Homework 3

Economics 7103

You have access to imaginary data on an energy-efficiency retrofit program in Atlanta *kwh.csv* (the same as the previous homework) and you are interested in whether the program reduced energy use. In your dataset is the following information: After recruiting the households for the program, you assigned them to

Variable	Description
electricity	kWh of electricity used by the household in the month
sqft	Square feet of the home
retrofit	= 1 if the home received a retrofit
temp	The outdoor average temperature (° F) during the month at the home's location

Table 1: Variable descriptions for homework 3.

treatment and control groups. Treatment homes received the retrofits on the first of the month and control homes did not have any work done.

1 Stata or Python

1. Suppose that for a home i, you think the underlying relationship between electricity use and predictor variables is $y_i = e^{\alpha}\delta^{d_i}z_i^{\gamma}e^{\eta_i}$ where e is Euler's number or the base of the natural logarithm, d_i is a binary variable equal to one if home i received the retrofit program, z_i is a vector of the other control variables, η_i is unobserved error, and $\{\alpha, \delta, \gamma\}$ are parameters to estimate.

   (a) Show that $ln(y_i) = \alpha + ln(\delta)d_i + \gamma ln(z_i) + \eta_i$

   (b) What is the intuitive interpretation of δ?

   (c) Show that $\frac{\Delta y_i}{\Delta d_i} = \frac{\delta - 1}{\delta^{d_i}} y_i$. What is the intuitive interpretation of $\frac{\Delta y}{\Delta d_i}$?

   (d) Show that $\frac{\partial y_i}{\partial z_i} = \gamma \frac{y_i}{z_i}$. What is the intuitive interpretation of $\frac{\partial y_i}{\partial z_i}$ when z_i is the size of the home in square feet?

   (e) Estimate the log-transformed equation via ordinary least squares on the transformed parameters using any algorithm you would like. Save the coefficient estimates and the average marginal effects estimates of z_i and d_i $\left(\frac{dy_i}{dz_i} \text{ and } \frac{\Delta y}{\Delta d_i}\right)$. Bootstrap the 95% confidence intervals of the coefficient estimates and the marginal effects estimates using 1000 sampling replications (note that each bootstrap replication should perform both the regression and the second stage calculation of the marginal effect). Display the results in a table with three columns (one for the variable name, one for the coefficient estimate, and one for the marginal effect estimate). Show the 95% confidence intervals for each estimate under each number.

   (f) Graph the **average marginal effects** of outdoor temperature and square feet of the home with bands for their bootstrapped confidence intervals so that they are easy to interpret and compare.