Determinants of Industrial Electricity Usage

Application of the Ordered Probit Model

Ted Pintea

5/8/2013

This paper analyzes the factors that influence electricity usage for industrial customers, and contributes to previous research by utilizing a new data set for large commercial and industrial customers. This analysis is important because these customers represent the largest portion of electricity consumption. Using an ordered probit model, I estimate the probability of an industrial customer falling into one of four electricity usage quartiles. Significant factors to estimate electricity usage are having access to a backup generator and usage of a chiller. Having access to a backup generator increases the probability of an industrial customer will fall into the lower usage quartiles (1, 2), and it decreases the probability of falling into the higher usage quartiles (3, 4). Having backup generation is a substitute for electricity supplied by Ameren UE. Using a chiller decreases the probability of falling into the first quartile, and increases the probability of falling into the third quartile because using a chiller is a large electricity consuming appliance.

I. Introduction

Large commercial and industrial customers use high levels of electricity consumption, and the potential for electricity reduction measures making an impact in this segment is great. Any reduction in these high consumption customers would be significant because of their size. Efficiency measures are difficult to implement for these customers because their electricity usage is a function of their business activities, and often it may seem out of their control to reduce their usage. The purpose of this analysis is to empirically investigate the determinants of electricity usage for industrial customers.

While these customers are of great interest, the reliability and availability of the data is an issue. The characteristics of these customers are closely guarded because they represent a large portion of utilities' customer base. The limited data has not been conducive to research in this area. The nature of this data is almost exclusively categorical. Categorical data was used because of the difficulty of finding continuous data on these customers. Some of the data used by this analysis are the industrial customers' access to alternative generation methods and fuels, age of the buildings, number of buildings, and usage characteristics of large electricity consuming appliances. The constraints of the data create the need for an estimation technique that has the ability to analyze categorical data.

The contribution of this study is an application of the ordered probit model on categorical variables that describe electricity usage. The results of this analysis will differ from an analysis using continuous variables. The coefficients of the explanatory variables are interpreted as the change in the probability of a discrete outcome of the dependent variable given the change in a unit of the explanatory variable. An ordered probit model uses the categorical variables that describe electricity usage as explanatory variables. From the model, we predict the probability of an industrial customer falling into one of four quartiles of electricity consumption. The first quartile customers have the lowest levels of electricity consumption and the fourth quartile customers have the highest levels.

In section II, the paper discusses the theoretical framework of applying the ordered probit model, and includes the rationale for selecting the explanatory variables. In section III, describes the survey conducted by Ameren UE to collect the data and an overview of the data's simple summary statistics. Section IV shows the ordered probit with all of the explanatory variables and an analysis of the expected signs of each of the explanatory variables. Section V is an analysis of the results of the ordered probit model and an interpretation of the marginal effects for the explanatory variables. It also compares the results of Model 1 to Model 2. Section VI has a conclusion including a summary of the results and suggestions for future research.

II. Theoretical Framework

This section of the analysis focuses on the intuition and theoretical background used to select the determinants of industrial customers' electricity usage included in the empirical model. Based on previous research we consider five factors that affect usage: the number of buildings (building) , access to alternative generation (gen), access to alternate fuels (fuel), the type of heating and cooling appliance (HVAC), and the year that the building was built (age).

The age of the construction can be an important factor in the level of electricity usage by the businesses. Costa and Kahn (2011) found a significant relationship between electricity consumption and the year a home was built. The study uses residential data, and similar to my study uses categorical variables that are represented as a vector of year binary variables. The homes that were built in periods when electricity prices were low, such as the mid-1970s, consumed more electricity compared to the base group of homes built from 1998-2000, when electricity prices were higher. The rational can be extended to industrial and large commercial customers as well. Factors such as building codes, electricity conservation programs, and customer preferences that influence the level of electricity efficiency are taken into consideration when constructing commercial buildings. Therefore periods of high electricity prices will result in more electricity efficient large commercial and industrial buildings. Intuitively we expect age to have a negative effect on electricity usage, but given the results of the Costa study the age of the building must also be considered alongside the price of electricity and contemporaneous building codes.

Electricity using appliances are influential determinants of electricity consumption in the short run for both residential and industrial customers. The study by Fisher and Kaysen (1962) identified "white goods" as the stock of electricity-consuming capital goods, and it studies industrial and residential data using continuous variables. The model showed that the demand for electricity is a function of the sum of the products of the average stock and average intensity of "white goods". The most significant electricity-consuming capital goods for industrial and large commercial customer segments are their heating and cooling systems because of the intensity of their use. Some type of heating or cooling appliance must be used during the majority of operating hours. I expect more inefficient heating and cooling equipment to increase usage.

Access to alternative fuels has been used by other studies because they are substitutes to electricity usage. In a study by Denton, et al. (2003), their model consists of three demand equations one for buildings using only electricity, and two for buildings using electricity and natural gas. The buildings that have access to both electricity and natural gas have lower electricity usage compared to the electricity only buildings. Some of the activities or appliances that would be using electricity can be using natural gas. The alternative fuels are included in the ordered probit model are natural gas, propane, steam, and oil. The purpose of including these variables is to test if access to these variables, like natural gas, will reduce electricity usage. I expect alternative fuels to have a negative relationship with usage.

The dependent variable quartile is divided into four usage groups based on the industrial customers' MWh usage per year. The first usage group Quartile 1 is less than 1.23MWh; the second usage group Quartile 2 is between 1.23 and 2.07 MWh. The third usage group Quartile 3 is between 2.07 to 3.58 MWh; the fourth usage group Quartile 4 is greater than 3.58 MWh:

- Quartile 1: ($y^* < 1.23$ MWh)
- Quartile 2: ($1.23 < y^* < 2.07$ MWh)
- Quartile 3: ($2.07 < y^* < 3.58$ MWh)
- Quartile 4:($y^* > 3.58$ MWh)

III. Data

The data was collected by a survey developed by Ameren UE and sent to their commercial customers to examine the interest in energy efficiency rebates. The participants in the survey were offered a $20 Visa gift card as payment for completing the survey. I dropped observations that did not respond to the age of the building to clean up the data, but it was less than 2% of the total observations.

The survey consisted of 174 questions about the companies' electricity usage, business operation, and energy efficiency. Some of the data collected by the survey includes the businesses characteristics about the industry, when the business open, the number of large appliances, the fuel type used by the appliances, alternative type of generation, when the building was created, the types of industrial processes used, exterior of the buildings, number and type of windows, and number and type of lighting.

The interest of the survey was to determine the effectiveness of implementing energy efficiency programs. The data is cross-sectional because the survey was only administered once. The questions are categorical in that the number assigned to the answers holds no information about the magnitude of the variable or its relationship to other variables. The categorical outcomes of each question in the survey are transformed into a binary variable to be used in the model.

The simple statistics for the variables used in the model are summarized in table 1. All of the variables in the sample have 372 observations. All of the variables are binary variables except age which can take three discrete values {1, 2, 3}. Age takes a value of 1 if the buildings were constructed before 1940, it takes a value of 2 if the age is from 1940-1970, and a value of 3 if the age is after 1970. The mean of the binary variables shows the percentage of the variable that has an observation of 1. For alternative generation sources 1.9% have onsite generation, 31.7% have backup generation, and 8.9% have some other generation. The percent of customers that have access to natural gas is 80.7%, oil is 61.8%, steam is 32.3%, and propane is 8.9%. The difference between single and multiple buildings is small, 51.6% of the customers' businesses have a single building and 48.4% have multiple buildings on their campus. In terms of large appliances 74.9% have packaged HVAC systems, 34.2% have air distribution systems, 26.2% have boilers, 24.5% have chillers, and 4.3% have some other HVAC system. The mean for age is 2.648 and this means that the majority of the customers' facilities were built after 1970. The observation for the building being constructed after 1970 is 3 so therefore the higher the value of the mean the newer the construction of the buildings.

IV. Empirical Model

Consider the probit model is based on a normal distribution, with an error term (ε) that has a standard normal distribution. The dependent variable (y*) is a latent variable that is unobservable in the data, but we know exists. The latent variable is a measure of electricity consumption that will take a value of {1, 2, 3, 4} in the ordered probit model. The probit model is based on the standard normal cumulative distribution function(Φ).

$$Prob(quartile = n) = \Phi(quartile_n - \boldsymbol{\beta x}) - \Phi(quartile_{n-1} - \boldsymbol{\beta x}), \quad n = 1 \dots 4$$

$$\begin{aligned}\boldsymbol{\beta x} = {} & \beta_1 onsitegen + \beta_2 backupgen + \beta_3 othergen + \beta_4 natgas + \beta_5 propane + \beta_6 steam \\ & + \beta_7 oil + \beta_8 age + \beta_9 singlebuild + \beta_9 multibuild + \beta_{10} packHVAC \\ & + \beta_{11} airdist + \beta_{12} boiler + \beta_{13} chiller + \beta_{14} otherhvac + \varepsilon\end{aligned}$$

The interpretation and the format I used to apply the ordered probit model to the electricity usage data is based on a paper by (Duncan, et al.). In that study they applied the ordered probit to transportation data and tried to determine the characteristics that could predict the severity of passenger and semi-truck collisions. The coefficients of the ordered probit model are not particularly meaningful because the magnitude cannot be easily interpreted because the explanatory variables and the latent variable do not have well-defined units of measurement. Also the ordered probit model is based on an exponential function rather than a linear function. The estimated coefficients are summarized in Table 2, but to make the interpretation of the model meaningful the marginal effects of the explanatory variables are evaluated. The marginal effects are calculated for each quartile and the results are summarized in Tables 3-5. The marginal effects can be expressed in the following way:

$$\frac{\delta Prob(quartile = n)}{\delta x} = -[\Phi(quartile_n - \beta x) - \Phi(quartile_{n-1} - \beta x)]\beta, \quad n = 1 \ldots 4$$

The vector for number of buildings contains two binary variables, single and multiple, that account for the number of buildings on the organizations campus. The variable multiple is the base group and is omitted in the analysis. The reasoning behind the building vector is that an organization where operations are entirely contained to one building is expected to characteristically have lower electricity usage than an organization that consists of two or more buildings on its campus. These variables are a proxy for the number of square feet that the organization occupies because that information is not directly available from the data set. The expected relationship between the number of buildings and electricity consumption is positive.

The vector for alternative generation contains three binary variables onsite, backup, and other generation, which account for substitutes to Ameren UE’s generation. The organization could have the ability to create their own electricity from backup generators, smaller power plants, renewable energy resources, or the ability to generate electricity from industrial processes such as incineration. The organizations could also choose to use backup generators if the spot price of electricity is more expensive during certain times of the day. None of the binary variables are omitted because an organization could have access to more than one type of alternative generation source. The expected relationship between alternative generation and electricity consumption is negative because they are substitutes.

The vector for fuel contains four variables (natural gas, propane, steam, and oil) which account for the type of fuel the company has access to on the property. The binary variables are not mutually exclusive because the organization may have access to more than one fuel source. Access to alternative fuel sources are substitutes to electricity, and the expected relationship between natural gas, propane, oil, and steam is negative. Access to these alternative fuel sources should decrease electricity consumption because businesses are more likely to use them instead of electricity, and it is more common to have appliances that use these fuel types.

The vector for HVAC contains five binary variables that represent major electricity-consuming appliances that are used by the organization. Again none of the binary variables are omitted because an organization could have access to more than one type of appliance. These variables include packaged HVAC systems, air distribution systems, boilers, chillers, and other. A packaged HVAC system is a single unit that provides both heating and cooling, and these units are found on rooftops and along the side of buildings. An air distribution system is the system that controls the temperature in a building by regulating the amount of cooled or heated air supplied. Boiler refers to water heaters that keep a quantity of water at a certain temperature, or produce hot water as it is demanded to distribute throughout the building. Chiller refers to a machine that removes heat from a liquid, and finally other represents a catchall category for heating or cooling system, the survey asked participants to write-in their own answer. The expected relationship between these variables and electricity consumption is positive. I expect that packaged HVAC is the most efficient appliance, and chiller is the least efficient.

The variable for age accounts for the year that the organization's building or majority of buildings on their campus were constructed. The variable can take a value of {1, 2, 3} these outcomes represent three sets of year ranges. If the variable takes a value of 1 then the construction occurred before 1940, and if the variable takes a value of 2 then the construction occurred sometime from 1940-1970. Finally if the variable takes a value of 3 then the construction occurred after 1970. There is a negative expected relationship between age and electricity consumption. The newer buildings should typically consume less electricity than older buildings and therefore the higher the value taken by age the lower its electricity consumption.

V. Results

We consider four models to evaluate each model's performance. The explanatory variables included in each model are summarized in Table 2. The first model includes all of the variables and the subsequent three models are smaller subgroups of the first model. The second model contains the binary variables for alternative fuel type and alternative generation. The third model contains the binary variables for the number of buildings and their age.

The fourth model contains the binary variables for the large electricity consuming appliances. These variables are separated because they represent the largest proportion of electricity usage. The models control for correlations between the binary variables, and examine the possibility of omitted variables. The models are separated to avoid inflated standard error because of multicollinearity.

The coefficients of the four ordered probit models are reported in Table 2. The only significant coefficient in Models 1 and 2 is backup generation at the 5% level of significance, and it is significant in both the first and second models.   The model that explains the most variation is Model 6 with a pseudo-R^2 of 0.0099. The significant coefficients in Model 4 are boiler and the interaction between age and boiler at the 5% and 10% level of significance, respectively.

The marginal effects for Model 1, the model that includes all of the explanatory variables, are presented in Table 3. The marginal effects summarize the effect that each explanatory variable has on the probability of the industrial customer falling in one of the four quartiles. Four variables are significant: backup generation, steam, oil, and chiller. Backup generation is significant at the 5% level for Quartiles 1 and 3, and it is significant at the 1% level for Quartiles 2 and 4. The sign for the marginal effect for backup generation in Quartiles 1 and 2 are positive which means that having access to a backup generator makes the industrial customer more likely to end up in the first or second quartiles. The industrial customer's access to backup generation makes them more likely to end up in the two lower electricity consumption quartiles. The sign for the marginal effect for backup generation in Quartiles 2 and 3 are negative, which means that having access to a backup generator makes the industrial customer less likely to end up in the third or fourth quartiles. Having a backup generator increases the probability the customer will have lower levels of electricity usage and in turn decreases the probability that the customer will have higher usage. The customers are more likely to use less electricity because they can substitute Ameren UE generation with their backup generation. Other generation follows a similar pattern, but it is not statistically significant. Onsite generation follows the opposite pattern, this suggests it is a complement, but this is not statistically significant.

The marginal effects for Model 1 are represented as graphs in Figures 1-14. The figures illustrate the direction of the marginal effects over all four quartiles of the dependent variable. If the line in the figure is above zero in a quartile, it is an increase in the probability that the industrial customer will have usage in that quartile, and if the line in the figure is below zero it is a decrease in probability.

The marginal effect for the binary variable for access to steam power is significant at the 1% level only in Quartile 2, and insignificant in the other quartiles. The sign for the marginal effect for steam is positive in Quartiles 1 and 2, and this can be interpreted as having access to steam makes the industrial customer more likely to have usage in the second quartile. Access to steam will increase the probability that the consumer will end up in a lower quartile because they may be able to substitute electricity usage with steam power. In all four quartiles natural gas, propane, and steam all follow the same pattern of electricity usage. Access to these fuel types increases the customers' probability of having lower electricity usage. The marginal effects for these fuel types are decreasing across all four quartiles.

The marginal effect for the binary variable for access to oil is significant at the 5% level only in Quartile 3, but is insignificant in any of the other quartiles. The sign for the marginal effect for oil is positive which can be interpreted as having access to oil will increase the probability that the industrial customer will have usage in the third quartile. Unlike access to steam access to oil results in higher usage because the cost of using oil may be more expensive than electricity usage even at higher usage rates. Therefore oil is a compliment, if oil is more expense the industrial customer will use more electricity. In contrast oil has the opposite pattern, the marginal effects for oil is increasing across all four quartiles.

The marginal effects for the binary variable chiller are significant at the 10% level in Quartile 1 and significant at the 5% level in Quartile 3. The sign for using a chiller is negative in Quartiles 1 and 2, and positive in Quartiles 3 and 4. Using a chiller decreases the probability that the industrial customer will end up in the low usage quartiles and increases the probability that they will end up in the high usage quartiles. The chiller has such high electricity demands that its use is an indicator that an industrial customer will end up being a higher usage customer. The marginal effects for using a chiller or an air distribution system is decreasing across all four quartiles using either of these appliances will increase the probability of higher usage. The marginal effects for using a packaged HVAC system, a boiler, or other HVAC system are increasing across all four quartiles using these appliances decrease the probability of higher usage. According to the summary statistics, Table 1, 74.9% of the customers have a packaged HVAC system, and access to these systems may lead to increased energy efficiency compared with having a boiler or a chiller. Other HVAC has the greatest probability of being a low usage customer.

The marginal effects for age are decreasing over all four quartiles. This can be interpreted as the newer the age of the building, if it was built after 1970, increases the probability of having lower electricity usage. The newer buildings are more energy efficient either from pressure from higher electricity prices when they were constructed or from legislation that mandates efficiency. The marginal effects of age in Model 1 are the same as in Model 3 where age is isolated with single and multiple buildings (Table 5).

Marginal effects for Model 2 are summarized in Table 4. In this model the alternative generation and fuel sources are separated from the Model 1. When these binary variables are isolated backup generation remains statistically significant at the 5% and 10% levels which are a slight drop in statistical significance compared to Model 1. The signs of the marginal effects in Quartiles 1 and 2 remain positive, and the signs in Quartiles 3 and 4 remain negative. Therefore access to a backup generator increases the probability that the industrial customer will have usage in the lower two quartiles, and it decreases the probability that they will have usage in the higher two quartiles. The marginal effect for the binary variable for access to steam is no longer statistically significant. The marginal effect for the binary variable for access to oil remains significant at the 1% level of significance, and the sign remains positive. Access to oil increase the probability that the industrial customer will have higher usage in the third quartile.

The marginal effects of natural gas in Model 1 change when alternative generation and fuel sources are isolated in Model 2 (Table 4). In Model 2 natural gas is negative in all quartiles except the fourth, and this means that natural gas increases the probability that the customer will be in the highest usage quartile. In contrast in Model 1 access to natural gas increases probability of being in the lower usage quartiles. The difference in the marginal effects in these two models suggests mixed results as to whether natural gas is a substitute or a complement to electricity usage. This change in the marginal effects for natural gas may be due to omitted variables bias because when more variables are included in Model 1 the results are consistent with intuition.

Marginal effects for Model 3 are summarized in Table 5. In Model 3 the age and number of buildings are isolated from Model 1. The coefficient for multiple buildings is omitted because of perfect collinearity, and the base group is the binary for single building. The marginal effects are not significant for any of the explanatory variables or any of the four quartiles. The direction of the marginal effects remains the same in both models 1 and 3.

The marginal effects for Model 4 are summarized in Table 6. In model 4 the large electricity consuming appliances are isolated from Model 1. The direction of the marginal effects remains the same as in Model 1. The only difference is that the marginal effects for chiller are no longer significant.

VI. Conclusions

In my analysis I used an ordered probit model to examine the factors of electricity usage by large commercial and industrial customers. The determinants that most influence electricity usage are access to backup generation and chillers. Backup generators decrease the probability of high electricity usage and having a chiller increase the probability of high usage. Another factor that influences electricity consumption is access to alternative fuels specifically steam and oil. Access to steam power decreases the probability of high usage while access to oil increases the probability to high usage. The statistical significance of these alternative fuel sources, as well as access to chillers, diminishes when the binary variables for alternative generation and fuel are isolated from the larger subset of binary variables in Model 1. The statistical significance of backup generation remains in both the Model 1 and the smaller Model 2.

Topics for future research would be to understand the effect that changes in the prices for alternative fuels such as natural gas, oil, and propane have on the demand for electricity. I want to determine at what prices would more large commercial customers substitute electricity consumption with these alternative fuels. I would also include weather patterns, average temperatures, and regions to see the effect that these factors have on energy efficiency. Those customers that are in warmer regions of the country would be expected to be less energy efficient because their use of air conditioning on more days of the year.

VII. References

Costa, Dora L., and Matthew E. Kahn. "Electricity Consumption and Durable Housing: Understanding Cohort Effects." *American Economic Review* 101.3 (2011): 88-92. *EconLit with Full Text*. Web. 14 Apr. 2013.

Daykin, Anne R., and Peter G. Moffatt. "Analyzing Ordered Responses: A Review Of The Ordered Probit Model." Understanding Statistics 1.3 (2002): 157. Academic Search Complete. Web. 14 Apr. 2013.

Denton, Frank T., Dean C. Mountain, and Byron G. Spencer. "Energy Demand with Declining Rate Schedules: An Econometric Model for the U.S. Commercial Sector." Land Economics 79.1 (2003): 86-105. EconLit with Full Text. Web. 14 Apr. 2013.

Duncan, C. S., A. J. Khattak, and F. M. Council. "Applying The Ordered Probit Model To Injury Severity In Truck - Passenger Car Rear-End Collisions." Transportation Research Record 1635 (1998): 63-71. British Library Document Supply Centre Inside Serials & Conference Proceedings. Web. 14 Apr. 2013.

Fisher, F.M. and Kaysen, C. "A study in Econometrics: The Demand for Electricity in the United States." Amsterdam: North Holland Publishing Co., 1962.

Taylor, Lester D. "The Demand for Electricity: A Survey." Bell Journal of Economics 6.1 (1975): 74-110. JSTOR with Full Text. Web. 14 Apr. 2013.

Wooldridge, Jeffrey M. Introductory Econometrics: a Modern Approach. Australia: South-Western College Pub., 2003. Print.

VIII. Tables

Table 1: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Onsite Generation	372	0.019	0.136	0	1
Backup Generation	372	0.317	0.466	0	1
Other Generation	372	0.089	0.285	0	1
Natural Gas	372	0.807	0.396	0	1
Propane	372	0.089	0.285	0	1
Steam	372	0.323	0.177	0	1
Oil	372	0.618	0.241	0	1
Age	372	2.648	0.576	1	3
Single Building	372	0.516	0.5	0	1
Multiple Buildings	372	0.484	0.5	0	1
Packaged HVAC	372	0.749	0.434	0	1
Air Distribution System	372	0.342	0.475	0	1
Boiler	372	0.262	0.44	0	1
Chiller	372	0.245	0.431	0	1
Other HVAC	372	0.043	0.203	0	1

Table 2: Ordered Probit: Coefficients

Quartile	Model 1	Model 2	Model 3	Model 4
Onsite Generation	0.036 (0.434)	0.071 (0.428)	-	-
Backup Generation	-0.368** (0.145)	-0.265** (0.132)	-	-
Other Generation	-0.249 (0.295)	-0.116 (0.289)	-	-
Natural Gas	-0.015 (0.211)	0.029 (0.208)	-	-
Propane	-0.044 (0.225)	-0.047 (0.222)	-	-
Steam	-0.375 (0.335)	-0.205 (0.318)	-	-
Oil	0.293 (0.247)	0.334 (0.246)	-	-
Single Building	0.005 (0.117)	-	0.012 (0.112)	-
Age	-0.001 (0.101)	-	-0.031 (.097)	
Packaged HVAC	-0.057 (0.149)	-	-	-0.037 (0.147)
Air Distribution System	0.041 (0.146)	-	-	0.021 (0.143)
Boiler	-0.154 (0.177)	-	-	-0.05 (0.164)
Chiller	0.315 (0.203)	-	-	0.108 (0.183)
Other HVAC	-0.219 (0.297)	-	-	-0.195 (0.293)
$Pseudo-R^2$	0.0099	0.0055	0.0001	0.0012

*,**,*** denotes significance at 1%, 5%, and 10%, respectively

(se): number in parenthesis is the standard error

Table 3: Model 1: Marginal Effects

Quartile	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Onsite Generation	-0.011	-0.003	0.003	0.012
	(0.133)	(0.04)	(0.033)	(0.14)
Backup Generation	0.124**	0.026*	-0.036**	-0.114*
	(0.049)	(0.01)	(0.017)	(0.041)
Other Generation	0.083	0.015	-0.026	-0.073
	(0.105)	(0.012)	(0.036)	(0.079)
Natural Gas	0.005	0.001	-0.001	-0.005
	(0.066)	(0.018)	(0.017)	(0.067)
Propane	0.014	0.004	-0.004	-0.014
	(0.072)	(0.017)	(0.02)	(0.069)
Steam	0.13	0.016*	-0.043	-0.104
	(0.126)	(0.006)	(0.048)	(0.079)
Oil	-0.084	-0.032	0.015**	0.1
	(0.063)	(0.032)	(0.007)	(0.09)
Age	0.0003	0.00008	-0.00008	-0.0003
	(0.032)	(0.009)	(0.008)	(0.032)
Single Buildings	-0.002	-0.0004	0.0004	0.002
	(0.037)	(0.01)	(0.01)	(0.037)
Packaged HVAC	0.018	0.005	-0.005	-0.018
	(0.046)	(0.014)	(0.011)	(0.048)
Air Distribution	-0.013	-0.004	0.003	0.013
System	(0.045)	(0.013)	(0.012)	(0.046)
Boiler	0.05	0.012	-0.014	-0.047
	(0.058)	(0.012)	(0.018)	(0.053)
Chiller	-0.093***	-0.031	0.02**	0.104
	(0.056)	(0.024)	(0.01)	(0.07)
Other HVAC	0.073	0.014	-0.02	-0.064
	(0.105)	(0.012)	(0.036)	(0.08)

*,**,*** denotes significance at 1%, 5%, and 10%, respectively

(se): number in parenthesis is the standard error

Table 4: Model 2: Marginal Effects

Quartile	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Onsite Generation	-0.022	-0.007	0.005	0.023
	(0.128)	(0.042)	(0.028)	(0.143)
Backup Generation	0.086**	0.019**	-0.024***	-0.082**
	(0.044)	(0.009)	(0.014)	(0.039)
Other Generation	0.038	0.009	-0.01	-0.036
	(0.097)	(0.018)	(0.029)	(0.086)
Natural Gas	-0.009	-0.002	-0.002	0.009
	(0.066)	(0.017)	(0.018)	(0.066)
Propane	0.015	0.004	-0.004	-0.015
	(0.072)	(0.017)	(0.02)	(0.069)
Steam	0.069	0.013	-0.021	-0.061
	(0.112)	(0.013)	(0.038)	(0.087)
Oil	-0.094	-0.037	0.015*	0.116
	(0.061)	(0.033)	(0.006)	(0.091)

*,**,*** denotes significance at 1%, 5%, and 10%, respectively

(se): number in parenthesis is the standard error

Table 5: Model 3: Marginal Effects

Quartile	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Age	0.01 (0.031)	0.003 (0.008)	-0.003 (0.008)	-0.01 (0.031)
Single Building	-0.004 (0.035)	-0.001 (0.009)	0.001 (0.009)	0.004 (0.036)

*,**,*** denotes significance at 1%, 5%, and 10%, respectively

(se): number in parenthesis is the standard error

Table 6: Model 4: Marginal Effects

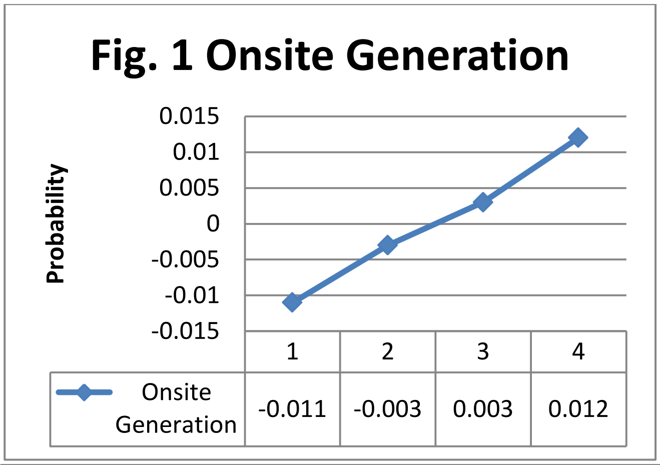
Quartile	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Packaged HVAC	0.012 (0.046)	0.003 (0.013)	-0.003 (0.011)	-0.012 (0.047)
Air Distribution System	-0.007 (0.045)	-0.002 (0.012)	0.002 (0.011)	0.007 (0.046)
Boiler	0.016 (0.053)	0.004 (0.013)	-0.004 (0.014)	-0.016 (0.051)
Chiller	-0.034 (0.056)	-0.01 (0.017)	0.008 (0.013)	0.035 (0.06)
Other HVAC	0.065 (0.065)	0.012 (0.013)	-0.019 (0.034)	-0.058 (0.081)

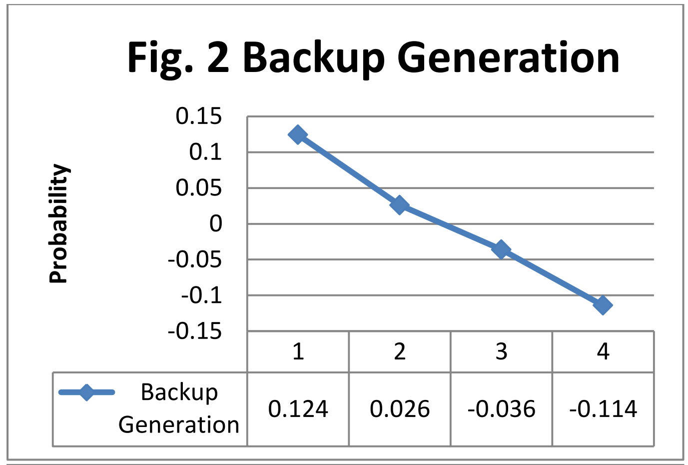
*,**,*** denotes significance at 1%, 5%, and 10%, respectively

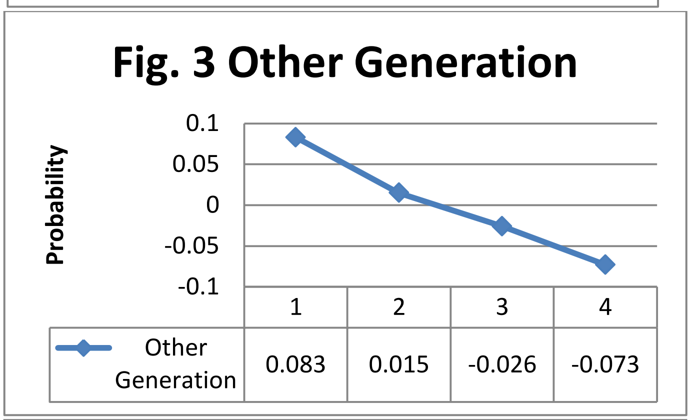
(se): number in parenthesis is the standard error

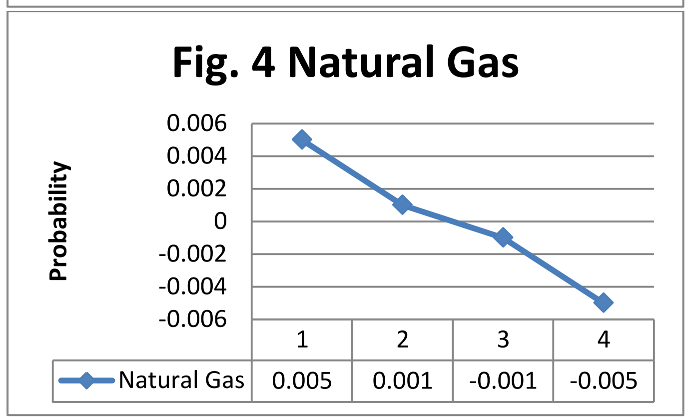
IX. Graphs

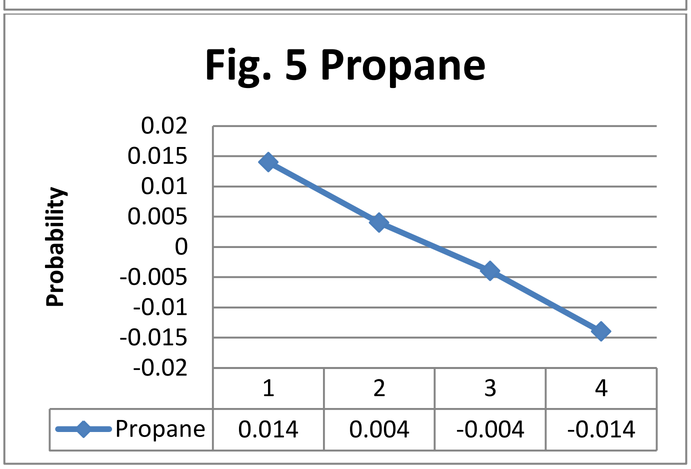
<u>Marginal Effects Model 1:</u>

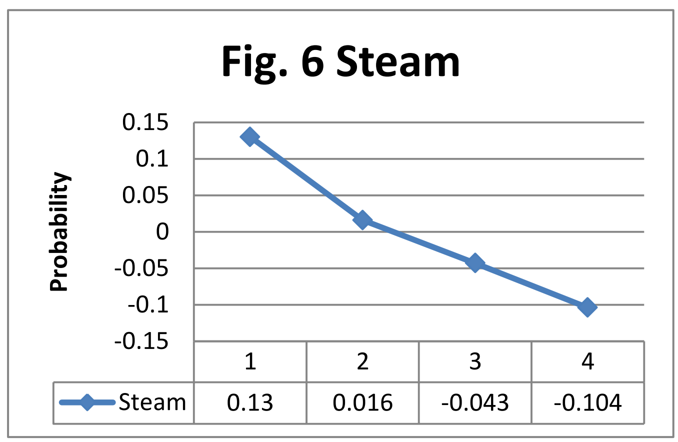












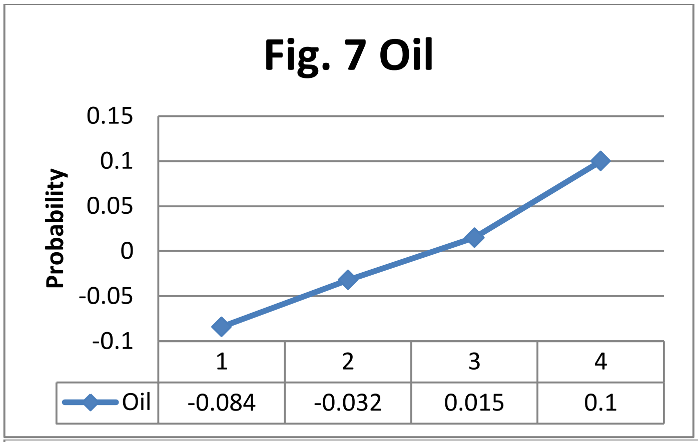


Fig. 7 Oil

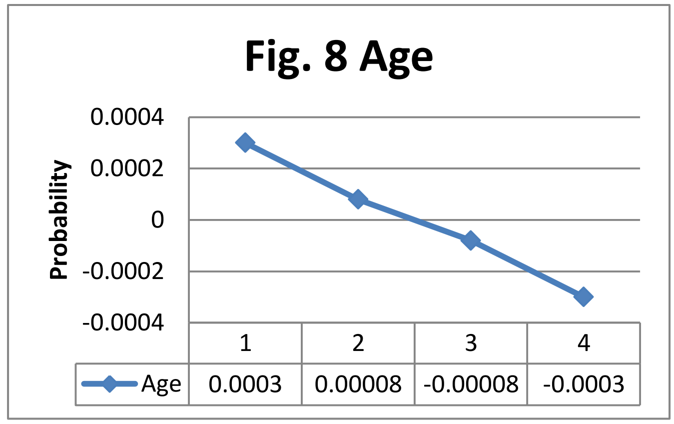


Fig. 8 Age

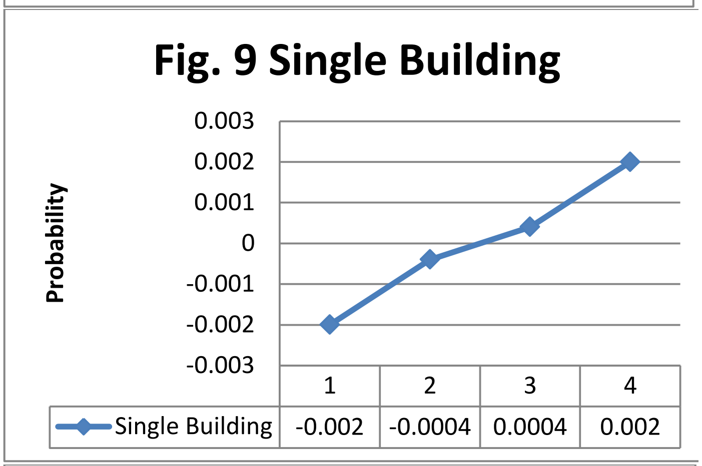


Fig. 9 Single Building

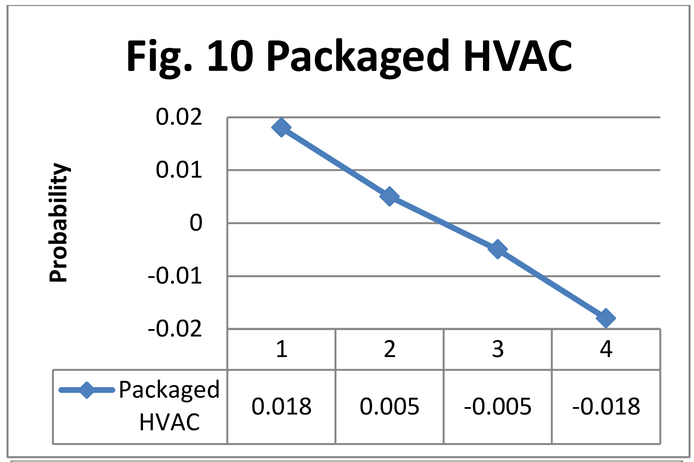


Fig. 10 Packaged HVAC

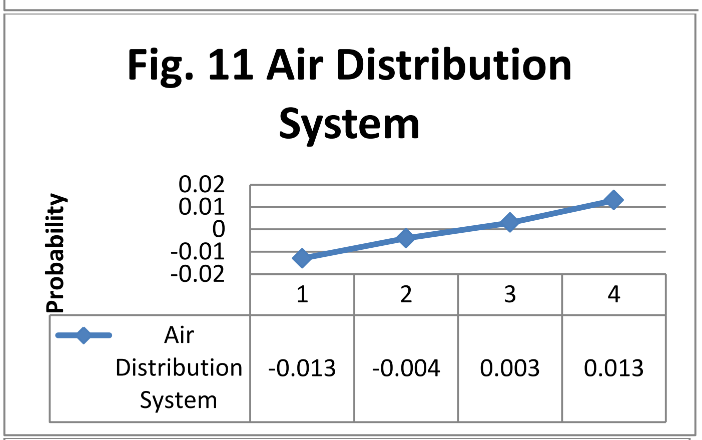


Fig. 11 Air Distribution System

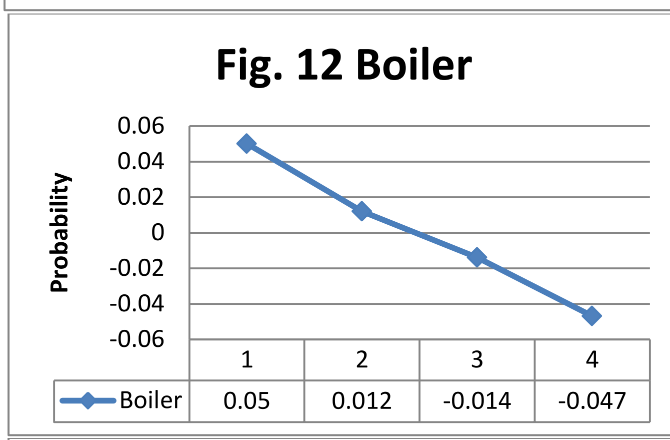


Fig. 12 Boiler

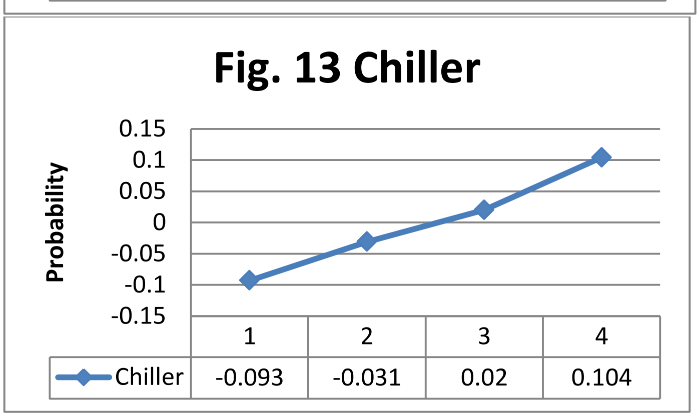


Fig. 13 Chiller

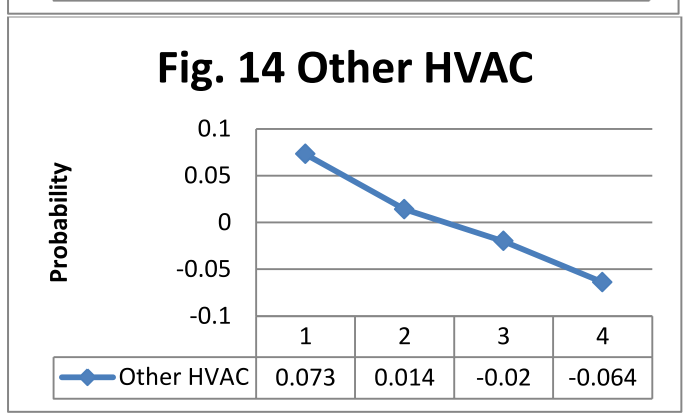


Fig. 14 Other HVAC