Homework 2

Lin Yang

February 1, 2023

1 Python

1.1 Check for balance between the treatment and control groups

	Control Mean (s.d.)	Retrofit Mean (s.d.)	Difference p value of t-test
Electricity	1181.33 (454.31)	1086.75 (423.96)	0.001
Sqft	1633.05 (682.90)	1657.55 (686.27)	0.572
Temp	79.89 (2.16)	79.89 (1.97)	0.987

Table 1: Summary Statistics by control and treatment group

The characteristics of Sqft and Temp are good measures for the households since the p-values of the two variables in the t-test is large, which means that the average Sqft and Temp among two groups are similar. And the p-value of Electricity is less than 1%, which means there is significant difference in Electricity use between two groups and the randomized experiment succeeded in creating the desired variation.

1.2 Kernel Density Plots

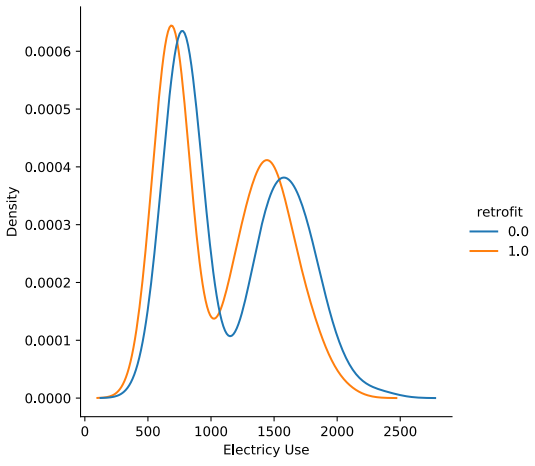

Figure 1: Kernel density plots of electricity use by control and treatment group

1.3 Estimation

(a) OLS by hand

	OLS (by hand) Coefficients
Constant	-83.602758
Sqft	0.615339
Retrofit	-109.666176
Temp	3.255075

Table 2: Coefficients from calculating OLS by hand

(b) OLS by simulated least squares

	OLS (by simulated least squares) Coefficients
Constant	-83.556600
Sqft	0.615338
Retrofit	-109.666059
Temp	3.254502

Table 3: Coefficients from simulated least squares

(c) OLS using a canned routine

	OLS (by StatsModels) Coefficients
Constant	-83.602758
Sqft	0.615339
Retrofit	-109.666176
Temp	3.255075

Table 4: Coefficients from using a canned routine

2 Stata

2.1 Check for balance between treatment and control group.

	Control mean/sd	Retrofit mean/sd	Difference p
electricity	1181.33	1086.75	0.001***
	454.31	423.96	
sqft	1633.05	1657.55	0.572
	682.90	686.27	
temp	79.89	79.89	0.987
	2.16	1.97	
Observations	501	499	1000

Table 5: Summary Statistics by control and treatment group

2.2 Scatter Plot with electricity consumption vs sqft

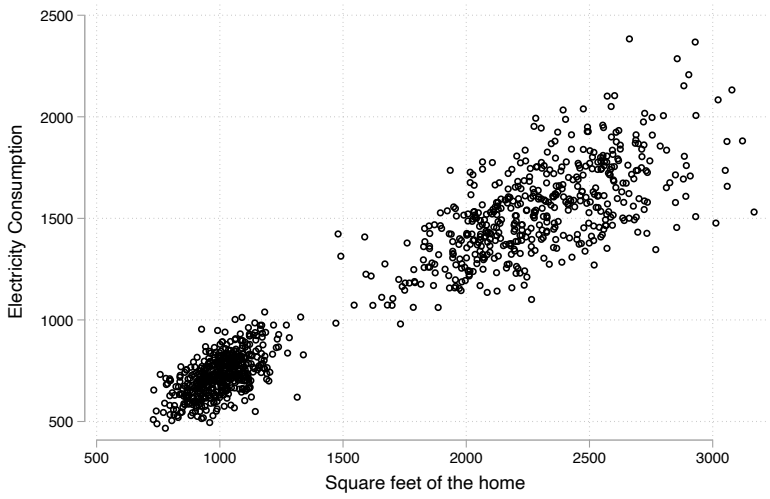

Figure 2: Scatter plot of electricity consumption versus square feet

2.3 Regression Results

VARIABLES	(1) Ordinary least squares
sqft	0.62**
	(0.01)
retrofit	-109.67**
	(7.94)
temp	3.26
	(1.93)
Constant	-83.60
	(154.69)
Observations	1,000
R-squared	0.92

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 6: Regression results table with heteroskedasticity standard errors