Burtraw, D., Palmer, K., Paul, A., Beasley, B., and Woerman M. (2013). Reliability in the U.S. electricity industry under new environmental regulations. *Energy Policy, 62*, 1078-1091.

Cameron, A. C., and Trivedi, P. K. (2009). *Microeconometrics using Stata.* Stata press.

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

Campbell, R. (2018). The smart grid: Status and outlook. Congressional Research Service Report. Retrieved from https://fas.org/sgp/crs/misc/R45156.pdf

Carliner, M. (2013). Reducing energy costs in rental housing: The need and the potential. Cambridge, MA: Joint Center for Housing Studies of Harvard University. Retrieved from www.jchs.harvard.edu/sites/jchs.harvard.edu/files/carliner_research_brief_0.pdf.

Chambers, R. G. (1998). Cost functions and duality for stochastic technologies. *American Journal of Agricultural Economics, 80*, 288-295.

Chambers, R. G. and Quiggin, J. (2000). *Uncertainty, Production, Choice, and Agency: The State-Contingent Approach.* Cambridge University Press

Chandler, A. (2016, June 8). Where the poor spend more than 10% of their income on energy. Hint: Almost everywhere in the United States. *The Atlantic.* Retrieved from https://www.theatlantic.com/business/archive/2016/06/energy-poverty-low-income-households/486197/

Chang, L., and Wu, Z. (2011). Performance and reliability of electrical power grids under cascading failures. *International Journal of Electrical Power and Energy Systems, 33,* 1410-1419.

Chen, W. and Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22*(3), 265–289.

Coleman-Jensen, A., Rabbit, M. P., Gregory C. A., and Singh, A. (2018). *Household Food Security in the United States in 2017*, ERR-256, U.S. Department of Agriculture, Economic Research Service

Colton, R. (2003). Measuring the outcomes of low-income energy assistance programs through a home energy insecurity scale. Technical Report, U.S. Department of Health and Human

Services, Administration for Children and Families. Retrieved from https://www.acf.hhs.gov/sites/default/files/ocs/measuring_outcome_0.pdf

Cook, J. T., Frank, D. A., Casey, P. H., Rose-Jacobs, R., Black, M. M., Chilton, M., and ... Cutts, D. B. (2008). A brief indicator of household energy security: associations with food security, child health, and child development in US infants and toddlers. *Pediatrics, 122*(4), 867-875.

Crabtree, G., Misewich, J., Ambrosio, R., Clay, K., DeMartini, P., James, R., ... and Slakey, F. (2011, November). Integrating renewable electricity on the grid. In *AIP Conference Proceedings* (Vol. 1401, No. 1, pp. 387-405). AIP.

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

Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin, 52*(4), 281.

Davis, L. W., Fuchs, A., and Gertler, P. (2014). Cash for coolers: Evaluating a large-scale appliance replacement program in Mexico. *American Economic Journal: Economic Policy*, *6*(4), 207-38.

Dear, K. B., & McMichael, A. J. (2011). The health impacts of cold homes and fuel poverty. *BMJ 342*, 2807.

Deaton, A., and Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge; New York: Cambridge University Press

Delarue, E., and Morris, J. (2015). Renewables intermittency: operational limits and implications for Long-term energy system models. MIT Joint Program on the Science and Policy of Global Change.

DeNavas-Walt, C., and Proctor, B. D. (2014). Income and poverty in the United States: 2013. Retrieved from https://www.census.gov/library/publications/2014/demo/p60-249.html

Department of Energy (DOE). (2010). Weatherization installer/technician mobile homes. Retrieved from https://www.energy.gov/eere/wipo/weatherization-installertechnician-mobile-homes

Department of Energy (DOE). (2015). Energy storage. Retrieved from https://www.energy.gov/

Department of Energy (DOE). (2016). Advanced metering infrastructure and customer systems. Retrieved from https://www.energy.gov/sites/prod/files/2016/

Department of Energy and Climate Change (DECC). (2013). Annual report on fuel poverty statistics 2013. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/199833/Fuel_Poverty_Report_2013_FINALv2.pdf.

Division of Energy Assistance. (2008). LIHEAP Home Energy Notebook for Fiscal Year 2008. Technical Report U.S. Department of Health and Human Services, Administration for Children and Families.

Drehobl, A., and Ross, L. (2016). Lifting the high energy burden in America's largest cities: How energy efficiency can improve low income and underserved communities. ACEEE Report. Retrieved from https://energyefficiencyforall.org/sites/default/files/Lifting%20the%20High%20Energy%20Burden_0.pdf

Dutta, I., and Gundersen, C. (2007). Measures of food insecurity at the household level. In Food Security Indicators, Measurement, and the Impact of Trade Openness: Series: WIDER Studies in Development Economics.

Edison Electrical Institute (EEI). (2018). LIHEAP funding. Retrieved from http://www.eei.org/issuesandpolicy/liheap/Pages/default.aspx

Energy Guard. (2017). Should I get double or triple pane windows? Retrieved from https://www.energyguardwindows.com/get-double-triple-pane-windows/

Energy Information Administration (EIA). (2016). Residential energy consumption survey (RECS). Retrieved from https://www.eia.gov/consumption/residential/background-technical.php

Eto, J. (2003). A framework and review of customer outage costs: Integration and analysis of electric utility outage cost surveys. Retrieved from https://emp.lbl.gov/sites/all/files/lbnl-54365.pdf

Eto, J. H. and LaCommare, K. H. (2008). Tracking the reliability of the U.S. electric power system: An assessment of publicly available information reported to state public utility commissions. Ernest Orlando Lawrence Berkeley National Laboratory

Eto, J. H., K. H. LaCommare, P. Larsen, A. Todd, and Fisher, E. (2012). Distribution-level Electricity Reliability: Temporal Trends Using Statistical Analysis. *Energy Policy, 49,* 243-252.

Everitt, B.S., Landau, S. and Leese, M. (2001) Cluster Analysis. 4th Edition, Arnold, London.

Federal Energy Regulatory Commission (FERC). (2015). Energy Primer, a Handbook of Energy Market Basics. *Federal Energy Regulatory Commission: Washington, DC, USA*

Felder, F.A. (2011). Examining Electricity Price Suppression Due to Renewable Resources and Other Grid Investments. *The Electricity Journal, 24*(4), 34–46.

Feleke, S. T., Kilmer, R. L., and Gladwin, C. H. (2005). Determinants of food security in Southern Ethiopia at the household level. *Agricultural Economics, 33*(3), 351-363.

Fell, M. J. (2017). Energy services: A conceptual review. *Energy Research and Social Sciences, 27,* 129-140.

Fenrick, S. A., and L, Getachew. (2012). Formulating appropriate electric reliability targets and performance evaluations. *The Electricity Journal, 25*(2), 44-53.

Filippini, M. (1999). Swiss residential demand for electricity. *Applied Economics Letters, 6*(8), 533-538.

First Energy Corp. (2018). Generation system. Retrieved from https://www.firstenergycorp.com/content/fecorp/about/generation_system.html

Flavin, C., and Dunn, S. (1999). A new energy paradigm for the 21st century. *Journal of International Affairs, 53*(1), 167-190.

Fowlie M., Greenstone, M., and Wolfram C. (2018). Do energy efficiency investments deliver? evidence from the weatherization assistance program. *The Quarterly Journal of Economics, 133*(3),1597–1644.

Fu, Y., Z. Li and Shahidephpor, M. (2004). Profit-based generation resource planning. IMA *Journal of Management Mathematics, 15,* 273-289.

Georgia Power. (2018). Energy sources. Retrieved from https://www.georgiapower.com/company/energy-industry/energy-sources.html

Gerrity, S. and Lantero, A. (2014). Understanding the grid. Retrieved from https://www.energy.gov/articles

Gillingham, K., D. Rapson, and Wagner, G. (2016). The rebound effect and energy efficiency policy. *Review of Environmental Economics & Policy, 10*(1), 68-88.

Gillingham, K., Kotchen, M., Rapson, D., and Wagner, G. (2013) The rebound effect is over-played. *Nature, 493*, 475-476.

Goulder, L. H. and Parry, I. W. H. (2008). Instrument choice in environmental policy. *Review of Environmental Economics and Policy*, *2*(2), 152-174.

Greene, W. H. (2012). *Econometric Analysis* 7th edition Upper Saddle River, NJ

Gundersen, C. (2008). Measuring the extent, depth, and severity of food insecurity: An application to American Indians in the USA. *Journal of Population Economics 21*, 191-215.

Hahn, H., S. Meyer-Nieberg, and Pickl, S. (2009). Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research 199*(3), 902–907.

Halvorsen, R. (1975). Residential demand for electric energy. *The review of Economics and Statistics, 85*, 12-18.

Hamilton, W. L., & Cook, J. T. (1997). Household food security in the United States in 1995: technical report of the Food Security Measurement Project.

Hanemann, W. M. (1984). Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics 66*(3), 332-341.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica, 46*, 1251-1271.

Hernández D. (2013). Energy insecurity: A framework for understanding energy, the built environment and health among vulnerable groups. [Editorial]. *American Journal of Public Health, 103*(4).

Hernández, D. (2016). Understanding ‘energy insecurity’ and why it matters to health. *Social Science Medicine*, *16*, 1-10.

Hibbard, P., S. Tierney, K. Franklin. (2017). Electricity Markets, Reliability, and the Evolving US Power System. *Analysis Group.* Retrieved from http://info aee.net/electricity-markets-reliability-and-the-evolving-us-power-system-report

Hills, J. (2012). Getting the measure of fuel poverty: Final report of the fuel poverty review. London, Department of Energy and Climate Change (DECC). Retrieved from http://sticerd.lse.ac.uk/dps/case/cr/CASEreport72.pdf

Hines, P., J. Apt, and Talukdar, S. (2009). Large blackouts in North America: Historical trends and policy implications. *Energy Policy, 37*, 5249-5259.

Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy, 39*(2), 137-175.

Huang, C. J., and Liu, J. T. (1994). Estimation of a non-neutral stochastic frontier production function. *Journal of Productivity Analysis, 5*(2), 171-180.

Huffman, W. E. (2011). *Household production theory and models.* The Oxford Handbook of Food Consumption and Policy, 35-74.

Institute for Energy Research (IER). (2013). Germany’s green energy destabilizing electric grid. Retrieved from http://instituteforenergyresearch.org

International Energy Agency [IEA]. (2019). What is energy security? Retrieved from https://www.iea.org/topics/energysecurity/whatisenergysecurity/

Johansson, B. (2013). Security aspects of future renewable energy systems–A short overview. *Energy, 61*, 598-605.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika, 36*, 149-176.

Kapur, N., Hiller, J., Langdon, R., and Abramson, A. (2011, July). *Show me the money. Energy efficiency financing barriers and opportunities*. Environmental Defense Fund. Washington, D.C.

Kilanowski, J. F., & Lin, L. (2012). Rasch analysis of US household food security survey module in Latino migrant farmworkers. *Journal of Hunger & Environmental Nutrition*, *7*(2-3), 178-191.

Kumbhakar, S., Ghosh, S., and McGuckin, J. (1991). A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business & Economic Statistics*, *9*(3), 279-286.

LaFrance, J. T. (2000). *Duality theory for the household.* California Agricultural Experiment Station, Giannini Foundation of Agricultural Economics.

Lalor, J. P., Wu, H., and Yu, H. (2016). Building an Evaluation Scale using Item Response Theory. *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, 648-657.

Larsen, P. H., Boehlert. B., Eto, J., Hamachi-LaCommare, K., Martinich, J., and Rennels, L. (2017). Projecting future costs to U.S. electric utility customers from power interruptions. *Energy*, *147*, 1256-1277.

Larsen, P., Hamachi-LaCommare, K., Eto, J., and Sweeney, J. (2015). A*ssessing Changes in the Reliability of the U.S. Electric Power System*. Berkeley: Lawrence Berkeley National Laboratory.

Larsen, P., LaCommare, K., Eto, J., and Sweeney, J. (2016). Recent trends in power system reliability and implications for evaluating future investments in resiliency. *Energy*, *117,* 29-46.

Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28,* 563–575.

Levinson, A. (2016). How much energy do building energy codes save? Evidence from California Houses, *American Economic Review, 106*, 2867–2894.

Li, J. (2011). Modeling household energy consumption and adoption of energy-efficient technology using recent micro-data. (Unpublished doctoral dissertation). University of Maryland, College Park, MD.

Linn, J., and Richardson, N. (2013). Cutting carbon, take two: A brief guide to federal electricity-sector climate policy without cap-and-trade. *Applied Economic Perspectives and Policy, 35*(3), 377-397.

Malla, K. (2013). The effects of quality standards on electricity service reliability. Illinois State University [Thesis]. Retrieved from https://irps.illinoisstate.edu

Massey, C. (2017). State of the efficiency program industry: Budgets, expenditures, and impacts 2016. *The Consortium for Energy Efficiency.*

Matthewman, S. D., and Byrd, H. (2014). Blackouts: A sociology of electrical power failure. *Social Space*, 1-25.

McLeod, S. (2013). “What is reliability? Retrieved from https://www.simplypsychology.org/reliability.html

Middlemiss, L., & Gillard, R. (2015). Fuel poverty from the bottom-up: Characterizing household energy vulnerability through the lived experience of fuel poor. *Energy Research & Social Science, 6*, 126-154.

Moghaddam, M. M., Javidi, M. H., Moghaddam, M. P., and Buygi, M. O. (2014). Reliability-based generation resource planning in electricity markets. *Turkish Journal of Electrical Engineering and Computer Sciences, 22*, 909-923.

Morris, A. C. (2009). *Equity and Efficiency in Cap-and-Trade: Effectively Managing the Emissions of Allowance Supply*. Energy Security Initiative at Brookings.

Muellbauer, J. (1974). Household production theory, quality, and the "Hedonic Technique". *The American Economic Review*, *6*, 977.

Murray, A. G., and Mills, B. F. (2012). An application of dichotomous and polytomous Rasch models for scoring energy insecurity. *Energy Policy, 51*, 946-956.

Murray, B.C. (2015). Why have carbon markets not delivered agricultural emission reductions in the United States? *Choices*. Quarter 2. Available online: http://choicesmagazine.org/choices-magazine/theme-articles/climate-change-and-agriculture-revisiting-the-evidence-and-potential-solutions/why-have-carbon-markets-not-delivered-agricultural-emission-reductions-in-the-united-states

National Renewable Energy Lab (NREL). (2015). Renewable portfolio standards. Retrieved from https://www.nrel.gov/technical-assistance/basics-portfolio-standards.html

Nikoletatos, J., and Tselepis, S. (2015). Renewable energy integration in power grids: technology brief. IEA-ETSAP and IRENA technology Brief.

Nord, Mark. (2003). Measuring the food security of elderly persons. *Family Economics and Nutrition Review, 15*(1), 33-46.

Nord, M. (2009). Food Insecurity in Households with Children: Prevalence, Severity, and Household Characteristics. Economic Information Bulletin Number 56. *US Department of Agriculture.*

Nord, M. (2014). Introduction to item response theory applied to food security measurement: basic concepts, parameters, and statistics. Technical Paper, FAO, Rome.

North American Electric Reliability Council (NERC). (2013). Frequently asked questions. Retrieved from https://www.nerc.com/AboutNERC

Novan, K. and Smith, A. (2018). The incentive to overinvest in energy efficiency: Evidence from hourly smart-meter data, *Journal of the Association of Environmental and Resource Economists, 5*(3), 577-605.

O'Connor, T. (2018). What this summer's heatwaves tell us about America's electric grid. Environmental Defense Fund: Energy. Retrieved from https://www.edf.org/blog/2018/07/26/what-summers-heatwaves-tell-us-about-americas-electric-grid

O'Meara, G. (2016). A review of the literature on fuel poverty with a focus on Ireland. *Social Indicators Research*, *128*(1), 285-303.

Opsomer, J. D., Jensen, H. H., and Pan, S. (2003). An evaluation of the US Department of Agriculture food security measure with generalized linear mixed models. *The Journal of Nutrition*, *133*(2), 421-427.

Orlando, M., and Thissen, D. (2000). Likelihood- based item-fit indices for dichotomous item response theory models. *Applied psychological measurement*, *24*(1), 50–64.

Ozturk, U. A., Mazumdar, M., and Norman, B. A. (2004). A solution to the stochastic unit commitment problem using chance constrained programming. *IEEE Transactions on Power Systems, 19*(3), 1589-1598.

Palmer, K. and Burtraw, D. (2005). Cost-effectiveness of renewable electricity policies. *Energy Economics,* 27, 873-894.

Palmer, K., Paul, A., Woerman, M., and Steinberg, D. C. (2011). Federal policies for renewable electricity: Impacts and interactions. *Energy Policy*, *39*, 3975-3991.

Pean, E., Pirouti, M., and Qadrdan, M. (2016). Role of the GB-France electricity interconnectors in integration of variable renewable generation. *Renewable Energy, 99*, 307-314.

Perl, L. (2018, June). LIHEAP: Program and funding. Library of Congress, Congressional Research Service. Retrieved from https://fas.org/sgp/crs/misc/RL31865.pdf

Pless, J., and Fell, H. (2017). Bribes, bureaucracies, and blackouts: Towards understanding how corruption at the firm level impacts electricity reliability. *Resource and Energy Economics, 47*, 36-55.

Polit, D.F., Beck, C.T. and Owen, S.V. (2007). Is the CVI an acceptable indicator of content validity? *Appraisal and Recommendations. Research in Nursing & Health, 30*, 459-467.

Popovich, N. (2018, December 24). How does your state make electricity? *The New York Times.* Retrieved from https://www.nytimes.com/interactive/2018/12/24/climate/how-electricity-generation-changed-in-your-state.html

Rathi, A. (2017). Batteries can’t solve the world’s biggest energy-storage problem. One startup has a solution. *Quart.* Retrieved from https://qz.com

Reames, T. G. (2016). Targeting energy justice: Exploring spatial, racial/ethnic and socioeconomic disparities in urban residential heating energy efficiency. *Energy Policy, 97*, 549-558.

Rocha, P., Das, T. K., and Botterud, A. (2015). Impact of CO_2 cap-and-trade programs on restructured power markets with generation capacity investments. *Electrical Power and Energy Systems, 71*, 195-208.

Rodriguez-Alvarez, A., Orea, L., and Jamasb, T. (2016). Fuel poverty and well-being: a consumer theory and stochastic frontier approach. *EPRG* Working Paper 1628

Ross, L., Drehobl, A. and Stickles, B. (2018, July). The High Cost of Energy in Rural America: Household Energy Burdens and Opportunities for Energy Efficiency. Retrieved from https://aceee.org/research-report/u1806

Salvador, D. (2018, April). How we can make energy more affordable for low-income families. [Video File]. Retrieved from https://www.ted.com/talks/

Sanstad, A. H. (2011). Notes on the economics of household energy consumption and technology choice. Retrieved from https://www1.eere.energy.gov/

Scally, C., Gilbert, B., Hedman, C., and Posey, A. G. L. (2018). Rental housing for a 21st century rural America. *Research Report* Retrieved from https://www.urban.org/sites/default/files/publication/99045/rental_housing_for_a_21st_century_rural_america.pdf

Schmalnesee, R., and Stavins, R. N. (2017). Lessons learned from three decades of experience with cap and trade. *Review of Environmental Economics and Policy, 11*(1), 59-79.

Schmid, E., Pahle, M., and Knopf, B. (2013). Renewable electricity generation in Germany: A meta-analysis of mitigation scenarios. *Energy Policy, 6*,1151-1163.

Scott, S., Lyons, S., Keane, C., McCarthy, D., and Tol, R. J. (2008). Fuel poverty in Ireland: Extent, affected groups and policy issues. Working Paper 262, Dublin: Economic and Social Research Institute. Retrieved from https://www.esri.ie/UserFiles/publications/20081110114951/WP262.pdf.

Sierra Club. (2018). 100% Commitments in cities, counties, & states. Retrieved from https://www.sierraclub.org/ready-for-100/commitments

Skea, J., Anderson, D., Green, T., Gross, R., Heptonstall, P., and Leach, M. (2008). Intermittent renewable generation and the cost of maintaining power system reliability. *IET Generation, Transmission & Distribution, 2*(1): 82-89.

Solar Energy Industries Association (SEIA). (2018). U.S. solar market insight. Retrieved from https://www.seia.org/us-solar-market-insight

Stone, C. (2015). The design and implementation of policies to protect low-income households under a carbon tax. *Issue brief, Resources for the Future.*

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. *Journal of Graduate Medical Education, 3*(2), 119-20.

Sullivan, M., Collins, M. T., Schellenberg, J., and Larsen, P. H. (2018). *Estimating power system interruption costs: A guidebook for electric utilities.* Retrieved from Lawrence Berkeley National Lab.

Sullivan, M., Mercurio, M. and Schellenberg, J. (2009). *Estimated values of service reliability for electric utility customers in the United States*. Berkeley: Lawrence Berkeley National Lab

Sullivan, M., Schellenberg, J. and Blundell, M. (2015). *Updated value of service reliability estimates for electric utility customers in the United States.* Berkeley: Lawrence Berkeley National Lab

Tahanan, M., Ackooij, W. V., Frangioni, A. and Lacalandra, F. (2015). Large-scale unit commitment under uncertainty. *4OR-Q J Oper Res*, *13*(2), 115-171.

Thompson, P. B. (2002). Consumer theory, home production, and energy efficiency. *Contemporary Economic Policy, 20*(1), 50-59.

Thomson, H., Snell, C., & Bouzarovski, S. (2017). Health, Well-Being and Energy Poverty in Europe: A Comparative Study of 32 European Countries. *International journal of environmental research and public health, 14*(6), 584.

Trochim, W. M. (2006). *The Research Methods Knowledge Base*. 2nd Edition. Retrieved from https://socialresearchmethods.net/kb/citing.php

U.S. Department of Health and Human Services. (2018) LIHEAP factsheet. Retrieved from https://www.acf.hhs.gov/ocs/resource/liheap-fact-sheet-0

U.S. Energy Information Administration (EIA). (2012). Electric generator dispatch depends on system demand and the relative cost of operation. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=7590

U.S. Energy Information Administration (EIA). (2012). Most states have renewable portfolio standards. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=4850

U.S. Energy Information Administration (EIA). (2018). Electricity explained: Electricity in the United States. Retrieved from https://www.eia.gov/energyexplained/

U.S. Energy Information Administration (EIA). (2018). What is U.S. electricity generation by source? Retrieved from https://www.eia.gov

U.S. Environmental Protection Agency (EPA). (2016). Sources of greenhouse gas emissions. Retrieved from https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

U.S. Environmental Protection Agency (EPA). (2018). Emissions & generation resource integrated database (eGRID). Retrieved from https://www.epa.gov/sites/production/files/styles/large/public/2018-02/egrid2016_nerc_regions.jpg

U.S. Environmental Protection Agency(EIA). (2018). Quantifying the multiple benefits of energy efficiency and renewable energy. Retrieved from https://www.epa.gov/sites/production/files/2018

United States Geological Survey (USGS). (2018). Hydroelectric power water use. Retrieved from https://water.usgs.gov/edu/wuhy.html

Varian, H. R. (1992). *Microeconomic Analysis.* New York: W.W. Norton

Villumsen, J. C., Clausen, J., and Pisinger, D. (2011). *Pricing and capacity planning problems in energy transmission networks*. Kgs. Lyngby, Denmark: Technical University of Denmark (DTU).

Vyas, S. and Kumaranayake, L. (2006). Constructing socio-economic status indices how to use principal component analysis. *Health Policy and Planning, 21*, 459-468.

Walstad, W. B., Rebeck, K., and Butters, R. B. (2013). The test of economic literacy: development and results. *Journal of Economic Education, 44*(3), 298–309.

Wangdee, W. (2014). Reliability Impact of intermittent renewable energy source integration into the power system. In *Electrical Engineering Congress (iEECON), 2014 International*, pp. 1-4. IEEE.

Watson, D., & Maitre, B. (2015). Is fuel poverty in Ireland a distinct type of deprivation? *Economic and Social Review, 46*(2), 267-291.

Wilkinson, P., Pattenden, S., Armstrong, B., Fletcher, A., Kovats, R. S., Mangtani, P., and McMichael, A. J. (2004). Vulnerability to winter mortality in elderly people in Britain: population based study. *BMJ, 329*(7467), 647.

Wiser, R., Barbose, G., Heeter, J., Mai, T., Bird, L., Bolinger, M., Carpenter, A., Heath, G., Keyser, D., Macknick, J., Mills, A. and Millstein, D. (2016). *A Retrospective Analysis of the*

Benefits and Impacts of U.S. Renewable Portfolio Standards. Lawrence Berkeley National Laboratory and National Renewable Energy Laboratory. NREL/TP-6A20-65005.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT Press.

Wooldridge, J. M. (2015a). Control function methods in applied econometrics. *Journal of Human Resources, 50*(2), 420-445.

Wooldridge, J. M. (2015b). *Introductory econometrics: A modern approach*. Nelson Education.

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research, 25*(5), 508–518.

Yacobucci, B. D. (2016). Energy policy: 114th congress issues. Library of Congress, Congressional Research Service. Retrieved from https://fas.org/sgp/crs/misc/R42756.pdf

Zaidi, K. R. (2007). Wind energy and its impact on future environmental policy planning: powering renewable energy in Canada and abroad. *Albany Law Environmental Outlook Journal, 11*.

Zubin, J. (1938). A technique for measuring like-mindedness. *The Journal of Abnormal and Social Psychology, 33*(4), 508-516.