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

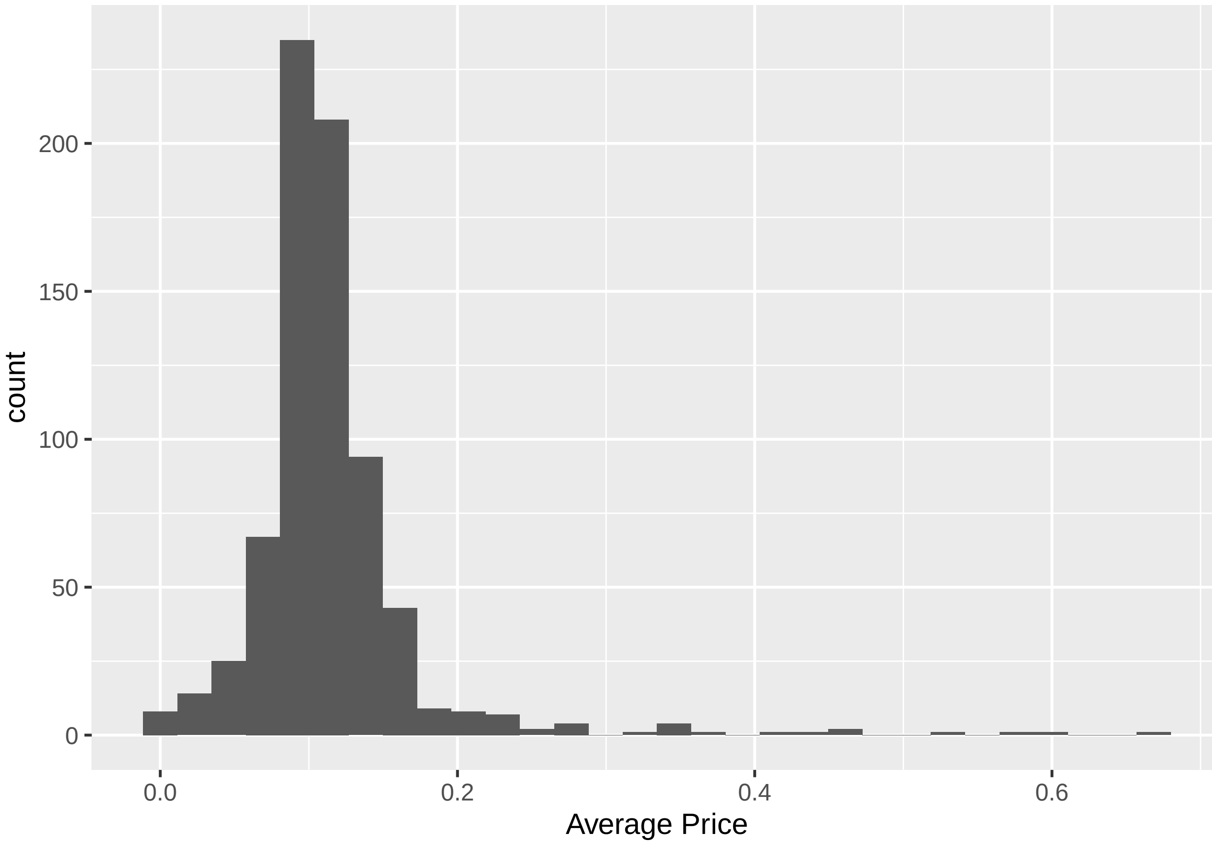
The data we will use for this analysis comes primarily from the US Energy Information Administration (EIA), a government agency that gathers data about the country's operational power plants. The EIA collects generator-specific data on the electricity generated throughout the nation through surveys each 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.

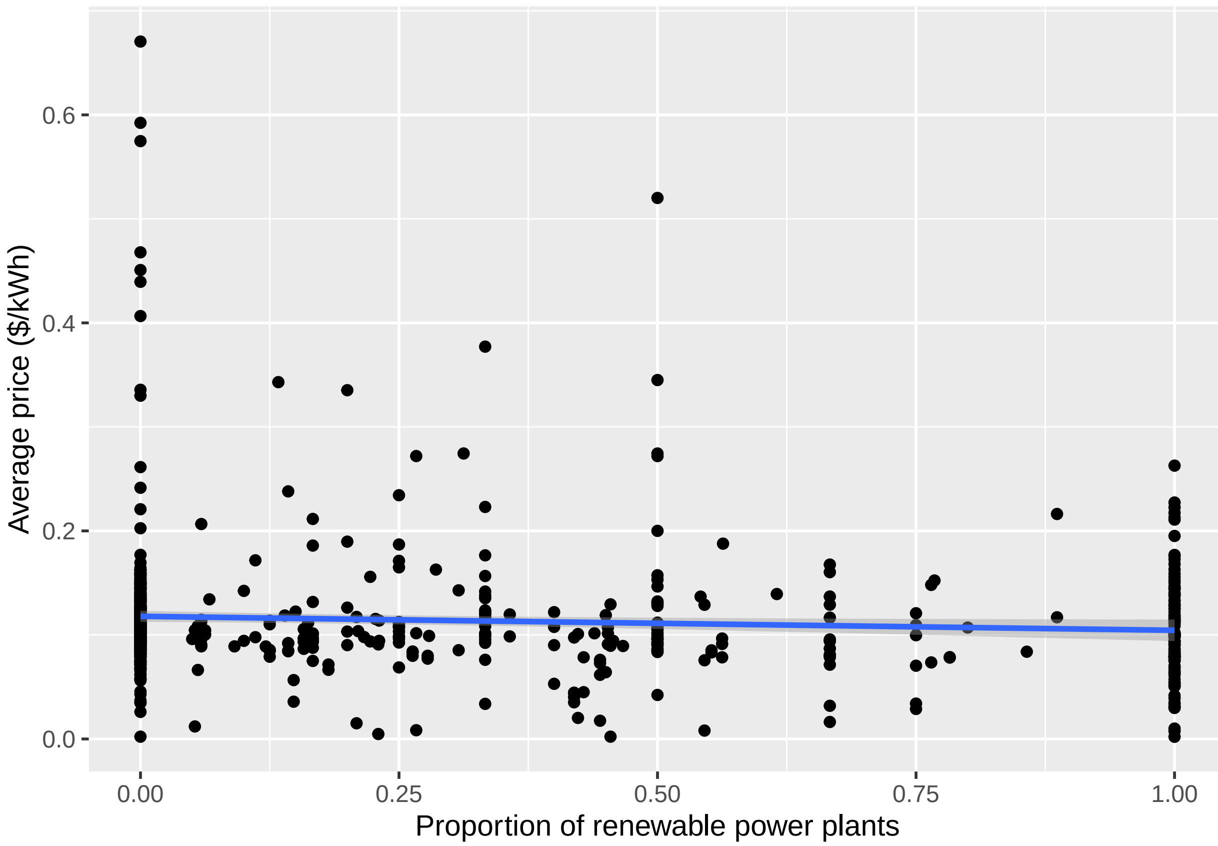
Variables

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. 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 most energy all of its customers want. If the peak demand is high, it is likely the utility will be able to charge higher prices for its energy. 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 don't sell as much electricity as larger entities do.

EDA





Methodology

To fit these models, we will fit ordinary least squares regression models. We chose this type of model because when examining our EDA, the response variable may have a linear relationship with our predictor variables. Other models that were considered include a Bayesian model, where we would incorporate some prior belief into our model.

To find the impact of renewables on the average price of electricity, we will use the following model:
$y_i = \beta_0 + \beta_1 NERCregion_i + \beta_2 PeakSummerDemand_i + \beta_3 ProportionIndustrialCustomer_i + \beta_4 SmallScaleUtility_i$

In the above model, y is the average price of electricity for the i^{th} utility.

To find the impact of renewables on the net electrical output, we will use the following model:
$y_i = \beta_0 + \beta_1 NERCregion_i$

In the above model, y is the electricity generated for the i^{th} power plant.

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. 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 Normal QQ-plot. The last assumption is that there is independence between observations.

We plan to try other models to ensure that our findings are robust.

Unfortunately, because the same company could have built many of the same kind of power plant, we cannot be sure that power plants will be independent of each other. However, for sensitivity analysis, we will try other models to ensure that our findings are robust.

```
## [1] 9776
```

```
## [1] 3177
```

```
## [1] 4570
```

```
## [1] 841
```

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

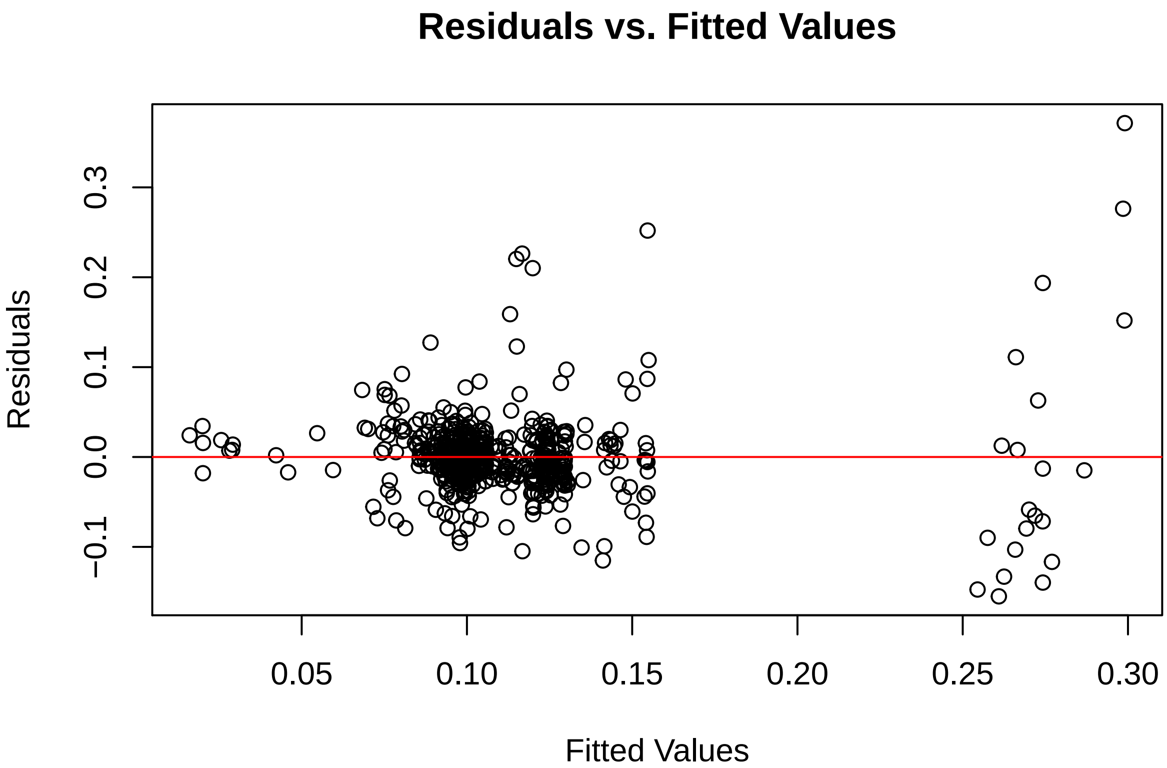
From our results, we reach the conclusion that, at a significance level of 0.05, the average price of electricity sold by a utility is significantly associated with the proportion of renewable energy power plants run by that utility. We find that for every one percentage point increase in the proportion of renewable energy power plants, the average price of electricity decreases by (INSERT NUMBER). 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.

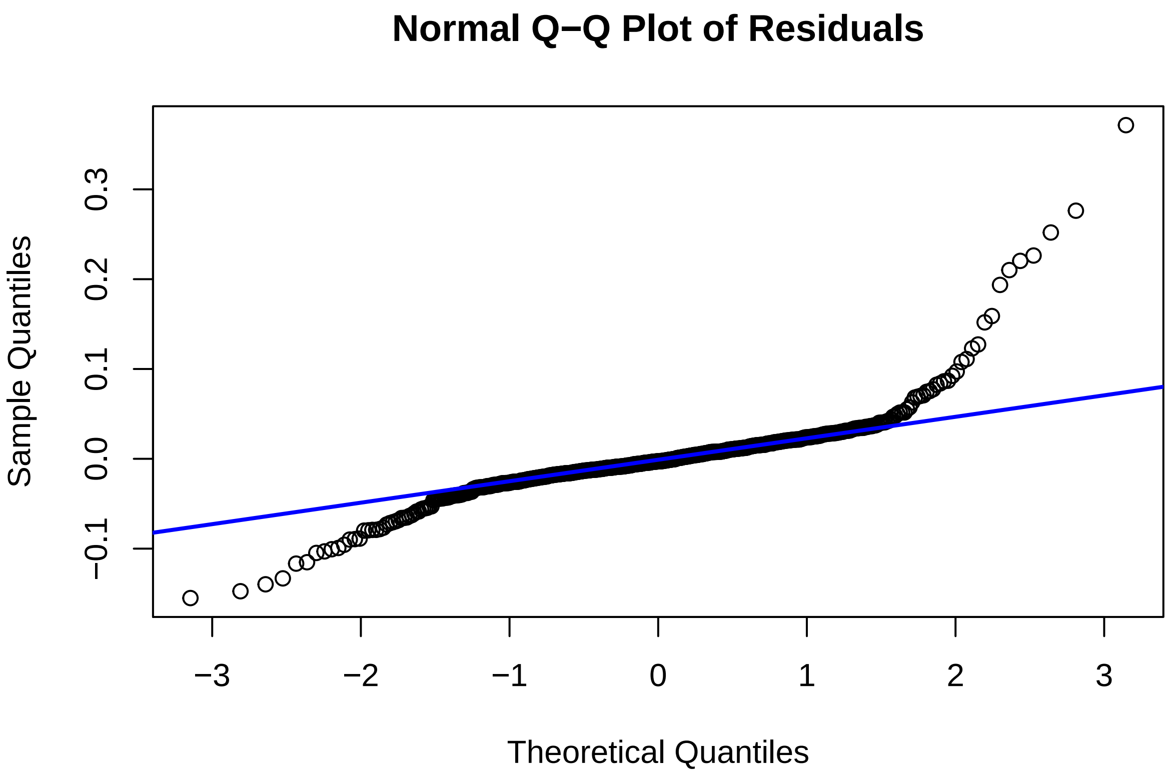
Some of the limitations of this analysis stem from the way the data was collected. Because financial data is reported to the EIA at the utility level, it is unclear how much revenue each individual power plant generates. In the future, we could try to collect more detailed financial data that would allow 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





References

"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