ECN 102: Analysis of Economics Data

Chapter 15: Marginal Effects

Remy Beauregard

Marginal Effects: Overview

Previously, interpreting the partial effect of a regressor has been as simple as looking at its associated coefficient. However, we may want to estimate **nonlinear models** with quadratic forms, log transformations, or interaction terms. In these cases, computing the **marginal effect** of x on y will require a bit more math.

As before, we denote the marginal effect as:

$$ME_x = \frac{\Delta \hat{y}}{\Delta x}$$

Marginal Effects: Partial Derivatives

To compute our marginal effect, we generally take the partial derivative of our predicted values with respect to our regressor of interest.

If we estimated a quadratic model such as $y = \beta_1 + \beta_2 x + \beta_3 x^2 + u$, we would compute:

$$\frac{d}{dx}\hat{y} = b_2 + 2xb_3$$

As we see, this marginal effect includes an x in it. We need to know which value of x to plug in to compute the proper marginal effect.

Marginal Effects: MEM, MER, and AME

There are three common types of marginal effects we compute:

1. Marginal effect at the mean (**MEM**) where we plug in $x_j = \bar{x}_j$ for each regressor present in our derivative
2. Marginal effect at a representative value (**MER**) where we plug in $x_j = x_j^*$ for each regressor
3. Average marginal effect (**AME**) where we take the average of marginal effects at each value of x for each regressor

We can write the AME as $\frac{1}{n}\sum_{i=1}^{n} \text{MER}(x_j = x_{ji})$

Marginal Effects: Interacted Regressors

We may also have a regression with **interacted regressors**: terms that depend on more than one regressor. In this case, our derivative may include more than one regressor for which we would plug in corresponding values.

If a given *regressor* is interacted with others in our regression, we would need a joint F-test to determine whether or not that *variable* contributes explanatory power overall to our regression or not. We could not use a t-test as we are testing multiple linear restrictions (for multiple coefficients) at once.

Marginal Effects: Interaction Example and Joint F-Test

Suppose we estimate the regression:

$$y = \beta_1 + \underset{\uparrow}{\beta_2} x + \beta_3 z + \underset{\uparrow}{\beta_4}(x \times z) + u$$

$$\Rightarrow \frac{\partial}{\partial x}\hat{y} = b_2 + b_4 z$$

Marginal Effects: Interaction Example and Joint F-Test

Suppose we estimate the regression:

$$y = \beta_1 + \underset{\uparrow}{\beta_2} x + \beta_3 z + \underset{\uparrow}{\beta_4}(x \times z) + u$$

$$\Rightarrow \frac{\partial}{\partial x}\hat{y} = b_2 + b_4 z$$

If we wanted to test joint significance of all *regressors* invoking the *variable* x with an F-test $(q = 2)$, our (joint) hypotheses would be:

$$H_0 : \beta_2 = 0 \cap \beta_4 = 0 \leftrightarrow \text{var. } x \text{ has no impact on } y$$

$$H_A : \beta_2 \neq 0 \cup \beta_4 \neq 0 \leftrightarrow \text{var. } x \text{ has some impact on } y$$

Marginal Effects vs. Partial Effects

Above we have (jointly) tested significance of all *regressors* invoking the *variable* x on y using an F-test. How is this different from testing just the coefficient on our *regressor* x?

Marginal Effects vs. Partial Effects

Above we have (jointly) tested significance of all *regressors* invoking the *variable* x on y using an F-test. How is this different from testing just the coefficient on our *regressor* x?

A t-test on β_2 tests only the **partial effect** of x on y, the component from the standalone x term, equal to the marginal effect of x when $z = 0$. This is only *one part* of x's overall effect on y.

In this case, we would be (t- or F-)testing $(q = 1)$:

$$\left.\begin{array}{r} H_0 : \beta_2 = 0 \\ H_A : \beta_2 \neq 0 \end{array}\right\} \begin{array}{l} \text{There is or is not a significant} \\ \text{partial effect of } x \text{ by itself on } y \end{array}$$

Marginal Effects vs. Partial Effects

Above we have (jointly) tested significance of all *regressors* invoking the *variable* x on y using an F-test. How is this different from testing just the coefficient on our *regressor* x?

A t-test on β_2 tests only the **partial effect** of x on y, the component from the standalone x term, equal to the marginal effect of x when $z = 0$. This is only *one part* of x's overall effect on y.

In this case, we would be (t- or F-)testing $(q = 1)$:

$$\left.\begin{aligned} H_0 &: \beta_2 = 0 \\ H_A &: \beta_2 \neq 0 \end{aligned}\right\} \begin{array}{l}\text{There is or is not a significant} \\ \text{partial effect of } x \text{ by itself on } y\end{array}$$

Even if we fail to reject $\beta_2 = 0$, x can still affect y through the interaction ($\beta_4 z$). Testing whether x matters *at all* requires the **joint F-test** on β_2 and β_4 $(q = 2)$, not a t-test $(q = 1)$.

Marginal Effects vs. Joint Significance

More generally, we could imagine cases where tests of joint significance deliver different conclusions than individual tests of association:

Marginal Effects vs. Joint Significance

More generally, we could imagine cases where tests of joint significance deliver different conclusions than individual tests of association:

Failing to reject each individual t-test does not mean we fail to reject the joint F-test. When regressors are correlated, as x and $x \times z$ typically are, each coefficient is estimated imprecisely (large standard errors), so each t-test on its own may fail to reject.

Marginal Effects vs. Joint Significance

More generally, we could imagine cases where tests of joint significance deliver different conclusions than individual tests of association:

Failing to reject each individual t-test does not mean we fail to reject the joint F-test. When regressors are correlated, as x and $x \times z$ typically are, each coefficient is estimated imprecisely (large standard errors), so each t-test on its own may fail to reject.

Yet those regressors can *together* explain enough variation in y that the joint F-test rejects. We might find no significant *partial* effect from either t-test alone, while the F-test shows that x **does** affect y overall.

Marginal Effects vs. Joint Significance

More generally, we could imagine cases where tests of joint significance deliver different conclusions than individual tests of association:

Failing to reject each individual t-test does not mean we fail to reject the joint F-test. When regressors are correlated, as x and $x \times z$ typically are, each coefficient is estimated imprecisely (large standard errors), so each t-test on its own may fail to reject.

Yet those regressors can *together* explain enough variation in y that the joint F-test rejects. We might find no significant *partial* effect from either t-test alone, while the F-test shows that x **does** affect y overall.

This is the hallmark of **multicollinearity**: uninformative individual t-tests alongside a significant joint F-test.

Marginal Effects: Log-Linear Models

For log-linear models, we estimate $\frac{\Delta \ln \hat{y}}{\Delta x} = b_2 \approx \frac{\Delta \hat{y}/\hat{y}}{\Delta x}$. To isolate our desired marginal effect of $\frac{\Delta \hat{y}}{\Delta x}$, we need to rearrange our terms:

$$\frac{\Delta \ln \hat{y}}{\Delta x} = b_2$$

$$\frac{\Delta \ln \hat{y}}{\Delta x} \approx \frac{\Delta \hat{y}/\hat{y}}{\Delta x} = b_2$$

$$\hat{y}\left(\frac{\Delta \hat{y}/\hat{y}}{\Delta x}\right) = (b_2)\hat{y}$$

$$\underbrace{\frac{\Delta \hat{y}}{\Delta x} = b_2 \hat{y}}_{\text{Marginal effect}}$$

Marginal Effects: Linear-Log and Log-Log Models

Similarly for other models, we will estimate our marginal effects as:

$$\frac{\Delta \hat{y}}{\Delta x} = \frac{b_2}{x} \text{ (linear-log)}$$

$$\frac{\Delta \hat{y}}{\Delta x} = b_2 \times \frac{\hat{y}}{x} \text{ (log-log)}$$

$$\frac{\Delta \hat{y}}{\Delta x} = b_2 \hat{y} \text{ (log-linear)}$$

$$\frac{\Delta \hat{y}}{\Delta x} = b_2 \text{ (linear)}$$

When to Use Each Model

When would we use each of these models?

- We use **log transformations** when we are concerned with a *proportional* change in one or more variables rather than a level change or when we have a right skew
- We use a **quadratic model** if we think the effects of a regressor are nonlinear or change with values of x. A common example is *age* and *income*: average earnings initially increase with *age* but then decrease as we start to retire.
- We use an **interacted model** if we think there is relevant covariance between regressors that we want to model. For example, we may want to interact *age* with *education* to capture how the returns to aging vary with level of education.

End of lecture material

Knowledge Check 15

Compute the marginal effect at the mean (MEM) of a 1-unit change in x for the following models; $\bar{x} = 4$, $\bar{y} = 10$, $\bar{z} = 7$:

a) $\ln y = \beta_1 + \beta_2 x + u$

b) $y = \beta_1 + \beta_2 x + \beta_3 x^2 + u$

c) $y = \beta_1 + \beta_2 x + \beta_3 z + \beta_4 (x \times z) + u$

d) $\ln y = \beta_1 + \beta_2 \ln x + u$