

Homework 3 Solutions

Kelly Lifchez

January 27, 2025

1 Python

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

$$y_i = e^\alpha \delta^{d_i} z_i^\gamma e^{\eta_i} \quad (1)$$

$$\ln(y_i) = \ln(e^\alpha \delta^{d_i} z_i^\gamma e^{\eta_i}) \quad (2)$$

$$\ln(y_i) = \ln(e^\alpha) + \ln(\delta^{d_i}) + \ln(z_i^\gamma) + \ln(e^{\eta_i}) \quad (3)$$

$$\ln(y_i) = \alpha \ln(e) + \ln(\delta)d_i + \gamma \ln(z_i) + \eta_i \ln(e) \quad (4)$$

$$\ln(y_i) = \alpha + \ln(\delta)d_i + \gamma \ln(z_i) + \eta_i. \quad (5)$$

(b) What is the intuitive interpretation of δ ?

δ can be thought of as the direct effect of treatment, or the treatment effect on the treated group. In other words, δ represents a scaling or multiplier effect of the treatment on y_i . This is because when $d_i = 0$, δ drops out of the regression equation, and when $d_i = 1$, δ is included. Specifically:

When $d_i = 0$, $y_i = e^\alpha z_i^\gamma e^{\eta_i}$ or $\ln(y_i) = \alpha + \gamma \ln(z_i) + \eta_i$.

And when $d_i = 1$, $y_i = e^\alpha \delta z_i^\gamma e^{\eta_i}$ or $\ln(y_i) = \alpha + \ln(\delta)(1) + \gamma \ln(z_i) + \eta_i$.

That means that in this case, δ is the direct multiplicative change in kWh due to retrofitting, and $\ln(\delta)$ represents the percent change in kWh of electricity used in homes that receive the retrofitting treatment.

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

$$\Delta y_i = y_i(d_i = 1) - y_i(d_i = 0) \quad (6)$$

We know the potential outcomes model:

$$y_i = \begin{cases} e^\alpha z_i^\gamma e^{\eta_i} & \text{if } d_i = 0 \\ e^\alpha \delta z_i^\gamma e^{\eta_i} & \text{if } d_i = 1 \end{cases}$$

Therefore we know that

$$\Delta y_i = (e^\alpha \delta z_i^\gamma e^{\eta_i}) - (e^\alpha z_i^\gamma e^{\eta_i}) \quad (7)$$

Pull out $e^\alpha z_i^\gamma e^{\eta_i}$ and get:

$$\Delta y_i = (\delta - 1)(e^\alpha z_i^\gamma e^{\eta_i}) \quad (8)$$

Next, find the change in d_i :

$$\Delta d_i = 1 - 0 \quad (9)$$

So we can then calculate the change in y_i over the change in d_i :

$$\frac{\Delta y_i}{\Delta d_i} = \frac{(\delta - 1)(e^\alpha z_i^\gamma e^{\eta_i})}{1 - 0} = (\delta - 1)(e^\alpha z_i^\gamma e^{\eta_i}) \quad (10)$$

(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?

Star with:

$$\ln(y_i) = \alpha + \ln(\delta)(1) + \gamma \ln(z_i) + \eta_i \quad (11)$$

Exponentiate both sides:

$$y_i = \exp(\alpha + \ln(\delta)(1) + \gamma \ln(z_i) + \eta_i) \quad (12)$$

Use the chain rule to take the partial derivative of y_i with respect to z_i :

$$\frac{\partial y_i}{\partial z_i} = \left(\frac{d}{dz}\right)(\exp(\alpha + \ln(\delta)(1) + \gamma \ln(z_i) + \eta_i)) \quad (13)$$

$$= \left(\frac{\alpha}{z_i}\right)(y_i) \quad (14)$$

$$= \alpha \frac{y_i}{z_i} \quad (15)$$

In other words, a 100% change in z creates a $100 \times \alpha$ percent change in y . This can also be thought of as a 1% change in home square footage leading to an α percent change in kWh of electricity use.

(e) Table of coefficient estimates and marginal effects

Table 1 shows OLS coefficient estimates and the average marginal effects of each of the variables in the regression model on kWh of electricity. Coefficient estimates provide insight into the direct relationship between an explanatory variable and the outcome variable, while the average marginal effects describe the overall impact of an explanatory variable on an outcome in the context of the entire model.

Confidence intervals represent the range in which an estimated parameter will fall 95% of the time if repeatedly sampled. Confidence intervals in Table 1 are constructed using bootstrapping with 1,000 replications.

Table 1: Coefficient estimates and average marginal effects

	Coefficients	Marginal effects (dy/dx)
Retrofit	-0.10 (-0.11, -0.09)	-113.98 (-0.79, -0.61)
Log(Sqft)	0.89 (0.88, 0.91)	0.63 (0.00, 0.00)
Log(Temp)	0.28 (0.04, 0.53)	4.00 (0.00, 0.05)
Constant	-0.77 (-1.88, 0.33)	
Observations	1000	1000

(f) Graph of average marginal effects

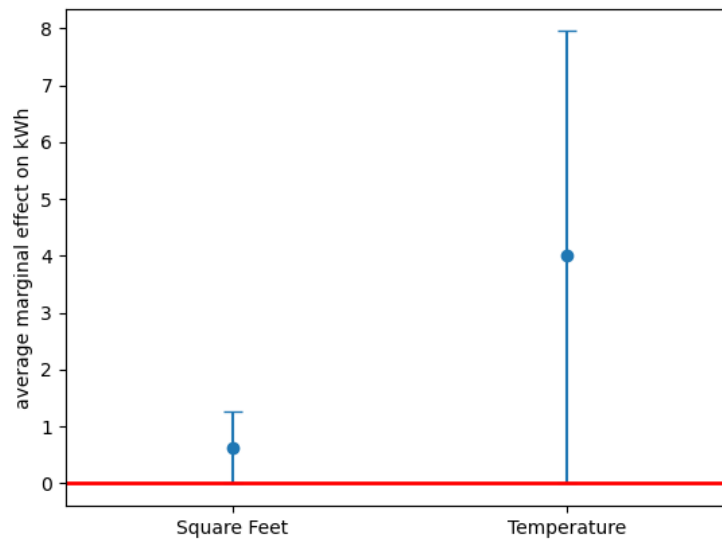


Figure 1: Average marginal effects of home square footage and outdoor temperature

Note: The figure depicts the average marginal effects of square footage in a home and outdoor temperature on kWh. Bar caps represent 95% confidence intervals bootstrapped with 1,000 replications.