Homework 2

Economics 7103

Januray 29, 2024

1 Python

1.1 Balance Table

Check for balance in treatment and control group. Does it appear that the randomization worked? If so, what can we say about the simple difference-in-means estimate?

The balance table suggests that the randomization worked well, because the covariates sqft and temp are quite similar for both the treatment and control groups. This tells us that a simple difference-in-means estimator will give us

Variable	Control	Treatment	P-value
sqft	1633.05 ± 682.90	1657.55 ± 686.27	0.571630
temp	79.89 ± 2.16	79.89 ± 1.97	0.987135

Table 1: Balance table from python.

1.2 Kernel Density Plot

Provide graphical evidence that the retrofits worked. Plot kernel density plots of the electricity use for treated group and control group on the same graph using Python. Make sure to label the histogram appropriately.
We can see the retrofits worked because the entire distribution of electricity usage is shifted lower from the group that received retrofits (treatment group).

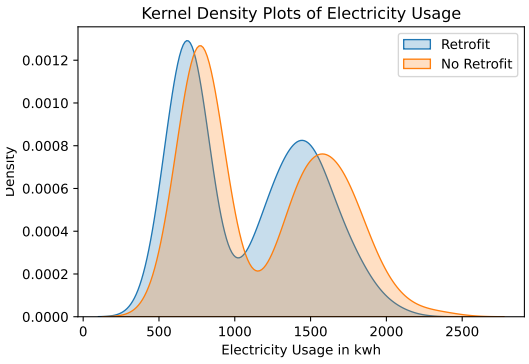

Figure 1: Sample kernel density plot of electricity usage for the treatment and control groups

1.3 OLS

Use the following methods to estimate $\hat{\beta}$, presenting coefficients in a single table with a column for each estimation technique

Method	Intercept	sqft	retrofit	temp
OLS by Hand	-83.602758	0.615339	-109.666176	3.255075
Simulated Least Squares	739.743120	0.611999	-92.653303	-7.080147
StatsModels OLS	[-83.60275758]	[0.61533854]	[-109.66617626]	3.255075

Table 2: Coefficients table from python.

2 Stata

2.1 Balance Table

	(1)		(2)		(3)	
	mean	sd	mean	sd	b	t
sqft	1633.052	682.904	1657.551	686.271	-24.499	(-0.566)
temp	79.891	2.163	79.893	1.968	-0.002	(-0.016)
N	501		499		1000	

Table 3: Balance table from Stata.

2.2 Scatterplot

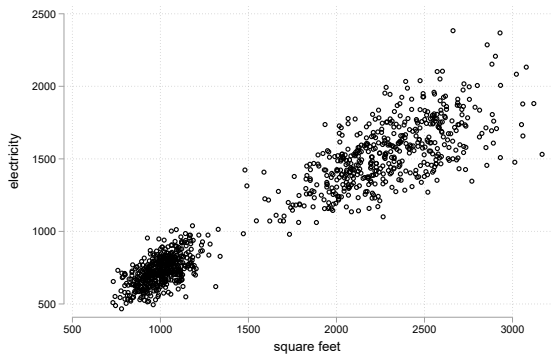

Figure 2: Scatterplot of electricity usage (kwh) and square footage

2.3 Regress

[]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 4: Regression output from Stata.