Chapter 6, Log, Quadratic Term, Interaction Term and Adjusted R-squared (Appendix A2-A4)

(I) Consider the simple regression model: $y = \beta_0 + \beta_1 x + u$. Calculus indicates that this model assumes constant marginal effect of x on y: $\frac{dy}{dx} = \beta_1$. However, there are many examples in economics in which the marginal effect is not constant (e.g., diminishing marginal utility, increasing marginal cost). To fix this issue, consider a level-log model:

$$y = \beta_0 + \beta_1 \log(x) + u \qquad (level - log)$$

We can show the marginal effect is nonconstant: $\frac{dy}{dx} = \frac{\beta_1}{x}$. As x rises, the marginal effect goes to zero. That means this level-log model can be used to describe the diminishing marginal utility (let y be utility, and x be the quantity of goods). Moreover we can show

$$\beta_1 = \frac{dy}{dlog(x)} \approx \frac{\Delta y}{\Delta \log(x)} \approx \frac{\Delta y}{\frac{\Delta x}{x}} \rightarrow \Delta y \approx \beta_1(\Delta x/x) \qquad (2)$$

We use the fact that change in log approximates percentage change:

$$\Delta log(x) \approx \frac{\Delta x}{x} \qquad (3)$$

If x changes by one-percent then the change in y is ____________ In short we need to divide β_1 by 100 when interpreting β_1 in the level-log model.

Exercise 1: Consider the log-level model $\log(y) = \beta_0 + \beta_1 x + u$. The marginal effect is _______________. How to interpret β_1? ___________________________________.

We also have log-log model $\log(y) = \beta_0 + \beta_1 \log(x) + u$. In this case, β_1 is the ratio of percentage change, or β_1 measures elasticity.

Table 2.3 on page 44 (5th edition) summarizes how to interpret β_1 in various models.

(II) There are two limitations of using log variable. First we cannot take log if the variable takes negative value or zero. Second, the model implies that there is no bound or no turning point. The model with quadratic term x^2 does not have these limitations

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + u \qquad (regression\ with\ quadratic\ term)$$

The key to understand the regression with quadratic term is to draw a graph for the quadratic function. We see a cup facing down if $\beta_2 < 0$, and there is a (minimum maximum). We see a facing-up cup when $\beta_2 > 0$, and there is a (minimum maximum).

Exercise 2: the marginal effect is (constant nonconstant)

Exercise 3: suppose $\beta_2 < 0$. The value of x that returns maximum value of y is__________________

In practice, the quadratic regression can be used to find the optimal class size, optimal number of bathroom, etc.

(III) Consider the regression that includes the product of two regressors(xw). This product is called interaction term.

$$y = \beta_0 + \beta_1 x + \beta_2 (xw) + u \qquad (regression\ with\ interaction\ term)$$

In this case the marginal effect of x on y depends on w:

$$\frac{dy}{dx} = \beta_1 + \beta_2 w$$

In short, the regression with interaction term allows interaction between regressors.

The quadratic regression and interaction-term regression have the drawback that it becomes hard to interpret β_1. Simply put, β_1 alone does not measure the marginal effect, or β_1 measures the marginal effect only when $x = 0$ (quadratic model) or $w = 0$ (interaction-term model). Nevertheless, Dr. Wooldridge suggests a smart solution. Read example 6.3 on the textbook.

(IV) In practice we can run many regressions. Now the question becomes which regression is the best. This problem is called model selection. For example, consider running three regressions

$$y = \beta_0 + \beta_1 x + u \qquad (A)$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + u \qquad (B)$$

$$y = \beta_0 + \beta_1 \log(x) + u \qquad (C)$$

Model (A) is a special case of (B) by imposing the null hypothesis __________________________; In this case we call model A is nested in Model B, and the F test or t test can be used for selecting the right model.

Exercise 4: how to show whether an interaction term is necessary in a regression?

Models (B) and (C) are non-nested in the sense that no one is special case of the other. In this case, we want to pick the model that has greater adjusted R-squared.

$$adjusted\ R^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$$

Sometimes, our economic knowledge can help. For example, if we know y is bounded, or there is a turning point, then model C becomes inappropriate because the log function is monotonic.

Finally, adjusted R-squared can be used to compare models only when the models have same dependent variables (so TSS is comparable). In other words, we cannot use R-square to compare a level-log model to a log-level model because their dependent variables are different. We can still use our economic knowledge though.

(V) The residual can be used for many purposes. For example, we can find outlier quickly by sorting the residual and paying attention to observations with the minimum and maximum residuals. As a second example, residual can help economic decision. A house is overpriced if the true price is greater than the fitted price, or if the residual is positive.

Exercise 5: how to recommend a stock to your parent?