Household Energy Expenditure and Consumption Patterns in the United States

Joyance Meechai
The Pennsylvania State University

Abstract
Using data from the Bureau of Labor Statistics 2015 Consumer Expenditure Survey and the U.S. Energy Information Administration, this study researches variations in energy expenditure and consumption patterns in the United States.

The Economist released a data tool comparing energy consumption by state, illustrating that energy usage varied widely. Inspired by this data tool, this project aims to investigate the relationship between household energy expenditure and usage patterns with not just geographic location, but also with sociodemographic characteristics.

This study begins with a cluster analysis to group households by characteristics including housing size, number of cars, and education level. After identifying these clusters, analyses of variance are performed for differences in energy consumption patterns among the clusters. Additionally, the chi-square test is used to study associations between energy type use with other defining variables such as geographical region and housing tenure.

In the face of climate change, there is a call for energy conservation goals. With this study, we seek to discover what factors are associated with certain energy use patterns, and by extension, affect the environment.

Key Words: Energy, environment, consumption, cluster analysis

1. Introduction

The Economist Intelligence Unit released an interactive data tool comparing the per capita energy consumption by state, including breakdowns by energy source. The infographics produced by the tool made clear that energy usage patterns varied widely among the states. For example, in 2014 New York had the lowest energy consumption per capita of the nation at 190 MMBtu, which was mostly in the form of oil or natural gas; whereas in North Dakota, the energy consumption per capita was 865 MMBtu almost half of which was in the form of coal.

Inspired by the Economist data tool, this paper expands upon this topic by investigating the variations in energy expenditure and consumption patterns by not just state, but also by household sociodemographic characteristics. This study uses data from the Consumer Expenditure Survey conducted regularly by the United States Bureau of Labor Statistics (BLS).

We begin with a cluster analysis to bundle households based on their similarities in a number of quantitative sociodemographic characteristics such as housing type and other lifestyle variables. The result of the cluster analysis are groupings of household with distinctive sociodemographic properties upon which a table of demographic and energy expenditure profiles is created. After identifying the sociodemographic clusters, we then compare the energy consumption patterns of the clusters. Analyses of variance are performed to test for significant differences among the clusters in energy consumption levels. The models are further developed with the consideration of additional categorical factors including building type, family makeup, and occupancy tenure.

In addition, chi-square tests are used to determine the significance of associations between a household's use of a particular energy type and categorical variables such as geographical region, building type, family makeup, and household tenure.

1.1 Overview of Data

Data collection is carried out by the United States Census Bureau under contract with the United States Bureau of Labor Statistics. For the Consumer Expenditure Survey, interviews are conducted every three months over four calendar quarters and the consumer units become part of the survey sample on a rolling basis.

Data are recorded on hundreds of variables including household characteristics such as geographic region, family composition, income, education levels, as well as expenditures by category such as expenses for food, clothing, housing, utilities, etc. The data we will be using in this project is taken from the 2015 interviews of the Consumer Expenditure Survey.

In particular, this project focuses on a subset of sixteen variables on over 22,000 survey responses. Surveys for which information is incomplete have been removed. The BLS records expense variables for previous quarter as well as current quarter. In this study, we opted to use the previous quarter amounts as the data are more complete. As the interview files used correspond to the second, third, and fourth quarters of 2015, as well as the first quarter of 2016, the expenditure amounts covered in this paper represent costs spanning all quarters of 2015.

There are three categories of variables being studied:

Energy Expenditure Variables (quantitative):

- Natural Gas (`NTLGASPQ`): Natural gas expenses for previous quarter.
- Electricity (`ELCTRCPQ`): Electricity expenses for previous quarter.
- Fuel Oil (`FULOILPQ`): Fuel oil expenses for previous quarter.
- Other Fuels (`OTHFLSPQ`): Other fuels expenses for previous quarter.
- Gasoline (`GASMOPQ`): Gasoline and motor oil expenses for previous quarter.

Sociodemographic Variables (quantitative):

- Family Size (`FAM_SIZE`): Number of members in the household.
- Number of Automobiles (`NUM_AUTO`): Number of owned automobiles.
- Number of Rooms (`ROOMSQ`): Number of rooms in the housing unit, including finished living areas and excluding all baths.
- Highest education level (`HIGH_EDU`): Highest education level attained within the household expressed in approximate years of schooling.

Sociodemographic Variables (categorical):

- Urban/Rural Classification (`BLS_URBN`): Urban or Rural
- Geographic Region (`REGION`): Northeast, Midwest, South, or West.
- Building Type (`BUILDING`): Apartment, Detached, Non-Detached, or Other.
- Occupancy Tenure (`CUTENURE`): Owner, Renter, or Other.
- Family Type (`FAM_TYPE`): Married Couple Only, Married Couple with other occupants, Single Person, Single Parent, or Other.

In addition to the fourteen variables listed above, the data set used in this study also includes the variable for `CUID` which is the consumer unit identification number used by the BLS to identify the household, as well as a column for `QTRINVMO` which is the interview month. Some of the field response values have been simplified. See the Appendix for details of the simplification.

1.2 Exploratory Data Analysis

An exploratory data analysis quickly reveals that the distributions of the energy expenditure amounts are neither evenly nor normally distributed. Table 1 shows the basic summary statistics for the five energy expenditure variables included in the survey which suggest a right skew to the data. In particular, note that both fuel oil and other fuels have a third quartile expenditure value of zero. Delving deeper, it is seen that in 2015, only 4.5% of households surveyed claimed fuel expenditures. These households tend to be concentrated in the Northeast, and even within this region, represent a minority of the households surveyed. Therefore, we will focus on natural gas, electricity, and gasoline and motor oil.

Table 1: Summary Statistics for Household Energy Expenditures

	Natural Gas	Electricity	Fuel Oil	Other Fuels	Gasoline and Motor Oil
Min	0.00	0.0	0.00	0.00	0.0
1st Qtr	0.00	97.0	0.00	0.00	100.0
Median	12.00	189.0	0.00	0.00	240.0
Mean	73.66	243.2	9.23	5.78	344.7
3rd Qtr	100.00	330.0	0.00	0.00	450.0
Max	3000.00	3400.0	4724.00	3000.00	12075.0

A geographic variation in energy usage is also apparent, and not unexpected considering climate differences over the country. Figure 1 below shows an example of the differences in consumption patterns between two very different states. The popularity of natural gas popularity is very different between Hawaii and Minnesota. There is also a difference in how their energy usage levels vary over the four quarters of a year.

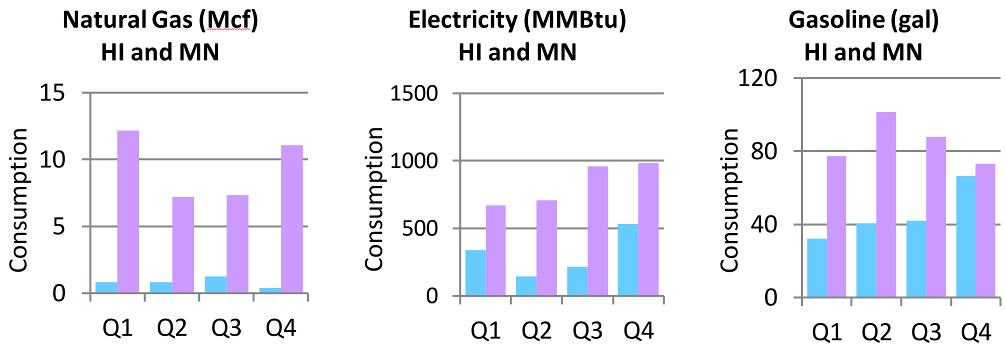

Figure 1: Energy Consumption Levels in Hawaii (blue) and Minnesota (purple)

2. Methods

2.1 Data Preparation

2.1.1 Consumption versus Expenditure

As energy prices vary geographically and our interest is primarily in studying consumption, the energy expenditure variables recorded for each household are converted into energy consumption per capita by considering family size and local retail prices available from the Energy Information Agency (EIA). Average 2015 retail natural gas and electricity residential prices by state were available and used in determining household consumption levels. Gasoline and motor oil consumption levels were approximated using the average regular all formulations gasoline retail prices for a state, when available (as in California or Colorado), or broader region when it was not (as in the use of the Rocky Mountain regular formulation price for Idaho and Wyoming).

2.1.2 Transformation

As noted in the exploratory data analysis, there are strong right skews seen in all the expenditure data. Even after applying the prices per energy unit conversion to the expenditure data, histograms of the energy consumption levels still display strong right skews. To prepare the data for our analyses so that regression assumptions are met, the log transformation $\ln(y+1)$ is applied to the energy consumption response variables which allows us to conduct the regressions and analyses of variance.

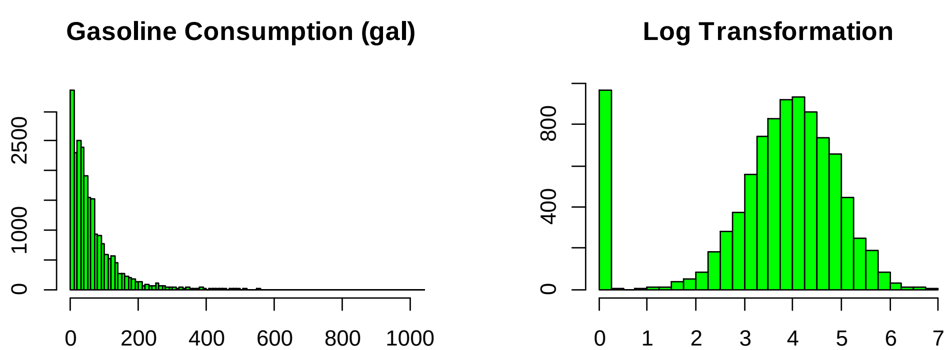

Figure 2: Example histogram illustrating the right skew of the distributions of gasoline consumption (left) which prompts the application of a logarithmic transformation (right).

2.1.3 Data Trimming
Households that reported negative income and negative expenditures were excluded. Households with survey values more than three interquartile range beyond the first quartile and third quartile were also excluded as outliers.

Additionally, there are a number of households that claimed zero expenditures for one of natural gas, electricity, or gasoline and motor oil. These households will be set aside and studied separately.

2.2 Cluster Analysis
Given a household's sociodemographic variables, state, and quarter(s) during which the survey was administered, there are many ways to parse the data. One way to reduce variables is to apply a cluster analysis to the quantitative sociodemographic variables.

In order to group the households by lifestyle characteristics, we will perform a cluster analysis on the following sociodemographic variables: family size, number of automobiles, number of rooms, years of education, and urban/rural classification. Years of education is based on the highest level of schooling within the household. This variable was recorded in the Consumer Expenditure Survey as a categorical factor (HIGH_EDU) which was converted to an approximate number of years in school. See appendix A for details on the conversion used. While urban/rural classification is not a quantitative variable, it will be treated as a binomial variable in the cluster analysis.

A complete linkage clustering is applied to the five sociodemographic variables resulting in a dendrogram which can be partitioned into various numbers of clusters. To determine the appropriate number of clusters, we compare how the various number of clusters would perform in a single factor ANOVA (with cluster as the treatment factor) for each response variable for mean per capita energy consumption. The F-statistics from the ANOVA's for each number of clusters are compared with the objective of choosing the number of clusters which correspond to a strong F-statistic without being too unwieldy.

2.3 Regression Analysis and Analysis of Variance
While this paper is primarily a study of energy consumption, it is based on expenditure survey responses. The question on the relationship between energy price and consumption naturally arose and a regression analysis was done comparing consumption levels and local retail price.

To address the question of whether geography and sociodemographic cluster groups have any distinct energy use patterns, several analyses of variance (ANOVA) were performed for each of the energy types. Initially, a simple single factor ANOVA was done on the log transformed values of per capita energy consumption as a function of state. A separate ANOVA was done on the factor for sociodemographic cluster.

As it was determined that the question of energy consumption patterns was more complicated than a simple consideration of one's state of residence and sociodemographic cluster alone, it became more meaningful to look focus on households within a state and consider other categorical lifestyle variances, such as housing type or occupancy tenure, in addition to the sociodemographic clusters based on qualitative variables in the model.

Because the Consumer Expenditure Survey entails quarterly interviews in which households are rolled onto the roster of participants, there are up to four responses from any given household over the year. To mitigate possible correlational effects from multiple survey records from the same household, we treat this study as a repeated measures design and compare models based on unstructured, variance components, and compound symmetry correlation assumptions. Below is a sample of the SAS code used to conduct the ANOVA for natural gas consumption with state as the fixed variable and consumer unit ID, which identifies the households, as the random variable that is repeated.

```
PROC MIXED data= fmli15;
class state cuid;
model adj_ng = state / ddfm=kr;
repeated / subject=cuid type=cs rcorr;

ods output FitStatistics=FitCS (rename=(value=CS))
FitStatistics=FitCSp;
title 'Compound Symmetry';
RUN;
```

2.4 Chi-Square Test of Association

The analyses of variance study was restricted only to households that reported positive expenditures for natural gas, electricity, and gasoline and motor oil. In the interest of characterizing what distinguishes households that had reported zero expenditures of an energy type, several chi-square tests of associations were applied on the categorical sociodemographic variables. Significant findings are illustrated through mosaic plots.

Similar questions were posed to determine any distinguishing associations between households that claimed fuel oil and other fuel expenditures, from those that did not.

3. Results

3.1 Inelasticity of Energy Demand

Scatterplots and regression analyses on 2015 average retail price values published by the EIA versus consumption levels show weak correlations across all energy types.

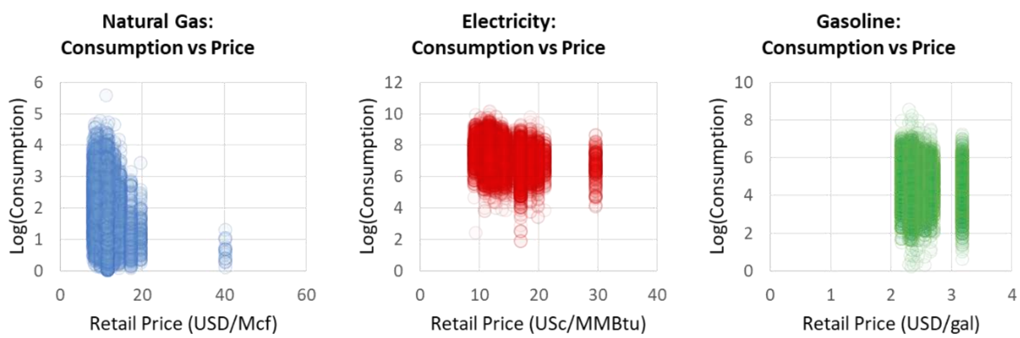

Figure 3: Energy Consumption vs Retail Price Scatterplots

Table 2: Energy Consumption vs Retail Price Regression Results

Regression on Price	Adjusted R-Square	Slope Coefficient
Natural Gas	5.51%	-0.081
Electricity	10.74%	-0.095
Gasoline	0.86%	-0.259

3.2 Cluster Analysis

The complete linkage clustering process resulted in a dendrogram which can be partitioned to various numbers of clusters. To determine what the number of clusters into which we partition the households, the *F*-statistics from a series of one-way ANOVA's were compared for each energy expenditure variable. Fortunately, the optimal number of clusters was fairly clear and consistent over the three energy consumption variables. Table 3 summarizes the results and given the strength of the *F*-statistics, we will proceed with a five group clustering.

Table 3: Table of *F*-Statistic values from one-way ANOVA

Number of Clusters	***F*-Statistic from ANOVA for Energy Response Variable**		
	Natural Gas	Electricity	Gasoline
2	0.13	51.67	0.20
3	0.12	25.90	0.79
4	0.57	20.45	2.19
5	157.60	233.40	98.48
6	126.09	187.19	78.88
7	115.34	168.34	72.85

Based on the five cluster partitioning, the clusters can be characterized as per their defining characteristics in Table 4 below:

Table 4: Summary of Cluster Characteristics

Cluster	**Percent of Households**	**Characteristics**
1	27.5%	Urban; Most automobiles; Most rooms; Most years of education
2	55.0%	Urban; Small family size; Few automobiles; Fewest rooms
3	15.5%	Urban; Largest family size
4	0.5%	Urban; Small family size; Fewest years of education
5	1.5%	Rural

3.3 Analysis of Variance

3.3.1 Analysis of Variance by State

Before taking sociodemographic characteristics into account, a single factor ANOVA was performed to see whether mean energy consumption levels were significantly different by state. The results below are based on the compound symmetry covariance structure for repeated measures.

Table 5: ANOVA Results on Energy Consumption by State

ANOVA by State	***F*-Statistic**	***p*-value**
Natural Gas	33.76	<0.0001
Electricity	27.71	<0.0001
Gasoline	3.83	<0.0001

3.3.2 ANOVA by Sociodemographic Cluster
Single factor analyses of variance were performed comparing the mean per capita consumption by sociodemographic cluster for each of the energy types. There are significant differences in mean consumptions between clusters as illustrated by the interval plots in Figure 4. Note, for example, that cluster 2 with its small family sizes and fewer automobiles has higher per capita consumption levels across the board than cluster 3 which consists of large families.

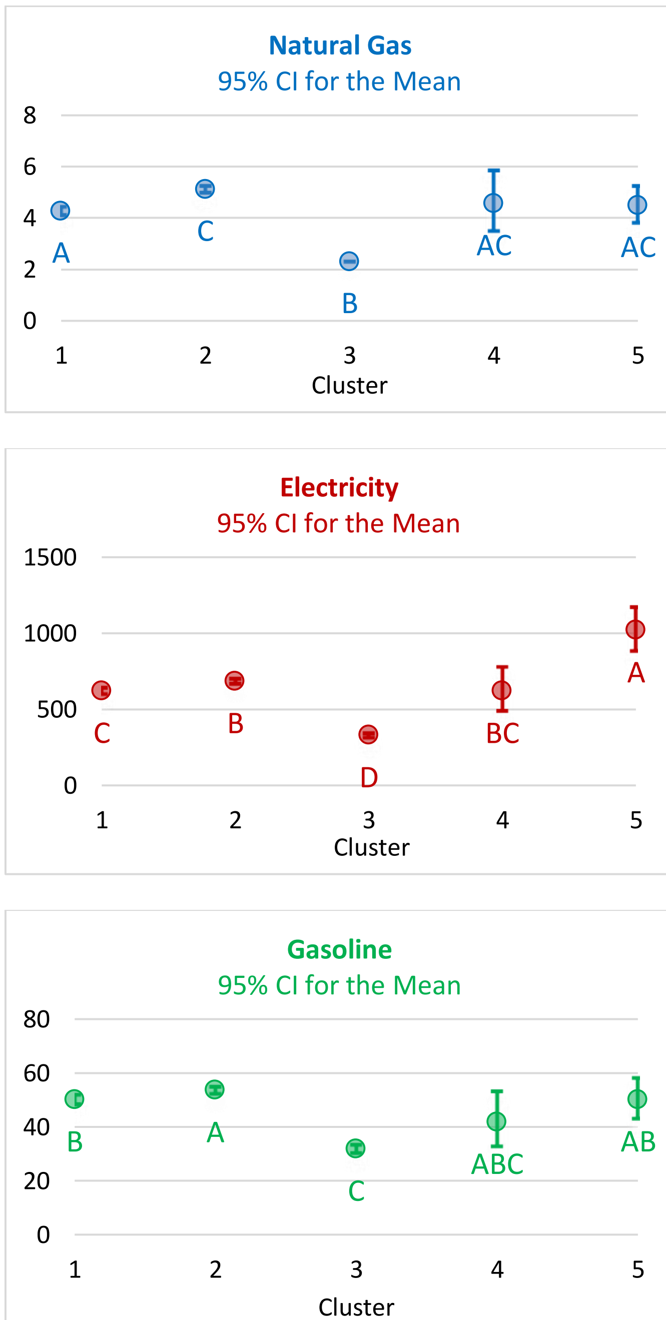

Figure 4: Mean Energy Consumption per Capita by Cluster

Across the fifty states that make up the country, some states are more similar to each other while others are quite disparate. In the interest of reducing complexity, we consider households within a particular state and perform a single factor ANOVA by sociodemographic cluster. Below are results from two states, Minnesota and New York. Again a compound symmetry covariance structure is assumed for the repeated measures. Note that no significant difference in mean gasoline consumption was detected between the different clusters in Minnesota.

Table 6: ANOVA Results on Consumption by Sociodemographic Cluster

State	**Energy Type**	***F*-Statistic**	***p*-value**
Minnesota	Natural Gas	3.25	0.0241
	Electricity	3.24	0.0242
	Gasoline	1.54	0.2066
New York	Natural Gas	12.65	<0.0001
	Electricity	10.49	<0.0001
	Gasoline	5.66	0.0009

3.3.3 ANOVA on Cluster and Other Categorical Variables

Although the results of our ANOVA comparing mean energy consumption across clusters resulted in low *p*-values, the adjusted *R*-square values associated were generally weak. While sociodemographic cluster is significant, there is more to the expenditure pattern picture than the variables we have considered so far. The tables below show the significance of including a second variable in the model, such as building type or occupancy tenure, for households in New York State. In both cases, there was no significant interaction between the second factor and cluster, and so the interaction terms are not included in the models below.

Table 7: ANOVA Results for Consumption in New York State on Cluster and Building Type

Energy Type	**Source**	***F*-Statistic**	***p*-value**
Natural Gas	Cluster	12.33	<0.0001
	Building Type	18.41	<0.0001
Electricity	Cluster	10.44	<0.0001
	Building Type	4.15	0.0065
Gasoline	Cluster	5.82	0.0007
	Building Type	2.10	0.0995

Table 8: ANOVA Results for Consumption in New York State on Cluster and Occupancy Tenure

Energy Type	**Source**	***F*-Statistic**	***p*-value**
Natural Gas	Cluster	10.67	<0.0001
	Occupancy Tenure	13.61	<0.0001
Electricity	Cluster	9.76	<0.0001
	Occupancy Tenure	1.72	0.1798
Gasoline	Cluster	5.49	0.0011
	Occupancy Tenure	0.18	0.8353

After considering several secondary categorical characteristics (building type, occupancy tenure, family type), building type was found to have the most significant contribution to the model. Still, building type was not a significant addition to the model for gasoline consumption. Figure 5 is an interval plot for mean electricity consumption in New York State by cluster and building type.

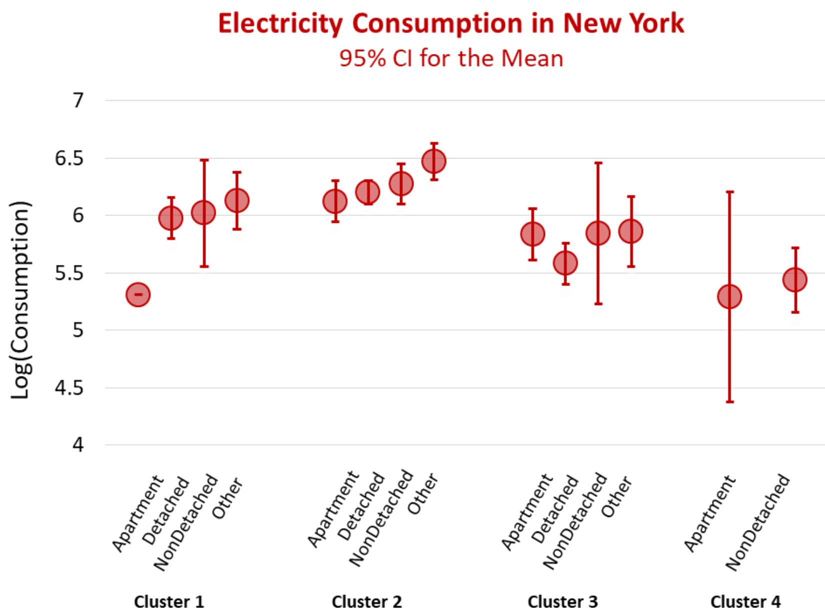

Figure 5: Interval Plot of Mean Electricity Consumption by Cluster and Building Type

3.4 Other Households

In this section, we consider the households that had been excluded from the previous analyses. Chi-square analyses are done to determine which households are more likely to use or not use a given energy type.

3.4.1 Households in which fuel oil and other fuel expenditures are claimed

It was previously noted that only 4.5% of households surveyed claimed any sort of fuel expenditure and that those who did were almost all from the Northeast. Additionally, significant association between building type and whether or not there were fuel expenditures was also found. See Table 9 below for observed distributions and Figure 6 for the corresponding mosaic plot.

Table 9: Comparison of Distributions of Households across Building Types between Those with Fuel Expenditures and Those Without.

Building Type	Apartment	Detached	Non Detached	Other
No Fuels	0.22	0.43	0.11	0.23
Uses Fuels	0.03	0.59	0.03	0.35
χ^2- Statistic: 340.0		***p*-value**: 0.000		

3.4.2 Zero expenditures claimed

The tables below show the different distributions of households that make up those who use a particular energy type versus those who opt out of that energy type completely.

Table 10: Comparison of Distributions of Households across Geographic Region between Those with Natural Gas Expenditures and Those Without.

Region	Northeast	Midwest	South	West
Uses Natural Gas	0.21	0.25	0.25	0.30
No Natural Gas	0.16	0.13	0.52	0.20
χ^2- Statistic: 1817.4		***p*-value**: 0.000		

Table 11: Comparison of Distributions of Households across Geographic Region between Those with Electricity Expenditures and Those Without.

Housing Tenure	Homeowner	Renter	Other
Uses Electricity	0.65	0.01	0.34
No Electricity	0.18	0.06	0.75
χ^2- Statistic: 1789.3	***p*-value**: 0.000		

Table 12: Comparison of Distributions of Households across Geographic Region between Those with Gasoline and Motor Oil Expenditures and Those Without.

Building Type	Apartment	Detached	Non Detached	Other
Uses Gasoline	0.18	0.47	0.10	0.25
No Gasoline	0.55	0.19	0.13	0.13
χ^2- Statistic: 1925.2		***p*-value**: 0.000		

Mosaic plots (Figure 6) neatly portray the proportional differences between households who claim expenditures for a fuel (in a colored scale) versus those who do not (grayscale). By comparing the heights of the horizontal bars between the colored scale and grayscale sides of the plots, one can readily compare the differences in popularity among household subgroups.

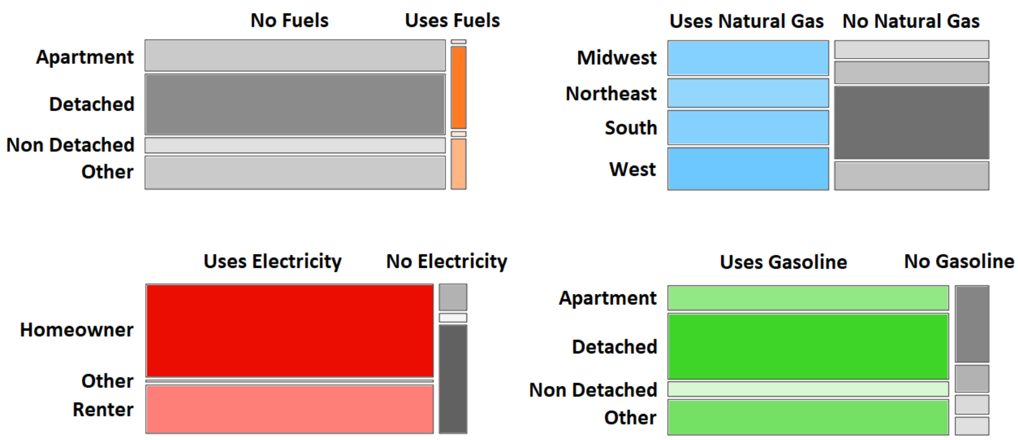

Figure 6: Mosaic Plots

4. Discussion

4.1 Inelasticity of Energy Demand

Despite the law of supply and demand, household energy consumption across state is not strongly tied to price. For example, one would expect higher energy prices to correspond to lower consumption levels. However, taking into account retail pricing data from the EIA, the weak correlations in the scatterplots suggest a relative inelasticity in energy demand, particularly with gasoline.

This seems somewhat surprising at first, but other studies have provided support for this phenomenon. A study by Eitches and Crain, which took a temporal rather than geographic view, looked at gasoline consumption levels from 2004 to 2014. During this time span, gasoline prices were notably volatile and yet, consumption levels remained fairly constant.

If a good is elastic, an increase in price would lead consumers to either find an alternative cheaper substitute, or to simply reduce consumption of that good. A difficulty in changing energy consumption habits or finding alternatives may explain this observation. In the case of gasoline, which is used for transportation, a basic necessity, households seem to be unwilling to reduce consumption, and it may be difficult to find alternatives. For example, public transport may seem like a suitable substitute, but local public transport systems are most likely not developed in places where its use is not already popular. Changing to an energy efficient car entails the purchase of a vehicle. Organizing carpooling may be difficult. In summary, changing habits is hard to do. New York's energy efficiency is credited to the fact that a quarter of the state's residents commute by public transport.

4.2 Significance of Geography and Sociodemographic Characteristics

Although the results of our ANOVA analyses resulted in low *p*-values, the *R*-squares associated were generally weak and it so the predictive values of the models are limited. While sociodemographic cluster is significant, there is more to the expenditure pattern picture than the variables we have considered so far.

To strengthen the model and add more nuance, other categorical sociodemographic characteristics were considered, and it was discovered that the addition of building type improves the predictive value of the model. However, the inclusion of neither family makeup nor occupancy tenure made significant contributions. By including building type, we can see tendencies such as apartment dwellers using less electricity than those living in houses. This generally, however, is not apparent in the cluster of households of larger family sizes.

The chi-square analyses and mosaic plots for expenses claimed for an energy type broken down by occupancy tenure showed that renters were much less likely to claim energy expenditures than homeowners. Perhaps this is more indicative of how energy billing is passed on to rental tenants than it is of actual energy consumption.

As was noted in the discussion of inelasticity of gasoline, the decision to use particular energy types is tied to geography and likely influenced by available infrastructure. Residential use of natural gas is primarily for heating, and thus its use is far less popular in the south than it would be in cooler regions. Beyond that, access to energy reserves and pipelines affect usage. Going back to the example illustrated in Figure 1, natural gas is barely consumed in Hawaii because the state does not produce it and only started receiving shipments of liquefied natural gas in 2014.

Conclusion

In 2015, the United Nations replaced their Millennium Development Goals with seventeen Sustainable Development Goals. There is an interest in energy consumption habits, especially as more societies develop and modernize. While the price of a good is integral to the principle of supply and demand, factors such as need, infrastructure, and accessibility play a role. It is hoped that with the information presented in this study, policy makers can consider what factors may or may not be conducive to energy consumption habits.

Acknowledgements

I would like to thank Dr. Manel Wijesinha and Dr. Eric Nord, from the Pennsylvania State University, for their review and advice. I would also like to acknowledge Dr. Mosuk Chow, also from the Pennsylvania State University, for her continued encouragement throughout my studies and during this project.

Lastly, I would like to thank the ASA Statistical Computing, Government Statistics, and Statistical Graphics sections and Dr. Wendy Martinez, Dr. Thesia Garner, and the team at the Bureau of Labor Statistics for sponsoring the 2017 Data Challenge for which this study was done.

The analyses in this project were performed using R, Minitab, and SAS.

References

Bureau of Labor Statistics. 2017. *Consumer Expenditure Survey Public Use Microdata*. Retrieved from: https://www.bls.gov/cex/pumd_data.htm

The Economist Intelligence Unit. 2015. *Tracking Energy Demand Trends*. Retrieved from: http://trackingenergydemandtrends.eiu.com

Eitches, E., and Crain, V. 2016. Using Gasoline Data to Explain Inelasticity. *In Beyond the Numbers*. Vol. 5, No. 5. Bureau of Labor Statistics. Retrieved from: https://www.bls.gov/opub/btn/volume-5/using-gasoline-data-to-explain-inelasticity.htm

National Oceanic and Atmospheric Administration. 2017. *National Weather Service: Climate*. Retrieved from: http://w2.weather.gov/climate/index.php

US Energy Information Administration. 2017. *Natural Gas Prices*. Retrieved from: https://www.eia.gov/dnav/ng/ng_pri_sum_a_EPG0_PRS_DMcf_m.htm

US Energy Information Administration. 2017. *Electricity Data Browser*. Retrieved from: https://www.eia.gov/electricity/data/browser/

US Energy Information Administration. 2017. *Gasoline and Diesel Fuel Update*. Retrieved from: https://www.eia.gov/petroleum/gasdiesel/

US Energy Information Administration (2011). *Household Heating Fuels Vary Across the Country*. Retrieved from: https://www.eia.gov/todayinenergy/detail.php?id=3690

US Energy Information Administration (2017). *State Profile and Energy Estimates*. Retrieved from: https://www.eia.gov/state/analysis.php

Appendix

Redefinition of Levels for BLS_URBN, REGION, BUILDING, CUTENURE, FAM_TYPE, and HIGH_EDU

The BLS data dictionary defines and records values for these variables on a numeric scale. The table below shows the original BLS definition and the simplification used in this study.

Variable Name	BLS Category Code (numeric)	Redefined Code (character)
BLS_URBN	1 Urban 2 Rural	1 Urban 2 Rural
REGION	1 Northeast 2 Midwest 3 South 4 West	1 Northeast 2 Midwest 3 South 4 West
BUILDING	1 Single family detached	Detached
	2 Row or townhouse 3 End row or end townhouse 4 Duplex 5 Triplex or 4-plex	Non Detached
	6 Garden 7 High-rise 8 Apartment	Apartment
	9 Mobile home or trailer 10 College dormitory 11 Other	Other
CUTENURE	1 Owned with mortgage 2 Owned without mortgage 3 Owned mortgage not reported	Homeowner
	4 Rented	Renter
	5 Occupied without payment of cash 6 Student housing	Other
FAM_TYPE	1 Married couple only	Married Couple Only
	2 Married couple, oldest child < 6yo 3 Married couple, oldest child 6-17yo 4 Married couple, oldest child >17yo 5 Married couple and others	Married Couple And Others
	6 One parent, male, own children 7 One parent, female, own children	Single Parent
	8 Single consumers	Single Person
	9 Other	Other
HIGH_EDU	0 Never Attended	0 years of schooling
	10 1st-8th Grade	6
	11 9th-12th Grade (no high school diploma)	10
	12 HS Graduate	12
	13 Some college, no degree 14 AA degree	14
	15 Bachelor degree	16
	16 Master, professional or doctorate degree	19