Homework 2

Ryan Ellis

January 24, 2023

1 Python

1.1 Comparison of means

Variable	Mean(D_{1i})	Mean(D_{0i})	Diff.-in-means (p-val)
Electricity	1086.75	1181.33	0.001
	(423.96)	(454.31)	
Square Footage	1657.55	1633.05	0.572
	(686.27)	(682.90)	
Retrofit	1.00	0.00	0.0
	(0.00)	(0.00)	
Temperature	79.89	79.89	0.987
	(1.97)	(2.16)	

It appears randomization worked here. The home locations (proxied by "Temperature") and sizes ("Square Footage") are similar across treatment and control. The difference in mean outcome, if indeed we have complete random assignment, is the Average Treatment Effect.

1.2 Kernel Density Plot

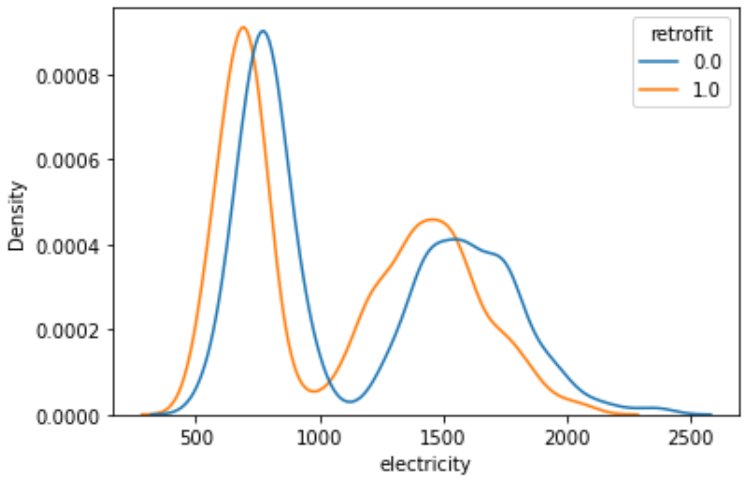

Figure 1: Treated households shown in orange

1.3 OLS Methods

	By hand	Simulation	Canned
Intercept	-83.6028	-83.611	-83.6028
retrofit	-109.666	-109.666	-109.666
sqft	0.615339	0.615339	0.615339
temp	3.25508	3.25518	3.25508

2 Stata

2.1 Comparison of means

	(1)		(2)		(3)			
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	T-stat.	p-value
electricity	1086.75	(423.96)	1181.33	(454.31)			94.584***	(3.404)
sqft	1657.55	(686.27)	1633.05	(682.90)			-24.499	(-0.566)
retrofit	1.00	(0.00)	0.00	(0.00)			-1.000	(.)
temp	79.89	(1.97)	79.89	(2.16)			-0.002	(-0.016)
N	499		501		1000			

Table 3: produced using Stata

2.2 Scatter plot

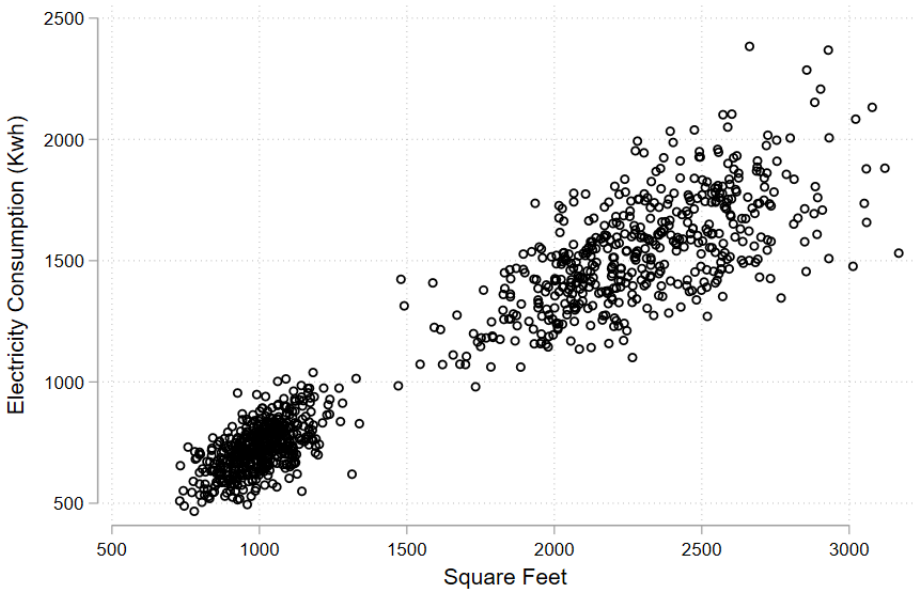

Figure 2: Scatter plot produced in Stata with *twoway*

2.3 Regression

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: produced using Stata