Price elasticity of demand for Electricity Consumption

In the Continental United States

Vineet Iyer, Jagadish Bapanapally,
Deekshith Mayan

5/13/16

BigData

Abstract

The purpose of this report is to identify the factors that influence electricity demand in the continental United States of America. This paper estimates the elasticity of electricity demand functions with particular attention paid to the demand side curve that is created by the capital-intensive nature of electricity consumption and to regional, and sectoral variation. The analysis uses a regression model for the short run elasticity that is estimated in a fixed effect framework. For the long run we have used a partial adjustment model for regression analysis considering lagged values. Such Price elasticities papers have been published time and again, but there is still a lack of a significant model which would deal with the endogeneity problem and factor in all of the variables that might affect the demand.

Key Words: HDD (Heating Degree Days), CDD (Cooling Degree Days), GSP (Gross State Product), Demand Elasticities, Regression,

Contents

Introduction

The electricity sector in the United States is large and growing. It makes up for 2.6% of the GDP of the US. 40% of the energy produced goes towards electricity production. However, the growth in electricity demand has been significantly slower than GDP growth for decades. In the 1950s, 1960s, and 1970s the use of electricity often increased more than 5% per year. It then slowed to 2% to 3% per year in the 1980s and 1990s, and over the past decade it has fallen to less than 1% per year. Over the next three decades, electricity use is expected to continue to grow, by 23%, but the rate of growth slows over time. The factors driving this trend include slowing population growth, market saturation of major electricity-using appliances, improving efficiency of several equipment and appliance types in response to standards and technological change, and a shift in the economy toward less energy intensive industry. This projection is visible even in the basic elasticity models where we see a higher impact in the short run but over the long run the impact of price seems to settle down.[1]

Because household and firm level data are costly or impractical for many regions, aggregated models are the most common models in electricity demand estimation.[2] We used this approach in term of data used and we adopt a regression formulation that includes responses to short-run changes in the determinants of demand, and also incorporates longer-run responses through a lagged consumption term. We commonly find that estimates of price elasticities vary widely across regions and that regional elasticities deviate from national averages.[3]

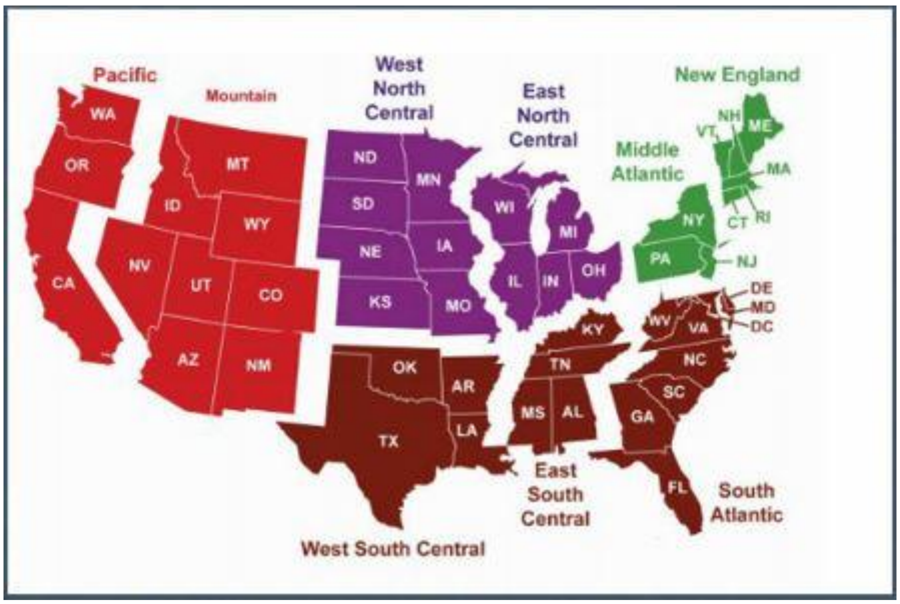
---

[1] (EIA.gov)
[2] (Xavier Labandeira)
[3] (Anthony Paul)

Such projection papers are readily available for view or reading purposes which use many different methods of estimating electricity demand but there is still no concrete model available with perfect projection for elasticity rates.

Elasticity is defined as a ratio of change in price with change in demand. With electricity we have come to realize that Price elasticity of demand for electricity is fairly inelastic. One difficulty in estimating electricity demand is the potential simultaneity between price and quantity. For the purpose of this paper we are going to consider "Quantity Sold" as the quantity demanded. This forgoes the endogeneity problem with the supply and demand curve.[4]

For the purpose of this project we decided to first separate our data into regions mainly 9 regions. And then into 3 sectors, residential, commercial, and industrial.



This gave us 27 Primary equations regression equations. However, for the purpose of finding the best fit of parameters we ran regression analysis for each individual parameter and combination of parameters. We will talk further about this in the result section of the paper.

[4] (Dr. John Garen)

Data and Estimations

The data in the model are monthly observations from several datasets over the period of January 1990 to March 2016. The bulk of the data for each state come from the Energy Information Administration (EIA) Form 826 "Monthly Electric Utility Sales and Revenue Report with State Distributions," which collects retail sales of electricity in gigawatt-hours, and associated revenue in millions of nominal dollars each month, from a sample of electric utilities in the United States. Prices for the model are monthly observed average retail prices, calculated as cents per kilowatt-hour, from the EIA Form 826. All variables measured in dollars are deflated by the Producer Price Index. Daylight hours are the average number of hours of daylight in each month. This data is available from the Astronomical Applications Department of the U.S. Naval Observatory. Heating and cooling degree-days are a measure of the difference in a day's average temperature in Fahrenheit (F) from 65°F, either above (in the case of cooling degree-days) or below. For example, if the mean temperature for a region was 67°F in a given day, there are 2 observed cooling degree-days. The data for heating and cooling degree days are available from the Climate Prediction Center of the National Oceanic and Atmospheric Administration (NOAA-CPC). The income variable is annual disposable income per capita for the residential class and commercial class, (we avoided the endogeneity problem by not using a dummy variable for commercial class which would have been the combination of GSP and personal income.) We used GSP for industrial class.[5]

Hence our models for the short run for residential and commercial used the same combination of parameter that are personal income, electricity price per unit, HDD, CDD, Day light hours. We however did not consider the use of daylight hours for industrial models and used GSP instead of personal income. We ran a multivariable regression based with our dependent variable as the electricity sales. In the long run we just added the lagged values the above parameters to the equation.

[5] (Anthony Paul)

The model that we use allows us to estimate both short-run and long run price elasticities of demand for each customer class, and region. The short-run price elasticity is simply the coefficient on contemporaneous electricity price in the demand equation. The long-run elasticity is calculated using the short-run elasticity and the coefficients on the lagged demand terms. The 27 estimated demand equations yield 81 short- and long-run elasticity estimates.

To set up the Model from the data set we first used a grouping algorithm and grouped the data into 9 different region. After we grouped the data we then filtered it into 3 subsets of Residential, Commercial and Industrial. For state wise data, we called the number assigned to the state in the Data.

Regression

In this section I will be briefing about the methods we have used to implement this project using simple examples using our results. At first we will learn about Linear Regression model and then we will delve into Multiple Linear regression, its methods and the parameters we consider to determine the best fit model.

Linear Regression model

Simple linear regression is a statistical method that allows us to summarize and study relationships between two continuous (quantitative) variables, it is basically a comparison of two models

One is where the independent variable does not even exist

The second model is where an independent variable and a dependent variable exist

Model 1: Single Variable Model

In this model there is just one variable and the best prediction for the other values is the mean of the dependent variable. In our project we have not implemented single variable model, so we have taken a sample example to explain this concept below.

Consider the tables shown in the example below, in model 1 there is just one variable which is Tip amount, In order to predict the value of the next tip amount we consider the mean of all the values as shown in graph below.[6]

Meal #	Tip amount ($)
1	5.00
2	17.00
3	11.00
4	8.00
5	14.00
6	5.00

Figure 1 Sample Example

[6] (Diez)

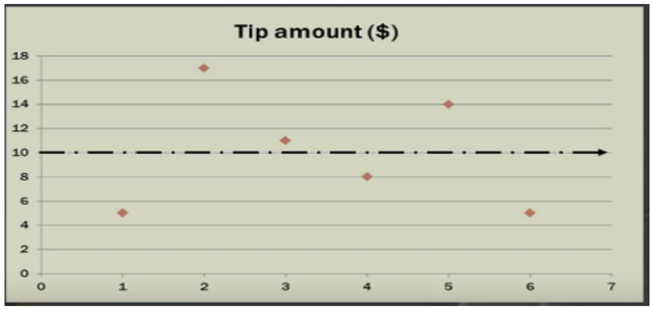


Figure 2Best fit regression line

The difference between the best fit line and the observed value is called the residual, the residuals are squared and then added together to generate the sum of squares

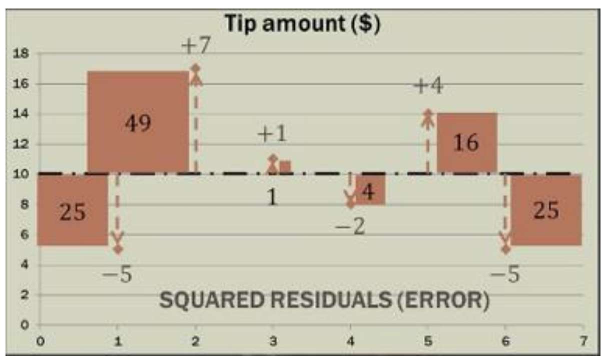


Figure 3 Squared Residuals

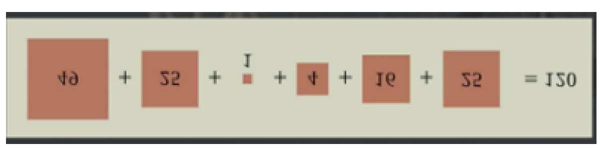


Figure 4 Sum of Squared error

The best model is determined based on the least number obtained from the sum of squared residuals. If for example if there is another model that has a sum of squared of errors less than 120 then that would be the best model .[7]

[7] (Foltz)

Model 2: Two Variable Model

In this model we have two variables, one is the dependent variable and the other is the independent variable based on the data obtained we find the regression line using the equation of a line with slope and intercept formula’s

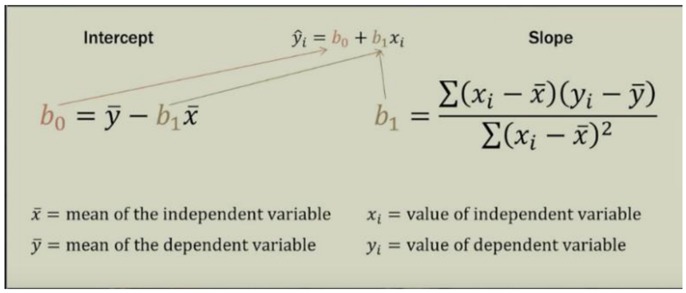


Figure 5 Formula to find the regression line using the slope and the intercept

In figure 6 we have given an example from our project where we are doing linear regression with HDD as independent variable and electricity demand as the dependent variable.

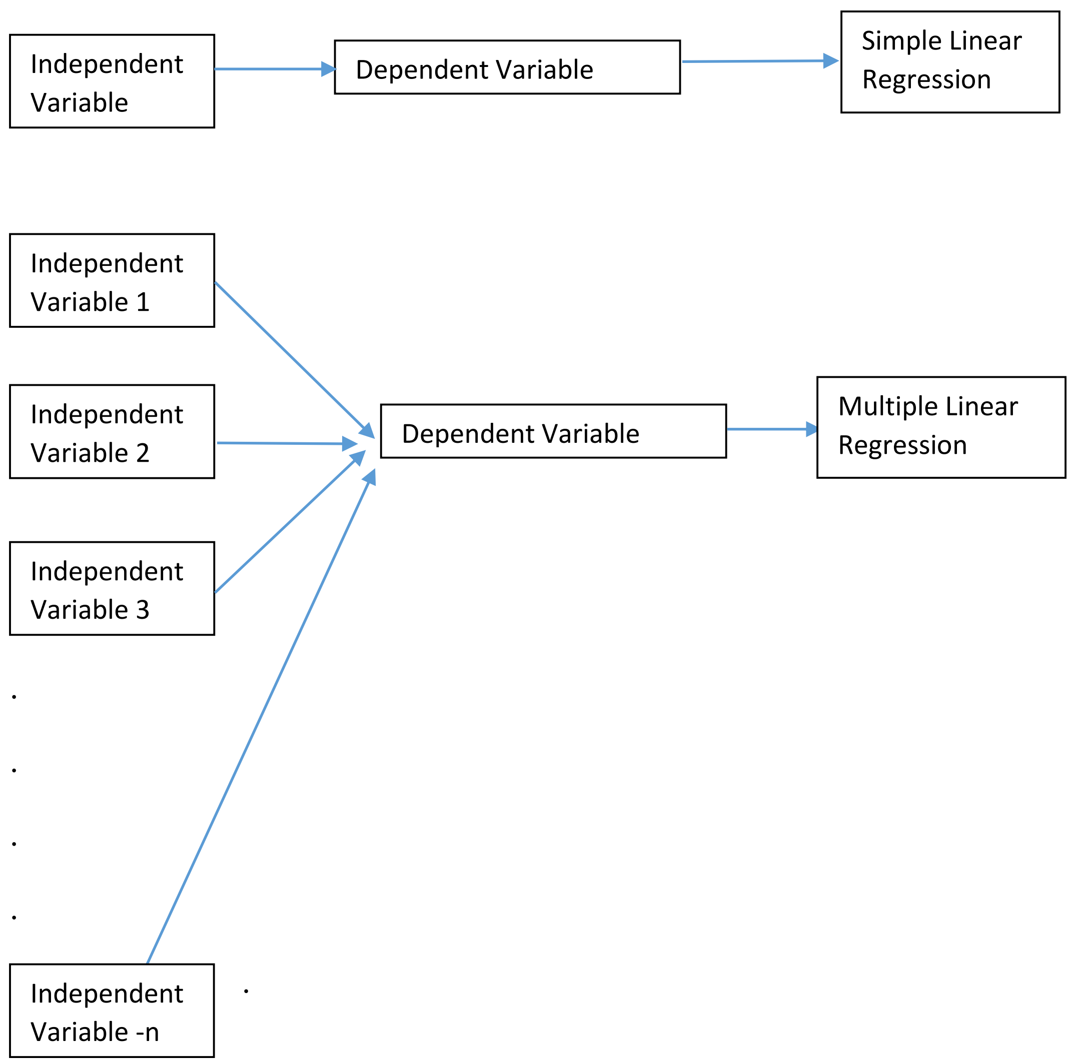
(Heating Degree Days)HDD	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*	*Intercept*
Residential	2.20E-16	0.1882	0.4597	3.83E-04	1.19E+01
Commercial	0.4315	0.2606	-0.001222	2.56E-05	1.25E+01

Figure 6 Example for two variable Linear Regression

We have calculated p-value, Standard error, R^2 Adjusted, Coefficient and Intercept using Simple linear regression model.

Multiple Linear Regression:

Multiple regression is just an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other independent variables. The variable we want to predict is called the dependent variable.



Problems Associated with Multiple Linear Regression

As we know in multiple linear regression there are many independent variables acting on a single variable during such cases there are some problems that arise they are

Over Fitting – Over fitting is caused by adding too many independent variables they account for more variance but add nothing to the model. An over fit model can too complicated for the data set. When this happens, the regression model becomes tailored to fit the quirks and random noise in your specific sample rather than reflecting the overall population. If you drew another sample, it would have its own quirks, and your original over fit model would not likely fit the new data.

We have considered only 5 independent variables against electricity demand for residential and commercial and 4 independent variables for Industrial. To prevent overfitting we did not consider natural gas price and coal price for getting the best fit regression model for electricity demand.

Multicollinearity – this issue usually occurs when you have two or more independent variables that are highly correlated with each other. This leads to problems with understanding which independent variable contributes to the variance explained in the dependent variable, as well as technical issues in calculating a multiple regression model.

In our project we faced the issue with multicollinearity when we considered the variables GDP and personal income, these two variables were strongly correlated with each other. We also found that GDP was more correlated with electricity demand than Personal income. So we have considered GDP instead of personal income for industrial sector.[8]

[8] (etinkaya-Rundel)

Apart from the issues we have discussed above, In order to obtain the best model using multiple linear regression we need to consider the following statements:

- Dependent variable should be measured on a continuous scale (i.e., it is either an interval or ratio variable).
- There needs to be a linear relationship between (a) the dependent variable and each of your independent variables, and (b) the dependent variable and the independent variables collectively. Whilst there are a number of ways to check for these linear relationships, we suggest creating scatterplots and partial regression plots
- Data needs to show homoscedasticity, which is where the variances along the line of best fit remain similar as you move along the line.
- Your data must not show multicollinearity, which occurs when you have two or more independent variables that are highly correlated with each other.
- Finally, we need to check that the residuals (errors) are approximately normally distributed (we explain these terms in our enhanced multiple regression guide).

In our Project we are doing simple regression for each independent variable individually. Each of the analysis that we do is our different models. The regression is conducted using R Studio.
Factors that are being considered in our project to determine the best model

- RSquare
- Adjusted R Square
- Standard Error
- P value

R Square – This is the proportion or percentage of variation in the dependent variable accounted for by independent variable . For ex if RSquare is 0.8442 then 84.42% of the variation in the dependent variable is caused by the independent variable

Adjusted R Square – It is the same as R Square but it gets adjusted with the number of independent variables

Standard Error – It is the average distance of the points from the regression line in the Dependent variable units. For Example if SE is 0.342 on average the data points are 0.342 away from the regression line.

P Value - The P-value is the evidence against a null Hypothesis. A small value of p indicates that it rejects the null hypothesis. Lesser the value of p stronger is evidence that the null hypothesis

A large value of p (>0.05) means the alternate hypothesis is weak so you do not reject the null hypothesis

If p>0.10 – "not Significant"

If p<0.10 – "Marginally significant"

If p<0.05 – "Significant"

If p<0.01 – "Highly Significant

The implementation of multiple linear regression is shown in the Results section.

Results

We have implemented Multiple Linear Regression for getting the best model for Electricity Demand for Residential, Industrial and Commercial in R Studio. We have done coding in R to get the results for all the 3 sectors. We have taken Wyoming state separately for our analysis to get the best model for short run and as well as long run. We have considered logarithm values of some factors to eliminate the units and make it linear.

Below are the factors considered:

CDD – cooling Degree Days

HDD – Heating Degree Days

DL – Logarithmic value of Day Light hours

Per inc- Logarithmic value of Personal Income

EP – logarithmic value of Electricity Price

GDP – Logarithmic value of Gross Domestic Product

Coefficient – This is the value by which electricity demand changes for 1 unit change in the Factor.

EP coefficient – This is the value by which Electricity demand changes for 1 unit change in Electricity price and if all the other factors are kept constant.

Wyoming:

For Wyoming we are giving results of combinations of all the factors that we considered. This is an example which shows in detail of how we applied Multiple linear regression to come up with best Regression model for Wyoming in all the three sectors. The rows which are marked as green in the following table are the best Regression model for Wyoming state for Residential and Commercial. Sectors.

Wyoming Short run for Residential and Commercial:-

Wyoming					
Cooling Degree Days(CDD)	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*	*Intercept*
Residential	6.40E-06	0.2481	0.06066	-0.0014667	12.15806
Commercial	0.1343	0.2599	0.004014	5.02E-04	1.25E+01
(Heating Degree Days)HDD	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*	*Intercept*
Residential	2.20E-16	0.1882	0.4597	3.83E-04	1.19E+01
Commercial	0.4315	0.2606	-0.001222	2.56E-05	1.25E+01

Day Light hours(DL hrs)	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*	*Intercept*
Residential	2.20E-16	0.2069	0.3468	-0.84309	17.67282
Commercial	0.1609	0.26	0.003127	-0.1155	13.2414
Personal Income(per inc)	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*	*Intercept*
Residential	2.20E-16	0.2035	0.3682	0.56345	6.85578
Commercial	2.20E-16	0.09537	0.8659	0.87658	4.28369
Electricity Price(EP)	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*	*Intercept*
Residential	1.98E-11	0.2385	0.1325	0.4671	11.18969
Commercial	2.20E-16	0.1574	0.6346	1.0516	10.57536

CDD + HDD	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1736	0.5405		
Commercial	0.01509	0.2577	0.02048		
CDD + DL hrs	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.2064	0.3498		
Commercial	0.01092	0.2575	0.02253		
CDD + per inc	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1895	0.4522		
Commercial	2.20E-16	0.0953	0.8661		

CDD + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.224	0.2397		
Commercial	2.20E-16	0.1573	0.6352		
HDD + DL hrs	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1885	0.458		
Commercial	0.2648	0.2601	0.002147		
HDD + per inc	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.102	0.8413		
Commercial	2.20E-16	0.09438	0.8686		

HDD + EP	_P-Value_	_Standard Error_	_R^2 adjusted_		
Residential	2.20E-16	0.126	0.7578		
Commercial	2.20E-16	0.1553	0.6442		
DL hrs + per inc	_P-Value_	_Standard Error_	_R^2 adjusted_		
Residential	2.20E-16	0.1391	0.7048		
Commercial	2.20E-16	0.09409	0.8694		
DL hrs + EP	_P-Value_	_Standard Error_	_R^2 adjusted_		
Residential	2.20E-16	0.1682	0.5683		
Commercial	2.20E-16	0.1545	0.6482		

Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1839	0.484		
Commercial	2.20E-16	0.09551	0.8655		
CDD + HDD + DL hrs	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1736	0.5402		
Commercial	0.02408	0.2577	0.02066		
CDD + HDD + Per inc	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.08144	0.8988		
Commercial	2.20E-16	0.09179	0.8757		

CDD + HDD + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1046	0.833		
Commercial	2.20E-16	0.1503	0.6671		
CDD + DL hrs + Per inc	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1393	0.7042		
Commercial	2.20E-16	0.0925	0.8738		
CDD + DL hrs + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1684	0.5673		
Commercial	2.20E-16	0.1511	0.6631		

CDD + Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1759	0.5281		
Commercial	2.20E-16	0.09545	0.8657		
HDD + DL hrs + Per inc	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.102	0.8414		
Commercial	2.20E-16	0.09424	0.869		
HDD + DL hrs + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1259	0.7583		

Commercial	2.20E-16	0.1547	0.647		
HDD + Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.102	0.8412		
Commercial	2.20E-16	0.09447	0.8684		
DL hrs + Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1349	0.7224		
Commercial	2.20E-16	0.09417	0.8692		
CDD+ HDD + DL hrs + Per inc	*P-Value*	*Standard Error*	*R^2 adjusted*		

Residential	2.20E-16	0.08002	0.9023		
Commercial	2.20E-16	0.09185	0.8756		
CDD+ HDD + Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficent*	
Residential	2.20E-16	0.08156	0.8985		
Commercial	**2.20E-16**	**0.09174**	**0.8759**	**5.99E-02**	
CDD+ HDD + DL hrs + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1028	0.8387		
Commercial	2.20E-16	0.1501	0.6676		

CDD+ DL hrs+ Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.1349	0.7223		
Commercial	2.20E-16	0.09249	0.8738		
HDD+ DL hrs+ Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*		
Residential	2.20E-16	0.102	0.8412		
Commercial	2.20E-16	0.09431	0.8688		
CDD+HDD+ DL hrs+ Per inc + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficent*	
Residential	**2.20E-16**	**0.0801**	**0.9021**	**0.0332802**	
Commercial	2.20E-16	0.09178	0.8758		

Wyoming Short run for Industrial:-

The line which is marked as green is the best Regression model for Electricity demand in Industrial Sector. We have considered GDP for Regression analysis as there is multicollinearity between Per inc and GDP.

Cooling Degree Days(CDD)	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*
Industrial	7.10E-01	0.1597	-0.002777	7.66E-05
(Heating Degree Days)HDD	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*
Industrial	6.83E-01	0.1597	-0.002687	8.13E-06
GDP	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*
Industrial	2.20E-16	0.09373	0.6544	0.4111

EP	*P-Value*	*Standard Error*	*R^2 adjusted*	*Coefficient*
Industrial	2.20E-16	0.08788	0.6962	0.57157
CDD + HDD	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	6.21E-01	0.1597	-0.003374	
CDD + GDP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.09385	0.6536	
CDD + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	

Industrial	2.20E-16	0.08799	0.6955	
HDD + GDP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.09376	0.6542	
HDD + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.08754	0.6986	
GDP + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.07913	0.7537	

CDD + HDD + GDP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.09391	0.6531	
CDD + HDD + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.1839	0.484	
CDD + GDP + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.07919	0.7533	

HDD + GDP + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficent*
Industrial	**2.20E-16**	**0.07889**	**0.7552**	**3.59E-01**
CDD + HDD + GDP + EP	*P-Value*	*Standard Error*	*R^2 adjusted*	
Industrial	2.20E-16	0.07897	0.7547	

Wyoming Long Run:

Below is the table which gives the best models for all the three sectors if we take previous month's electricity demand also as one of the factors.

Wyoming Long Run					
Residential	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.06371	0.9383	3.63E-01
Commercial	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.06308	0.9412	7.09E-01
Industrial	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*

		2.20E -16	0.07567	0.7754	2.10E-01

Electricity Demand for Residential:

Below is the table which gives the best regression model for Electricity model for each region for residential sector. We have considered CDD, HDD, DL, Per inc and EP factors for regression analysis for Residential.

Residential	**Electricity Demand Residential vs**				
East North Central	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.1579	0.8042	-6.33E-01
East South Central	*CDD + HDD + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.1208	0.8537	-2.11E-01
Middle Atlantic	*CDD + HDD + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*

		2.20E-16	0.1535	0.7914	-1.53E+00
Mountain	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.1386	0.9704	-2.55E-01
New England	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.1019	0.9878	-5.83E-02
Pacific	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.09649	0.9798	-4.73E-01
South Atlantic	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.1723	0.9493	-9.70E-01

West North Central	*CDD + HDD + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.0984	0.9716	-3.26E-01
West South Central	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.115	0.9829	-1.34E-01

Electricity Demand for Commercial:

Below is the table which gives the best regression model for Electricity model for each region for commercial sector. We have considered CDD, HDD, DL, Per inc and EP factors for regression analysis for Commercial.

Commercial	**Electricity Demand Commercial vs**				
East North Central	*CDD + HDD + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.07075	0.9513	1.62E-01
East South Central	*CDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20E-16	0.07084	0.929	-6.71E-03
Middle Atlantic	*CDD + HDD + Per inc + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*

		2.20E-16	0.06367	0.9544	6.06E-02
Mountain	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20E-16	0.1294	0.96	-6.84E-01
New England	*CDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20E-16	0.1156	0.9842	8.00E-02
Pacific	*CDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20E-16	0.07017	0.9937	-3.10E-01
South Atlantic	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20E-16	0.08642	0.9881	-3.85E-01

West North Central	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20E-16	0.0984	0.9579	-6.65E-02
West South Central	*CDD + HDD + DL hrs + Per inc + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20E-16	0.06394	0.9955	1.13E-01

Electricity Demand for Industrial:

Below is the table which gives the best regression model for Electricity model for each region for Industrial sector. We have considered CDD, HDD, GDP and EP factors for regression analysis for Industrial. We have GDP instead of Personal income as there is strong multicollinearity for these two factors.

Industrial	**Electricity Demand Industrial vs**				
East North Central	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20 E-16	0.1639	0.7195	-1.54E+00
East South Central	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	*R^2 adjusted*	*EP coefficient*
		2.20 E-16	0.1941	0.3679	-5.04E-01

Middle Atlantic	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20 E-16	0.4036	0.6954	-3.23E+00
Mountain	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20 E-16	0.183	0.5105	2.96E-01
New England	*GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20 E-16	0.4512	0.795	5.44E-01
Pacific	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*

		2.20 E-16	0.1011	0.971	-9.35E-01
South Atlantic	*CDD + GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
	same with CDD + HDD + GDP + EP so consider simpler one	2.20 E-16	0.3672	0.6973	-2.63E+00
West North Central	*CDD + GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
	same with CDD + HDD + GDP + EP so consider simpler one	2.20 E-16	0.2464	0.8861	-8.87E-02
West South Central	*CDD + HDD + GDP + EP*	*P-Value*	*Standard Error*	R^2 *adjusted*	*EP coefficient*
		2.20 E-16	0.1418	0.9696	-1.48E-01

Conclusion:

We find that demand for electricity is highly price inelastic in the short run and less so in the long run and that price elasticities vary by customer class, region, and season. Consistent with earlier literature, we find a national, annual average short-run price elasticity across all customer classes of –0.49. Industrial customers exhibit the greatest price responsiveness of demand in the short run, and residential and industrial customers have identical levels of price responsiveness in the long run. Commercial customers are the least price responsive of the three customer classes over both time frames.[9]

Ideally, electricity demand equations would be part of a structural model that includes information about capital stock and explicit decisions about capital turnover and utilization in response to changes in current and expected prices of electricity and other energy sources.

Currently, the lack of data required to estimate such a model for the full range of end-uses of electricity means that any such model must necessarily focus on a subset of end-uses and therefore be of limited value. The model developed in this paper provides a useful reduced-form representation of electricity demand that can be incorporated into an equilibrium model of the electricity sector to provide important information about short-run and long-run responses to changes in prices, conditions, and other factors.[10]

[9] (Anthony Paul)
[10] (Borenstein)

References

[1]Anthony Paul, Erica Myers, and Karen Palmer. "A Partial Adjustment Model of US Electricity Demand." *Resource for Future* (2007): 7-27.

[2]Borenstein, Severin. "To What Electricity Price Do Consumers Respond?" *Berkely* (2010): 7-31.

[3]Diez, David M. *OpenIntro Statistics*. openintro.org, n.d.

[4]Dr. John Garen, Dr. Christopher Jepsen, and James Saunoris. *The Relationship between Electricity Prices and Electricity Demand*. Thesis. Lexington: Gatton College of Business and Economics, 2011.

[5]"EIA.gov." October 2014. May 2016. <https://www.eia.gov/analysis/studies/buildings/energyuse/pdf/price_elasticities.pdf>.

[6]etinkaya-Rundel, Mine Ç. *Openintro*. n.d. May 2016. <openintro.org>.

[7]Griffin, M.A Bernstein and J. "Regional Differences in the Price-Elasticity of Demand for energy. ." *RAND corporation* (2006): 80-116.

[8]M.K. King, D.L.Weimer. "Price and Income elasticities of Demand for Energy." 3 March 2016. *Encyclopedia of life support systems.* May 2016. <http://www.eolss.net/sample-chapters/c08/e3-03-16.pdf>.

[9]Reinhard Madlener, Ronald Bernstein, Miguel Ángel Alva González. "Econometric Estimation of Energy Demand." *E.ON Energy reserch Center Series* (n.d.): 3-56.

[10]Robert I. Kabacoff, Ph.D. *Statistics Methods*. n.d. May 2016. <http://www.statmethods.net/stats/regression.html>.

[11]Ros, Dr. Agustin J. "nera.com." 9th June 2015. *Natural Econimic Research Associate. .* May 2016. <http://www.nera.com/content/dam/nera/publications/2015/PUB_Econometric_Assessment_Elec_Demand_US_0615.pdf>.

[12]Statistics 101. Perf. Brandon Foltz. 2013.

[13]Xavier Labandeira, José M. Labeaga and Xiral López-Otero. "Estimation of Elasticity Price of Electricity with Incomplete Information." January 2010. *Economics of Energy. .* May 2016. <eforenergy.org/docpublicaciones/documentos-de-trabajo/WP01-2010.pd>.

Project Member Roles:

Vineet Iyer: Was responsible for acquiring the data set from various different sources. Major help in this task was provided by Greg Torell, PhD candidate in the Economics Department. After acquiring the data, with help of Travis Dawry, Masters Graduate in Strategic Analytics, and I worked on sorting the data into regions and sectors. Then the statistics and Coding for results and output began. After the Results were acquired, I was responsible for making conclusions on the project.

Deekshith Mayanna – Learnt Linear and Multiple Linear Regression models. Assisted in implementing the models in the project. Helped in choosing various parameters and analyzing data based on that.

Jagadish Bapanapally – Written code in R Studio to load the data. Written code to choose various factors that is needed for Regression Analysis and storing all the outputs in the right format which helped in analyzing the data quickly.