Aidan Cassel-Mace
Machine Learning, Text Analysis and Economics
First Project

Exploring Nonlinearity for the Landlord-Tenant Problem

Introduction

Economics literature has identified an inefficiency in the cost structure of utilities for rental residences, known as the landlord-tenant problem. If the tenant is responsible for paying the price of utilities, the landlord has minimal incentive to invest in energy efficient appliances or insulation. Conversely, if the landlord is responsible for utility costs, the tenant has no incentive to ration energy. Papers such as "Energy use by apartment tenants when landlords pay for utilities" by Levinson and Nieman have found empirically that these effects exist (Levinson and Niemann 2004).

This project will seek to understand the relationship between a tenant's price of energy and their lease structure (whether they are responsible for paying for energy) on their energy consumption. Particular focus will be paid to non-linear relationships between price consumption to understand how it might be possible to optimally leverage lease structures to promote energy conservation.

Data Sources and Descriptions

To analyze the relationship between price, lease structure, and energy consumption, I use data from the 2015 Residential Energy Consumption Survey (RECS). The dataset is published publicly by the U.S. Energy Information Administration. It includes a representative sample of over 5,000 residences throughout the United States, and it includes 763 descriptors for each residence from their total energy use to the Energy Star status of each major appliance. Together, the 763 data points on each residence detail the residence's energy set-up and consumption for one year.

Unfortunately, the subset of data I am able to analyze is much smaller. Only a subset of the 2015 RECS dataset is about tenants, with the rest being homeowners. Furthermore, the dataset includes many imputed data points, flagged accordingly. Taking only observations corresponding to tenants with all essential data present, my analysis is left with a sample size of 569 residences. Of these, 134 have their natural gas included in rent, while 435 pay for their own natural gas. Appendix A shows the percentage of residences located in each of the nation's 10 census divisions for the data I use in my analysis and for the RECS data as a whole. Certain regions such as the East North Central are significantly more represented in the used subset than in the full RECS dataset, while others such as the South Atlantic. These differences likely reflect the differences in the prevalence of rentals between regions. Appendices B and C display the dataset's summary statistics and plot the dataset, respectively.

Methodology:

I explore two relationships in my analysis. The first is the relationship between the *actual* price of natural gas and the amount of natural gas used. The second is the relationship between the *effective* price of natural gas from the tenant's perspective and the amount used. The only difference between these metrics, is that for a tenant whose natural gas is included in their lease, the effective price of natural gas is zero.

I estimate each relationship employing two methods. First, I use OLS to regress the quantity of natural gas used on the price of natural gas (either effective or actual), including a linear, a squared, and a cubed term. I control for household income. More control variables would make the analysis more robust, but due to the small sample size they could cause overfitting. Second, I estimate the relationship between the quantity of natural gas and the price of natural gas using a multilayer perceptron. Again, I control for household income.

Results:

Appendix D tabulates the results of my OLS regressions. Appendix E plots each regression's curve of best fit against the scatterplot of data.

Unsurprisingly, for the actual price regression, I received a negative linear coefficient, suggesting that as prices increase natural gas consumption decreases. Notably, there was also a positive quadratic coefficient, suggesting that as prices continue to increase the marginal reduction in consumption decreases. This result likely reflects decreasing marginal utility for each unit of energy used; while tenants are likely to heat their homes to a certain livable level regardless of the price, additional heating for comfort is likely to be more price elastic.

Interestingly, in the effective price regression we do not see the same clear decline in demand. In fact, the linear coefficient is positive, suggesting that consumption increases as does price. That said, the regression's curve of best fit in Appendix E is generally quite flat. This difference could be reflective of a positive correlation between a landlord paying for utilities and a residence's energy efficiency.

Appendix F reports on the R^2 scores of the multilayer perceptron. Appendix G plots the models' curves of best fit against the scatterplot of data.

After many training experiments with different learning rates, activation functions, numbers of hidden layers, hidden layer sizes, my multilayer perceptrons were unable to pick up on trends beyond the mean. As seen in Appendix F, both perceptions have very near-zero R^2 scores, suggesting that they were unable to represent a significant pattern beyond the mean of the training data.

Conclusion:

While my results do not offer a solution to the tenant-landlord problem, they provide insight into the relationships at play. There is some evidence that consumption decreases as does price for the dataset as a whole, which is made up mostly of tenants who pay for their own energy. However, if we consider the *effective* price tenants pay, we no longer find the same relationship. It is possible that tenants who do not pay for their own utilities tend to live in more energy efficient homes, and by consolidating them at the bottom of the price spectrum we counteract the demand effect.

One notable result of my analysis is that there is a decaying decline in natural gas consumption as price increases. This observation could warrant the exploration of more nuanced pricing schemes such that the landlord and tenant each have some price burden and some incentive to invest in efficiency and conserve energy. Previous work such as "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing" by Koichiro Ito explores and ultimately does not recommend nonlinear utility pricing, but schemes where landlords and utilities share the cost of utilities have been less researched (Ito 2014).

Future work should consolidate RECS data over many years or merge with other data sources to compile a larger sample size. My analysis is subject to omitted variable bias, because I was only able to control for one variable without overfitting. With a larger sample size, this analysis could be significantly more robust and reliable.

Bibliography:

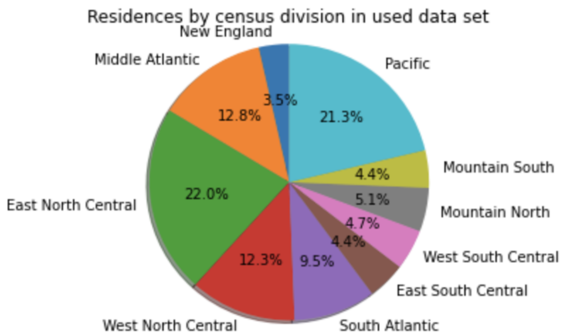
Arik Levinson and Scott Niemann, "Energy use by apartment tenants when landlords pay for utilities." *Resource and Energy Economics* 26 (2004): 51–75.

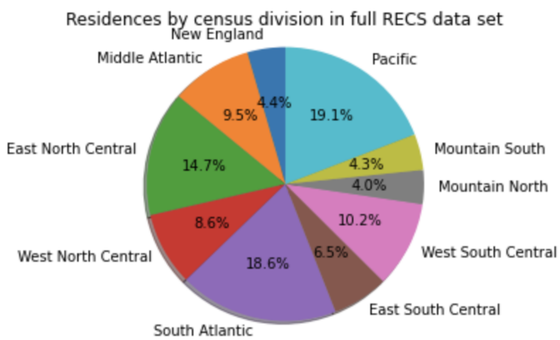
Koichiro Ito, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review* (2014): 537–563.

"Residential Energy Consumption Survey (RECS)." *U.S. Energy Information Administration* (2015).

Appendix:

Appendix A: Data by census division

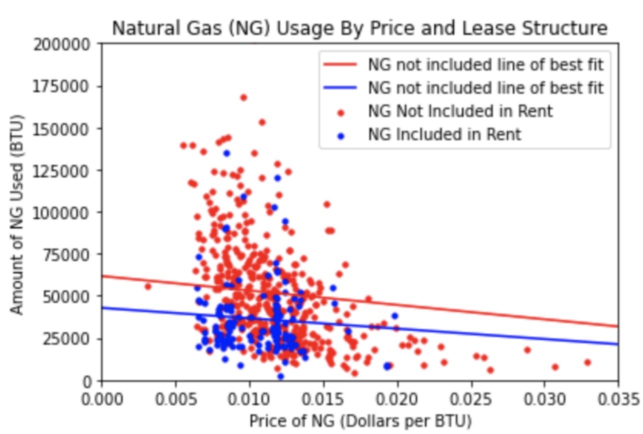




Appendix B: Data summary statistics

	NG Included	Count	Mean	STD
Amount of NG Used (BTU)	Yes	134	36312.7	21472.8
NG Price (Dollars/BTU)	Yes	134	0.0105331	0.0026163
Amount of NG Used (BTU)	No	435	51147	33794.7
NG Price (Dollars/BTU)	No	435	0.0123255	0.0102557

Appendix C: Data distribution



Appendix D: OLS regression results

Regression on Effective Price

OLS Regression Results

Dep. Variable:	BTUNG	**R-squared:**	0.030
Model:	OLS	**Adj. R-squared:**	0.023
Method:	Least Squares	**F-statistic:**	4.379
Date:	Thu, 31 Mar 2022	**Prob (F-statistic):**	0.00171
Time:	16:46:00	**Log-Likelihood:**	371.59
No. Observations:	569	**AIC:**	-733.2
Df Residuals:	564	**BIC:**	-711.5
Df Model:	4		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
CONST	0.1870	0.011	17.229	0.000	0.166	0.208
EFFNGPRICE	0.2213	0.245	0.905	0.366	-0.259	0.702
EFFNGPRICESQ	-4.1641	1.396	-2.982	0.003	-6.907	-1.422
EFFNGPRICECU	3.7856	1.245	3.041	0.002	1.341	6.230
MONEYPY	0.0244	0.020	1.198	0.231	-0.016	0.064

Omnibus:	225.793	**Durbin-Watson:**	2.082
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	949.787
Skew:	1.786	**Prob(JB):**	5.71e-207
Kurtosis:	8.225	**Cond. No.**	364.

Regression on Actual Price

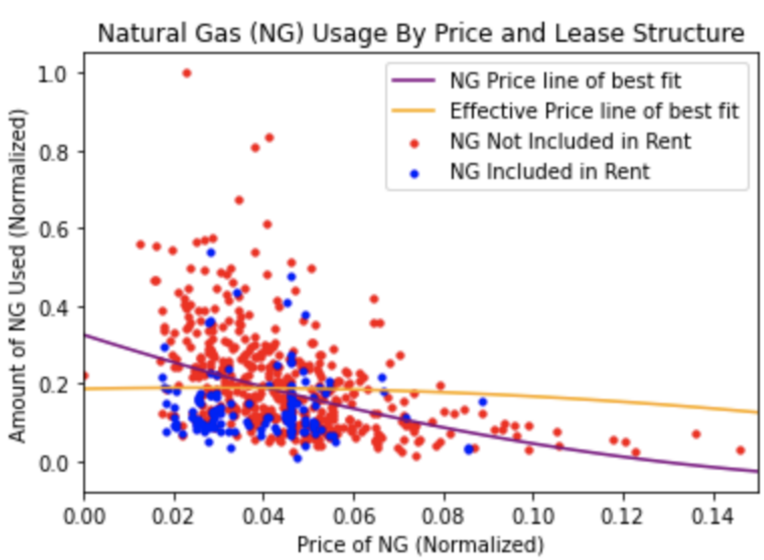
OLS Regression Results

Dep. Variable:	BTUNG	**R-squared:**	0.164
Model:	OLS	**Adj. R-squared:**	0.159
Method:	Least Squares	**F-statistic:**	27.76
Date:	Thu, 31 Mar 2022	**Prob (F-statistic):**	4.64e-21
Time:	16:46:01	**Log-Likelihood:**	414.01
No. Observations:	569	**AIC:**	-818.0
Df Residuals:	564	**BIC:**	-796.3
Df Model:	4		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
CONST	0.3258	0.016	20.550	0.000	0.295	0.357
NGPRICE	-3.7030	0.401	-9.246	0.000	-4.490	-2.916
NGPRICESQ	9.0820	1.682	5.400	0.000	5.779	12.385
NGPRICECU	-5.7082	1.386	-4.118	0.000	-8.431	-2.986
MONEYPY	0.0315	0.019	1.668	0.096	-0.006	0.068

Omnibus:	225.113	**Durbin-Watson:**	2.098
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	1079.204
Skew:	1.723	**Prob(JB):**	4.51e-235
Kurtosis:	8.800	**Cond. No.**	460.

Appendix E: Curves of best fit of OLS regressions



Appendix F: R^2 for multilayer perceptrons

Effective Price Training R Squared	-0.009235217065662438
Effective Price Testing R Squared	-0.06350978407537688
Actual Price Training R Squared	-0.0043851719182594895
Actual Price Testing R Squared	-0.0030468129243985675

Appendix G: Curves of best fit for multilayer perceptrons