Homework 3 Submission

David Wilson
Economics 7103
February 5, 2024

Stata and 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, and η_i is unobserved error, and $\{\alpha, \delta, \gamma\}$ are the parameters to estimate.

   (a) Show that $\ln(y_i) = \alpha + \ln(\delta)d_i + \gamma\ln(z_i) + \eta_i$
   Take the natural log of both sides.

   $\ln(y_i) = \ln(e^{\alpha}\delta^{d_i}z_i^{\gamma}e^{\eta_i})$
   $\ln(y_i) = \ln(e^{\alpha}) + \ln(\delta^{d_i}) + \ln(z_i^{\gamma}) + \ln(e^{\eta_i})$
   $\ln(y_i) = \alpha + d_i\ln(\delta) + \gamma\ln(z_i) + \eta_i$

   (b) What is the intuitive interpretation of δ?

   We could say that δ is the marginal treatment effect since d_i is the treatment dummy variable. It represents the change in electricity usage if house i received the retrofit program. If the program is expected to lower energy consumption, then we should expect the sign of δ to be negative.

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

   The intuitive interpretation of $\frac{\Delta y_i}{\Delta d_i}$ that it represents how a change in energy consumption is related to receiving the retrofit program. Since the retrofit program is aimed at reducing monthly energy consumption, it can be hypothesized that Δd_i should be negative.

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

   The intuitive interpretation of $\frac{\delta y_i}{\delta z_i}$ when z_i is the size of the home is that any increase in the size of the home will increase the size of monthly energy consumption.

   (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 $\left(\frac{\delta y_i}{\delta z_i} \text{ and } \frac{\Delta y_i}{\Delta d_i}\right)$. 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.

Table 1. Coefficient and Marginal Effect Estimates from Stata

	Parameter Estimates	Marginal Effect Estimates
=1 if home received retrofit	0.904	-121.566
	[0.893,0.915]	[-141.426,-101.706]
Size of home in ft^2	0.894	0.628
	[0.880,0.909]	[0.612,0.643]
Average outdoor temperature F°	0.281	2.425
	[0.030,0.533]	[-2.482,7.332]
Observations	1,000	1,000

Note: Values in the first column are the coefficient estimates for each variable while the second column displays the average marginal estimates for each coefficient. Both columns have the 95% confidence interval for each variable in parentheses.

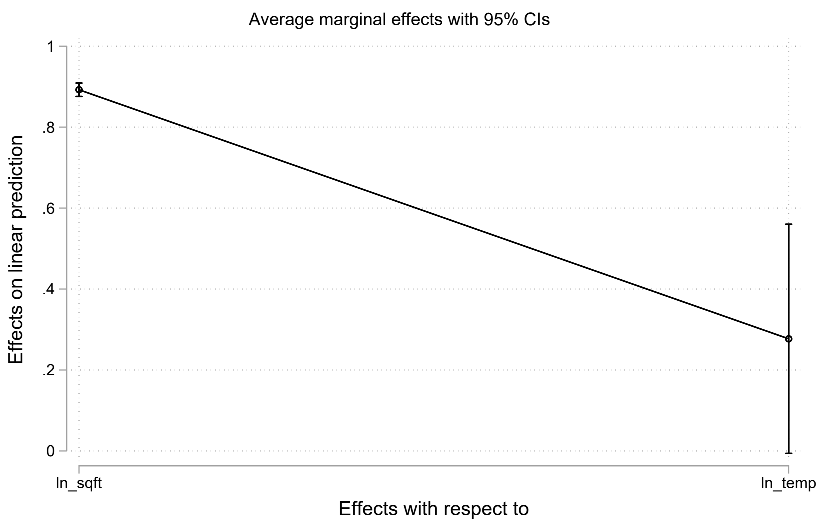

Fig. 1. This graph shows the average marginal effects of outdoor temperature and the square footage of the home.