Statistics 440 Individual Project

Austin Kao

3/3/2021

Introduction

One of the most pressing issues of today is man-made climate change. Climate change is spurred by the emission of greenhouse gases into the atmosphere. The energy sector, which burns fossil fuels to help generate power for electricity, heating, industry, and transport, contributes about two thirds of global greenhouse gas emissions, motivating a push to find renewable and sustainable energy sources to replace fossil fuels (Energy, 2020). In fact, according to the International Renewable Energy Agency (IRENA), operating new solar photovoltaic and onshore wind power plants cost less than operating existing coal-fired plants (Renewables, 2020).

We would like to further explore the IRENA's claim and its implications by assessing the economic performance of electrical utilities in the U.S. Namely, we want to determine if any significant relationship exists between the proportion of renewable energy power plants an electric utility owns and the average price of electricity sold by that utility. By doing so, we hope to gain insight into the state of the U.S. energy market and whether or not renewable energy technologies, in general, will be economically viable.

Dataset

The data we will use for this analysis comes primarily from the US Energy Information Administration (EIA), a government agency that gathers data about the country's operational power plants. The EIA collects generator-specific data on the electricity generated throughout the nation through surveys each power plant completes known as Form 923. It also collects data on the finances of electric utilities throughout the nation through surveys each utility completes known as Form 861.

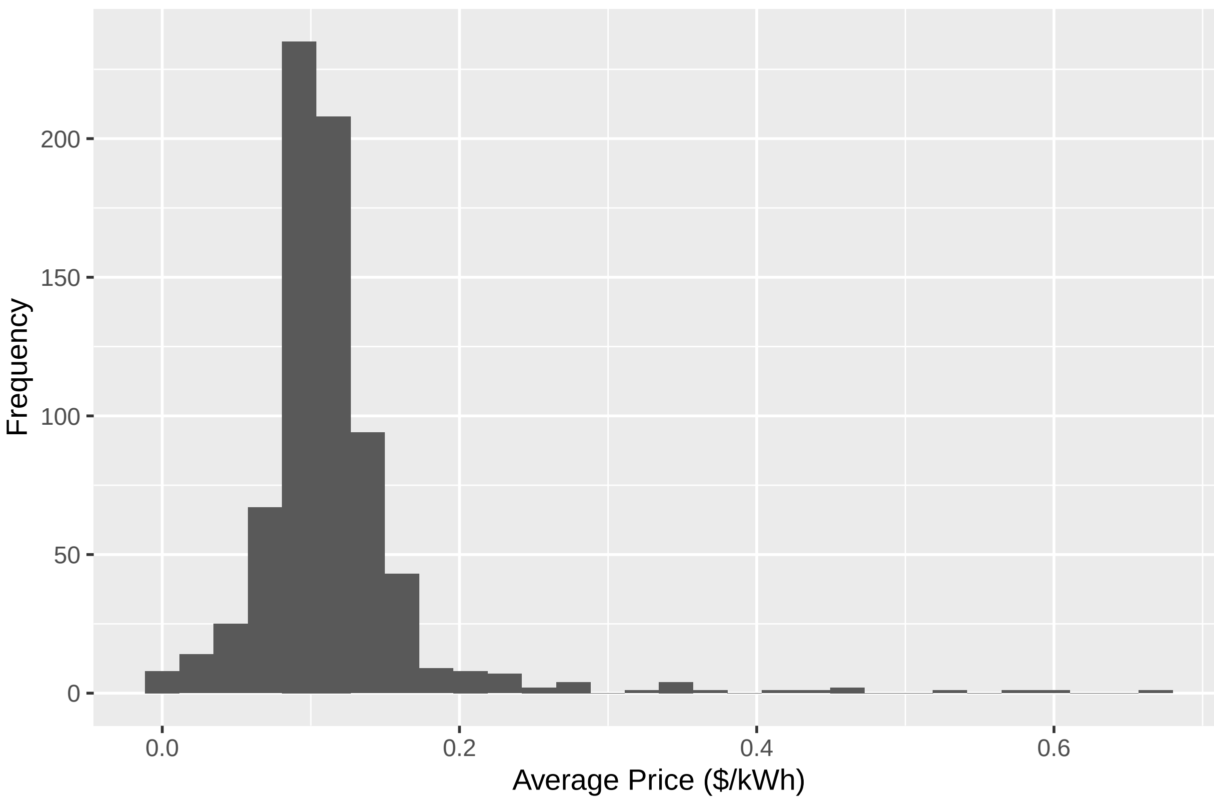
Variables

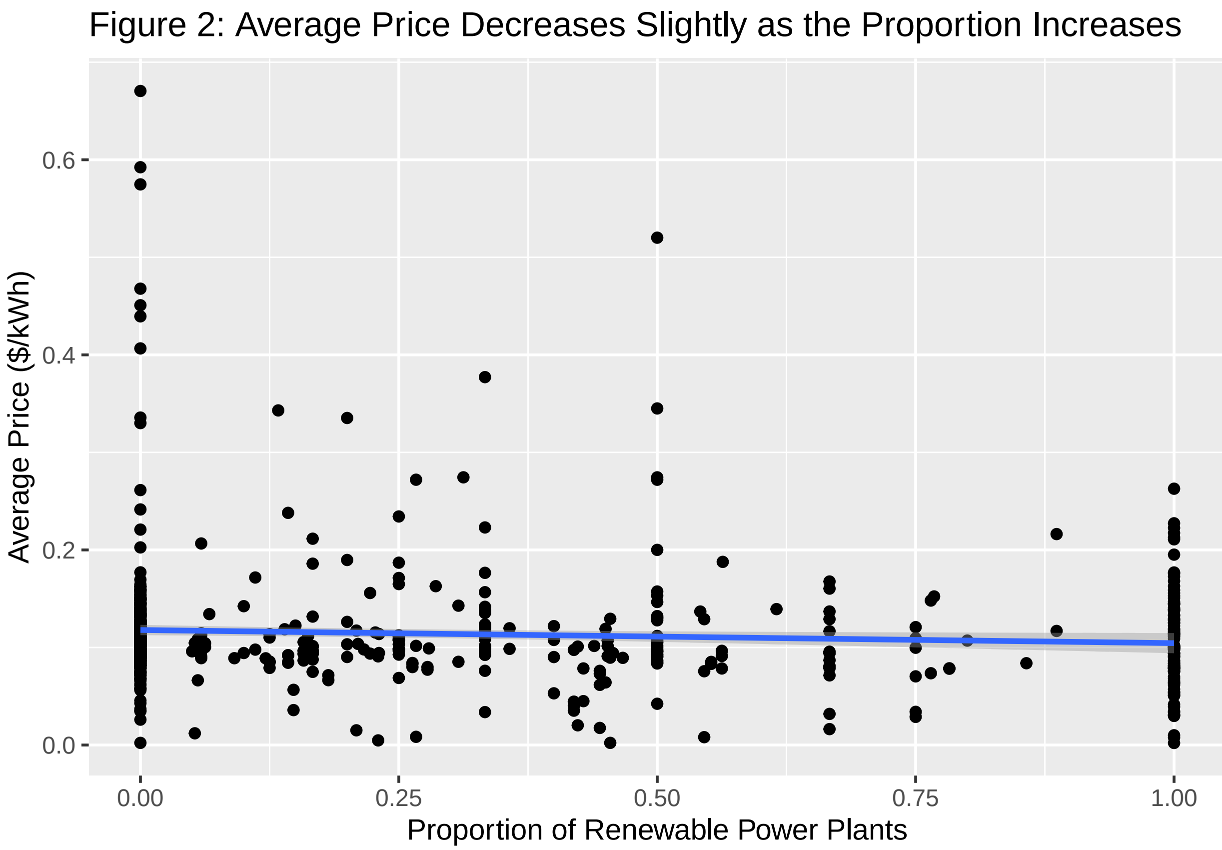
To model the performance of electrical utilities, for our response variable, we will use the average price of the electricity sold, measured in dollars per kilowatt-hour, by utilities in the year 2019. The utilities that we are primarily concerned with are those that have generated electricity at a power plant and sold that electricity to an end user, a customer who will use the electricity. Plenty of electrical utilities act the part of wholesalers, buying and selling electricity in bulk between electricity utilities. Many other utilities only buy their electricity from electrical wholesalers to sell to end users, producing no electricity of their own. We chose to focus on the utilities that generate their own electricity and sell to end users because the relationship between the energy produced and price becomes obscured once wholesalers are included. From the data, it is unclear which wholesalers buy from which power plants and sell to which utilities. Our main predictor is the proportion of renewable energy power plants that utilities own. We define renewable energy power plants to be those that use solar, wind, hydro, nuclear, and geothermal sources, as well as those that use solid, liquid, and gaseous biomasses for fuel.

Other predictors that we will include are the NERC (North American Electricity Reliability Corporation) region the utility is based in, the proportion of customers that are industrial, the peak demand of electricity, measured in terawatt-hours, and whether or not the utility is considered small. Each NERC region in the country sets its own regulatory standards and operating procedures, so the NERC region a utility is based in can affect prices greatly. Industrial customers are known to buy electricity at cheaper prices than commercial and residential customers do (Key, 2017), so if a utility sells exclusively to industry, that utility will be able to sell its electricity at lower prices than utilities that only sell to commercial and residential customers. The peak demand of electricity for a utility represents the maximum energy all of its customers want at once. If the peak demand is high, we theorize that the utility will be able to charge higher prices for its energy since it would be seen as a more valuable commodity. Finally, small-scale utilities are only required by the EIA to fill out a short form for the Form 861 survey. In general, the EIA defines these entities to be those that have annual retail sales of 200,000 megawatt-hours or less. We suspect that such entities, due to their small size, often have to charge higher rates for electricity since they do not sell as much electricity as larger entities do.

EDA

Figure 1: Average Price





When examining the plots above, we notice that the distribution of average prices of electricity for utilities is skewed to the right. If ignoring the skew, the distribution looks roughly normal, and the average price seems to lie around 0.10 dollars per kilowatt-hour. When plotting the average price against the proportion of renewable power plants for each utility, we notice that there is a slight negative relationship between the two variables. This provides some evidence that average price and the proportion of renewable energy power plants are negatively correlated.

Methodology

Based on our EDA, we will fit an ordinary least squares regression model for average price. We chose this type of model because when examining our EDA, average price had a slightly negative relationship with the proportion of renewable energy power plants. However, in our EDA, we also noticed a right skew in our data. To deal with this, we considered applying a log or square-root transformation to our response variable. However, because the main goal of this analysis is inference, having a model that is easy to understand is more imperative. We applied log and square-root transformations as a part of a sensitivity analysis, but we found similar results (See Appendix B,C). Only for a log transformation did we find that one of our other variables becomes significant. In the end, we decided to continue using the untransformed average price, which we find to be sufficient for our purposes.

We will use the following model:
$y_i = \beta_0 + \beta_1 NERCRegion_i + \beta_2 ProportionRenewablePowerPlant + \beta_3 PeakSummerDemand_i + \beta_4 ProportionIndustrialCustomer_i + \beta_5 IsASmallScaleUtility_i$

In the above model, y is the average price electricity sold by the i^{th} electrical utility. It is important to note that because there are nine NERC regions in the data, $NERCRegion_i$ is a vector of eight indicators for whether or not the utility is based in the corresponding region, while β_1 is a vector of the eight coefficients corresponding to each of the eight variables in $NERCRegion_i$.

A ordinary least-squares regression model assumes that a linear relationship exists between the response variable and a linear combination of the predictor variables as well as that the variance of the model residuals is constant for all sets of values for the predictor variables. To verify that these assumptions are reasonable, we plot the residuals of the model against the fitted values (See Appendix A). Another assumption we make is that the model residuals form a normal distribution around the mean of the response, which we will verify using a histogram of the model residuals (See Appendix A). While we noticed some potential issues, overall, these assumptions appear to be satisfied. The last assumption is that there is independence between observations, which is satisfied because utilities that own power plants and sell to end users are not likely to sell the electricity they produce between each other.

Results

Variable	**Coefficient**	**Standard Error**	**95% CI**	**p-Value**
Intercept	0.106	0.006	(0.093, 0.118)	<0.001
Proportion of Renewable Power Plants	-0.024	0.006	(-0.037, -0.012)	<0.001
WECC Region	0.014	0.008	(-0.001, 0.029)	0.057
RFC Region	-0.001	0.008	(-0.016, 0.014)	0.899
NPCC Region	0.049	0.010	(0.030, 0.068)	<0.001
MRO Region	-0.006	0.007	(-0.020, 0.008)	0.382
SPP Region	-0.005	0.008	(-0.021, 0.010)	0.522
TRE Region	-0.004	0.018	(-0.039, 0.031)	0.840
FRCC Region	0.009	0.016	(-0.022, 0.040)	0.586
ASCC Region	0.168	0.012	(0.146, 0.191)	<0.001
Peak Demand	-0.365	0.489	(-1.325, 0.595)	0.455
Proportion of Industrial Customers	-0.075	0.014	(-0.104, -0.047)	<0.001
Is a Small-Scale Utility	0.025	0.005	(0.016, 0.034)	<0.001

Table 1: Average Price model

Discussion

From our results, we reach the conclusion that, at a significance level of 0.05, the average price of electricity sold by a utility is significantly associated with the proportion of renewable energy power plants run by that utility. We find that for every one percentage point increase in the proportion of renewable energy power plants, the average price of electricity decreases by 0.024 dollars, or 2.4 cents, when holding all else constant. These results are encouraging to policy makers who are looking to encourage the construction of and switch to renewable energy sources. This means that electrical utilities who want to construct more renewable energy power plants and retire nonrenewable ones will become more price-competitive, which will make them more appealing to consumers. Unlike the IRENA report, which solely focused on solar, wind, and coal power, our results encapsulate all renewables and nonrenewables in general, showing that on average, owning renewable energy power plants will lead to lower prices for consumers than owning nonrenewable energy power plants.

In addition to the proportion of renewable energy power plants run by a utility, we found that the NPCC and ASCC regions are positively correlated with average price. The NPCC region, or Northeast Power Coordinating Council, covers much of New York, New England, and eastern Canada, so the higher prices could be a result of the highly developed metropolises and high standards of living in the area. The ASCC region, or Alaska Systems Coordinating Council, oversees the two grids in Alaska state. Unlike the continental states, Alaska's electrical system is not connected to any large, interconnected grid (Alaska, 2021). The lack of electrical infrastructure, as well as the low energy consumption and population, means that Alaska's electricity prices are often three to five times higher in its rural areas than in its urban areas, which would affect the average price of electricity in the state.

We also found that the proportion of industrial customers is negatively correlated with average price, while being a small-scale utility is positively correlated with average price. These results fall in line with what we predicted. We did not, however, find that the peak demand of electricity has a significant relationship with average price. This would indicate that the demand for electricity is price inelastic, perhaps because electricity is necessity for households, industry, and businesses in the United States.

Some of the limitations of this analysis stem from the way the data was collected. Because financial data is reported to the EIA at the utility level, it is unclear how much revenue each individual power plant generates. In the future, we could try to collect more detailed financial data that would allow us to determine these quantities, which would let us determine the average price of electricity sold by each power plant. Another limitation is the fact that many utilities that generate electricity do not end up selling the electricity to end users of that power, but to electric utility wholesalers, who act as middlemen that transport and resell the electricity to other electric utilities. Any future analysis would try to estimate the price of the electricity sold to wholesalers and incorporate that information.

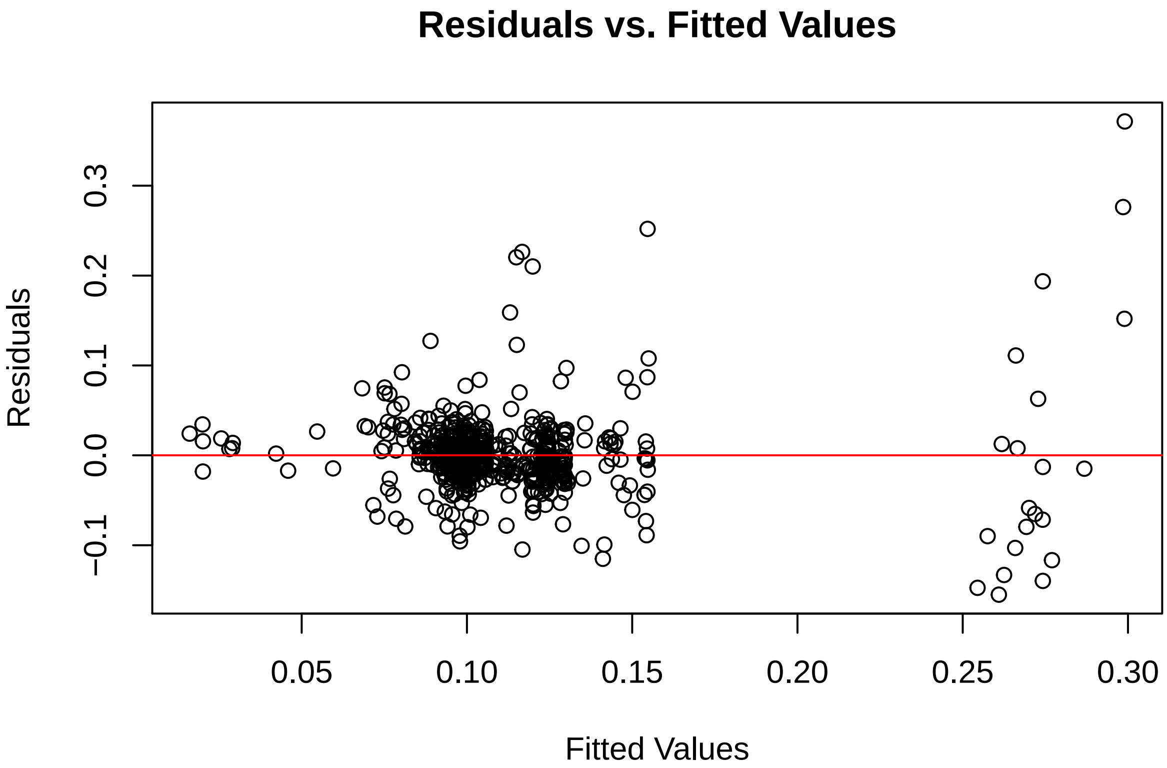
We also recognize that the results of this analysis can only be generalized to electric utilities in the U.S. Considering how climate change is a global issue, we would like to determine if we would observe the same results when applied to the global energy market. Another interesting point of analysis is to look at other economic indicators, such as the amount of profit or the rate of return for an electric utility. This would help us see if renewables can still outperform nonrenewables in other measures of economic performance.

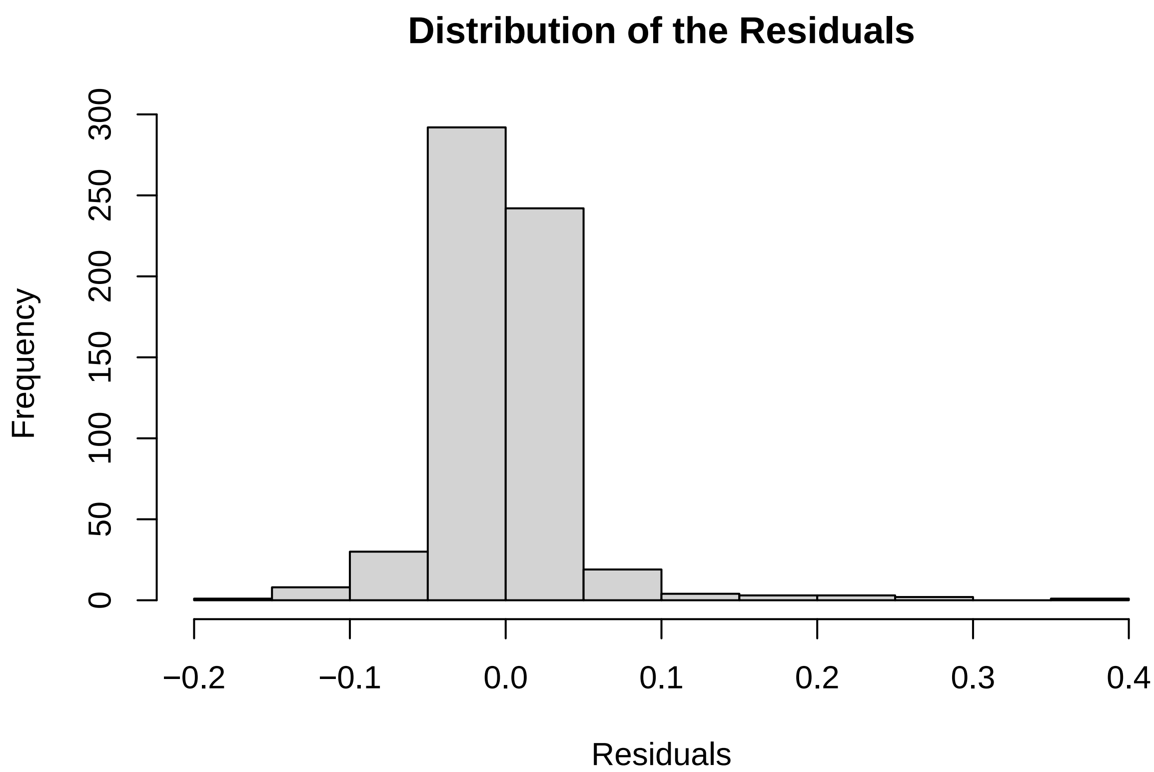
Another avenue for future work is to fit a Bayesian model to see if it is more suitable for this data. Because the distribution of average prices skews to the right, incorporating a prior for the variance of the distribution could make our model more accurate.

Appendix

A. Model Assumptions and Diagnositics

We notice that there are two clusters of points in the plot of residuals vs. fitted values. Also, the distribution of the residuals is somewhat skewed to the right. However, we think that this is likely due to how the average price is skewed to the right, which influenced the fitted values in our model. Overall, the model assumptions seem to be satisfied.





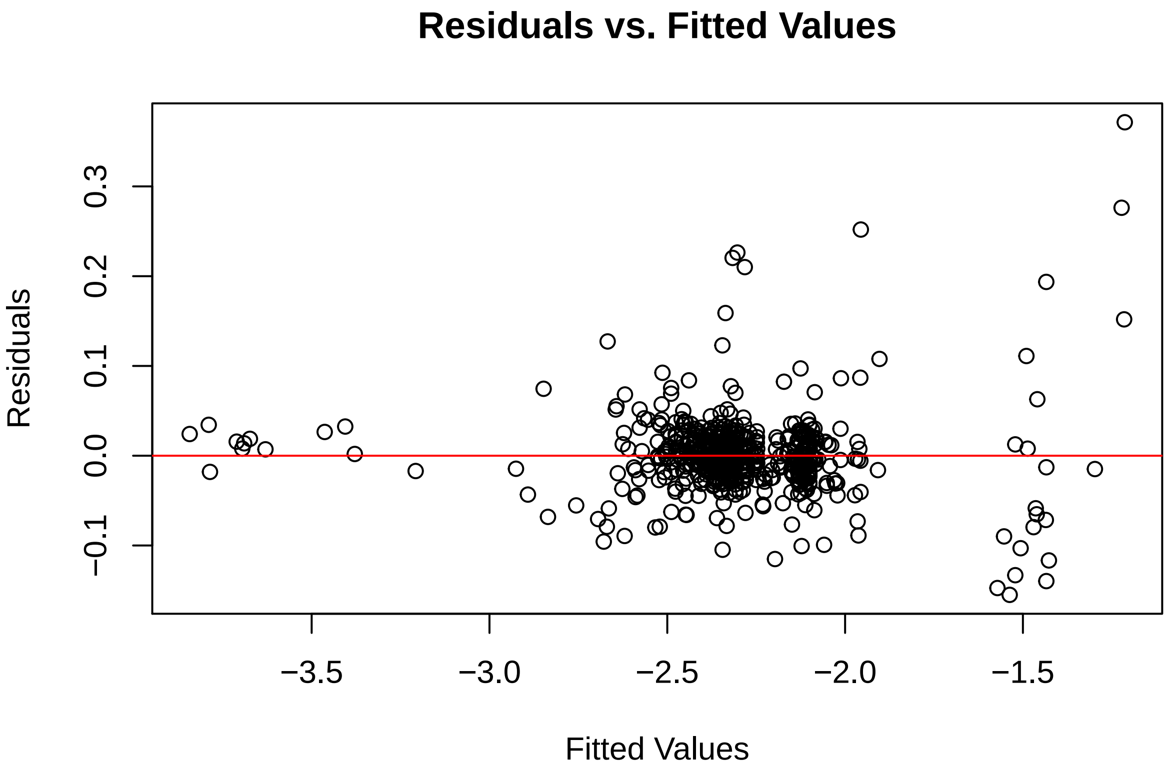
B. Applying a Log Transformation to Average Price

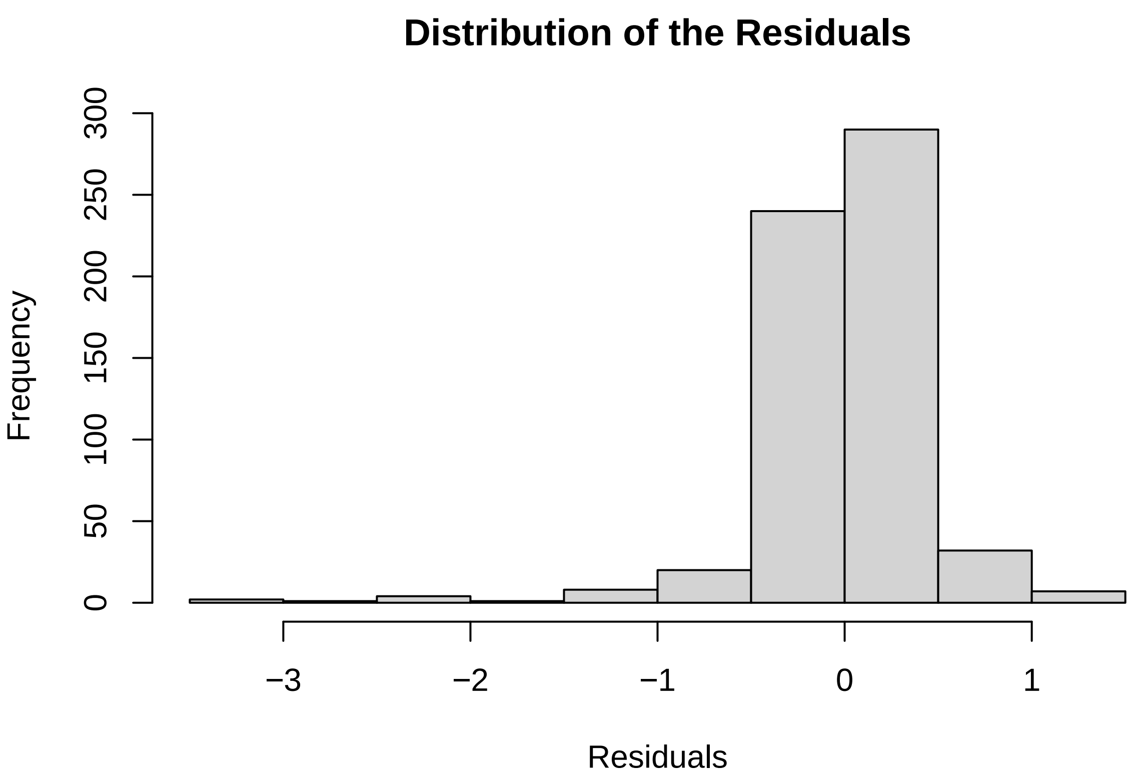
After applying a log transformation to our data, we find that all the variables that were significant in our average price model are significant again, with the notable addition of the peak demand variable. These results suggest that our model is robust.

We also notice that points in the plot of residuals vs. fitted values are clustered towards the center. Also, the distribution of the residuals is somewhat skewed to the left. These plots demonstrate some improvement over the untransformed average price.

Variable	Coefficient	Standard Error	95% CI	p-Value
Intercept	-2.247	0.062	(-2.370, -2.123)	<0.001
Proportion of Renewable Power Plants	-0.168	0.061	(-0.290, -0.046)	0.007
WECC Region	-0.030	0.074	(-0.177, 0.116)	0.685
RFC Region	-0.098	0.075	(-0.246, 0.049)	0.190
NPCC Region	0.291	0.094	(0.106, 0.476)	<0.001
MRO Region	-0.074	0.070	(-0.213, 0.064)	0.292
SPP Region	-0.057	0.078	(-0.212, 0.097)	0.466
TRE Region	-0.041	0.177	(-0.389, 0.308)	0.820
FRCC Region	0.105	0.158	(-0.205, 0.416)	0.506
ASCC Region	0.813	0.115	(0.587, 1.039)	<0.001
Peak Demand	-10.421	4.854	(-19.955, -0.887)	0.032
Proportion of Industrial Customers	-1.339	0.143	(-1.620, -1.058)	<0.001
Is a Small-Scale Utility	0.221	0.046	(0.130, 0.311)	<0.001

Table 2: Log Transformation of Average Price





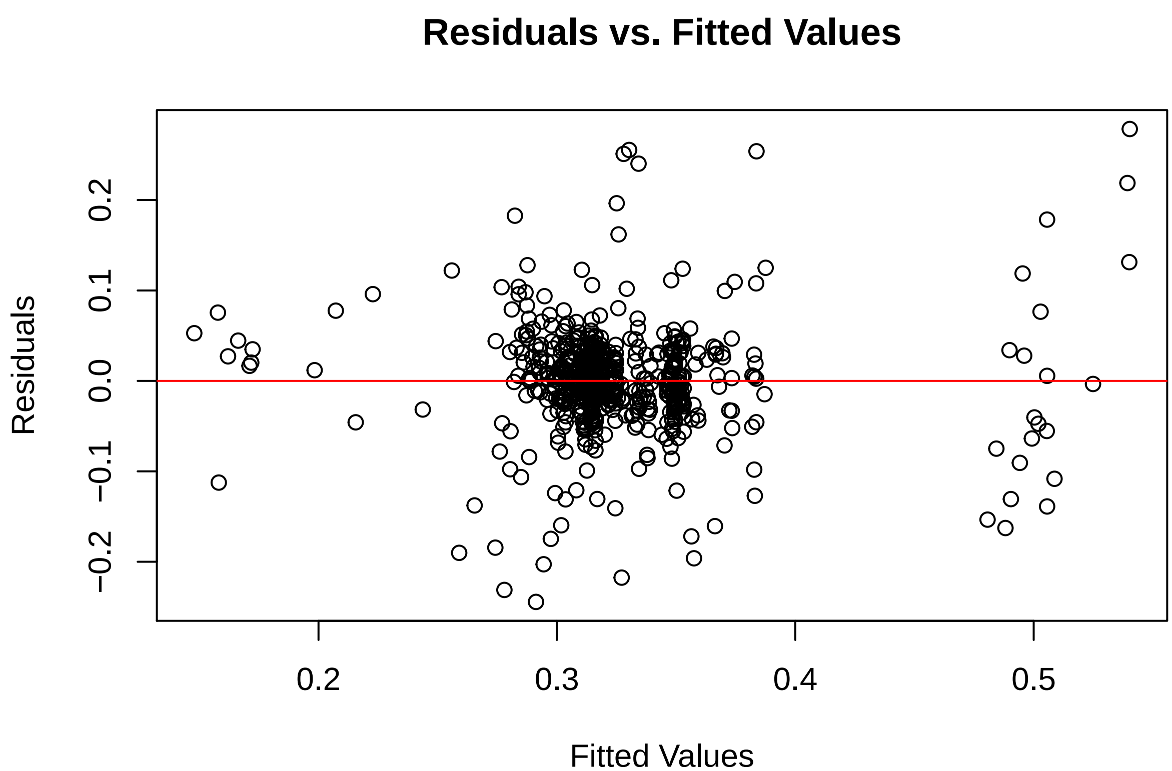
C. Applying a Square-Root Transformation to Average Price

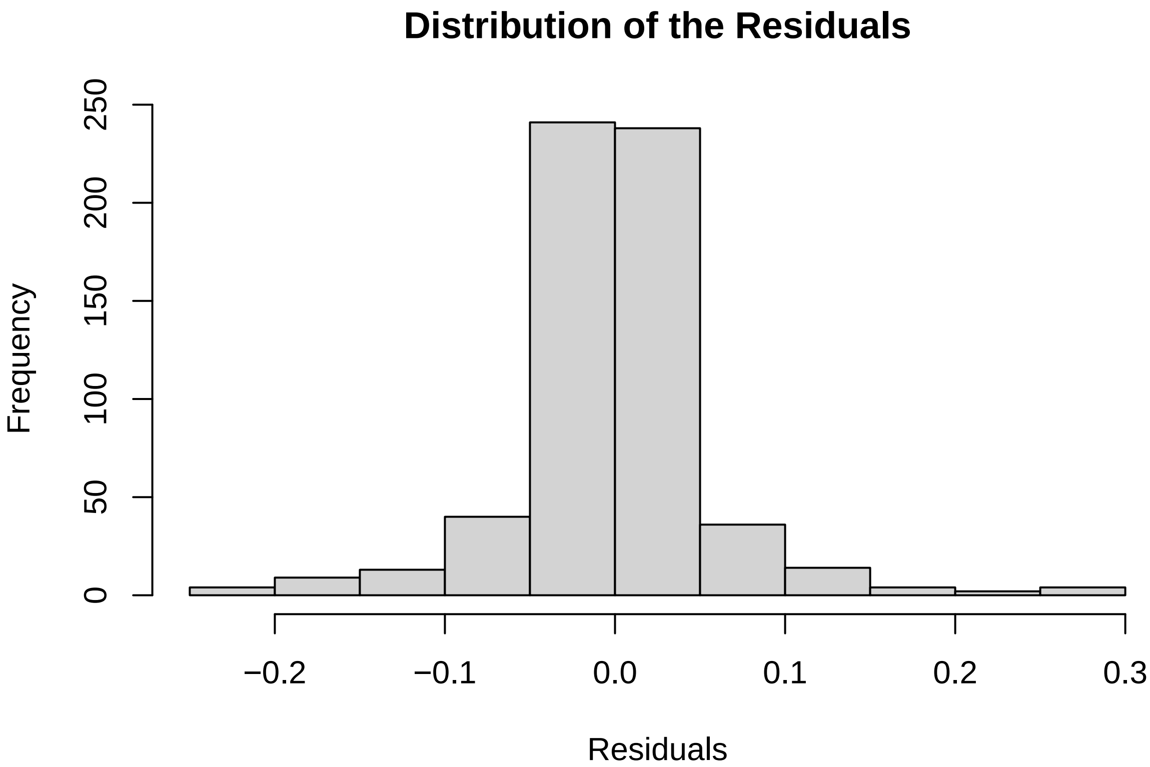
After applying a log transformation to our data, we find that all the variables that were significant in our average price model are significant again. These results suggest that our model is robust.

We also notice that points in the plot of residuals vs. fitted values are clustered towards the center. These plots demonstrate soem improvement over the untransformed average price.

Variable	Coefficient	Standard Error	95% CI	p-Value
Intercept	0.325	0.008	(0.308, 0.341)	<0.001
Proportion of Renewable Power Plants	-0.031	0.008	(-0.047, -0.014)	< 0.001
WECC Region	0.010	0.010	(-0.010, 0.029)	0.316
RFC Region	-0.006	0.010	(-0.026, 0.013)	0.527
NPCC Region	0.059	0.013	(0.034, 0.083)	< 0.001
MRO Region	-0.010	0.009	(-0.028, 0.008)	0.283
SPP Region	-0.008	0.010	(-0.029, 0.012)	0.430
TRE Region	-0.006	0.024	(-0.052, 0.040)	0.794
FRCC Region	0.015	0.021	(-0.026, 0.056)	0.473
ASCC Region	0.181	0.015	(0.151, 0.211)	< 0.001
Peak Demand	-1.069	0.644	(-2.333, 0.195)	0.097
Proportion of Industrial Customers	-0.146	0.019	(-0.183, -0.108)	< 0.001
Is a Small-Scale Utility	0.035	0.006	(0.023, 0.047)	< 0.001

Table 3: Square-Root Transformation of Average Price





References

"Alaska". Energy Information Administration, 21 January 2021, https://www.eia.gov/state/analysis.php?sid=AK#:~:text=Electricity,-Alaska%20ranked%20second&text=In%202019%2C%20natural%20gas%20fueled,for%203%25%20of

"Energy and Climate Change". European Environment Agency, 23 November 2020, https://www.eea.europa.eu/signals/signals-2017/articles/energy-and-climate-change.

"8 Key Factors That Impact Electricity Prices". ElectricChoice.com, 2017, https://www.electricchoice.com/blog/8-key-factors-impact-electricity-prices/

Prehoda, Emily, Joshua M. Pearce, and Chelsea Schelly. "Policies to Overcome Barriers for Renewable Energy Distributed Generation: A Case Study of Utility Structure and Regulatory Regimes in Michigan". Michigan Technological University, 20 February 2019, https://digitalcommons.mtu.edu/cgi/viewcontent.cgi?article=1181&context=social-sciences-fp

"Renewables Increasingly Beat Even Cheapest Coal Competitors on Cost". International Renewable Energy Agency, 2 June 2020, https://www.irena.org/newsroom/pressreleases/2020/Jun/Renewables-Increasingly-Beat-Even-Cheapest-Coal-Competitors-on-Cost#:~:text=Electricity%20costs%20from%20utility%2Dscale,USD%200.115%2FkW