Economics 7103 - Homework 2

Ana Mazmishvili

January 29, 2024

Python

Note: *While working on this HW, I got help from Afi only with some Python codes.*

Question 1.1

Response: Randomization worked which is demonstrated by comparison across control and treatment groups that indicates statistical balance in observables. Column 3 presents the differences in means and the standard errors of the differences in brackets. The differences are small and in case of electricity consumption statistically significant.

	Control	Treatment	P-value
Monthly electricity usage by HHs (kWh)	1181.33	1086.75	0.001
	454.31	423.96	[3.403]
Square feet of home	1633.05	1657.55	0.572
	682.90	686.27	[-0.566]
Outdoor average temperature (°F)	79.89	79.89	0.987
	2.16	1.97	[-0.016]
Observations	501	499	

Table 1: Summary Statistics for the treated and control groups.

Question 1.2

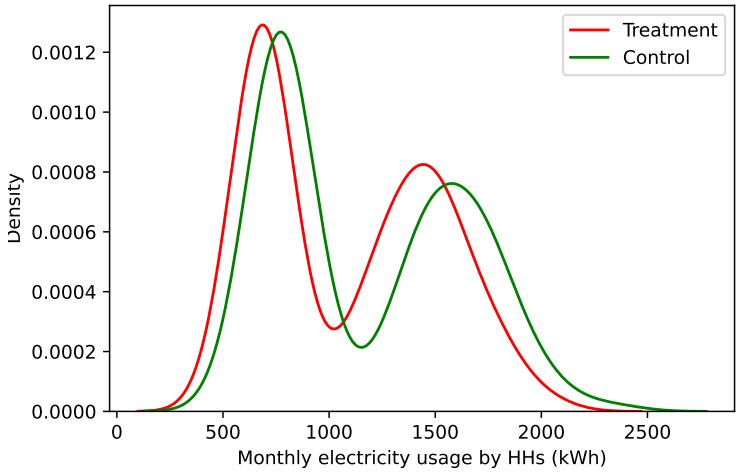

Figure 1: Kernel density plots of the electricity use for treated group and control group.

Question 1.3

(a) I used the Numpy package in Python to create an array X that is the 1000×4 matrix of the predictor variables (3) and a column of ones and an array Y that is the 1000×1 vector of the dependent variable. The codes are provided in the Python code file. I used matrix operations to calculate $\hat{\beta}$. Recall that

$$\hat{\beta} = (X^{'}X)^{-1}X^{'}Y$$

I obtained $\hat{\beta}$ that is a 4x1 vector with the following values [0.615 -109.666 3.255 -83.603]

(b) Missing. I was not able to generate a table including all three regression results. I will try today to fix it. (c) Missing

Stata

Question 2.1

I created a table that displays each variable's sample mean, sample standard deviation, and p-values for the two-way t-test between treatment and control group means. Please see the Table 2

	Control	Treatment	P-value
electricity	1181.33	1086.75	0.001
	(454.31)	(423.96)	[3.404]
sqft	1633.05	1657.55	0.572
	(682.90)	(686.27)	[-0.566]
temp	79.89	79.89	0.987
	(2.16)	(1.97)	[-0.016]
Observations	501	499	1,000

Table 2: Summary statistics produced using Stata

Question 2.2

I created a two-way scatterplot of electricity consumption and square feet of home data using Stata. Please refer to the Figure below.

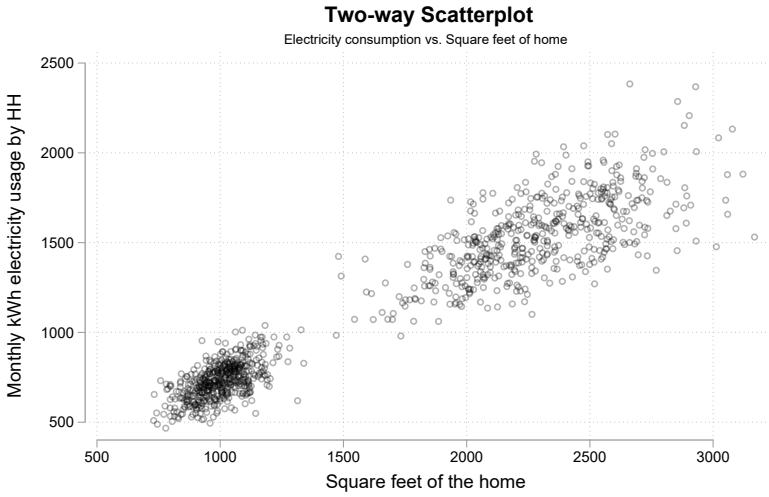

Figure 2: Scatterplot with electricity consumption and square feet of home

Question 2.3

I estimated model using OLS and obtained heteroskedasticity robust standard errors and coefficients. Please refer to the Table below.

[]article

VARIABLES	(1) electricity
retrofit	-109.7***
	(7.943)
sqft	0.615***
	(0.00678)
temp	3.255*
	(1.932)
Constant	-83.60
	(154.7)
Observations	1,000
R-squared	0.919

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: OLS regression results using Stata