Functional form

Econometrics (35B206), Lecture 7

Tunga Kantarcı, TiSEM, Tilburg University, Spring 2019

Functional form, quadratic in the variable

A regression model that is quadratic in a regressor is often of interest because it provides economic interpretations. E.g., it provides the decreasing marginal utility interpretation.

Functional form, quadratic in the variable

Consider the regression model

$$wage = \beta_0 + \beta_1 age + \beta_2 age^2 + u$$

which is quadratic in age.

Functional form, quadratic in the variable

The dependent changes approximately by the amount

$$\frac{\partial wage}{\partial age} = \beta_1 + 2\beta_2 age$$

if age changes by a small amount. It changes exactly by

$$\begin{aligned}\Delta wage = &(\beta_0 + \beta_1(age + \Delta age) + \beta_2(age + \Delta age)^2) \\ &- (\beta_0 + \beta_1 age + \beta_2 age^2)\end{aligned}$$

by some given change in age. The approximate change will be good for small changes but bad for big changes in *age*.

Functional form, quadratic in the variable

In reference to calculus, the former change is the tangent slope, and the latter is the secant slope. The tangent slope refers to a marginal change, and the secant slope refers to a discrete change in the nonlinear function. We take the tangent slope as an approximation to the secant slope. That is,

$$\frac{\delta y}{\delta x} \approx \frac{\Delta y}{\Delta x}.$$

Functional form, quadratic in the variable

In either case, the expression states that the partial effect of *age*, or the slope of the relationship between *wage* and *age*, depends on the value of *age*.

Functional form, nonlinearity in the variables

The logarithmic function also provides economic interpretations. It also has econometric implications but we do not cover them here. For example, it has implications if there is heteroskedasticity.

Functional form, logarithmic in the variable

Consider the following logarithmic change:

$$\ln(41) - \ln(40) \approx 0.024.$$

The proportionate change, or relative change, we frequently calculate is

$$\frac{41-40}{40} = 0.025.$$

The two quantities are very close. This shows that, for small changes, the logarithmic change closely approximates the proportionate change.

Functional form, logarithmic in the variable

Consider the following logarithmic change:

$$\ln(60) - \ln(40) \approx 0.405.$$

The proportionate change is

$$\frac{60 - 40}{40} = 0.500.$$

The two quantities are not really close. This shows that, for large changes, the approximation is not accurate.

Functional form, logarithmic in the variable

We can make use of this result in regression analysis. Assume a log-linear model taking the form

$$\ln(y) = \beta_1 x_1 + u.$$

We are interested in the change in $\ln(y)$ when we change x_1 by some unit.

Functional form, logarithmic in the variable

$$\ln(y) = \beta_1 x_1 + u.$$

For small changes in x_1, the change in $\ln(y)$ closely approximates the proportionate change in y. In this case we can consider the derivate

$$\frac{\partial \ln(y)}{\partial x_1} = \beta_1.$$

This change refers to the tangent slope.

Functional form, logarithmic in the variable

$$\ln(y) = \beta_1 x_1 + u.$$

For large changes in x_1, the approximation is worse. In this case the exact proportionate change can be calculated as

$$\Delta \ln(y) = \ln(y_1) - \ln(y_0) = \beta_1 \Delta x_1.$$

Taking the terms to the exponential gives

$$\begin{aligned}
e^{(\ln(y_1) - \ln(y_0))} &= e^{\beta_1 \Delta x_1} \\
e^{\ln \frac{y_1}{y_0}} &= e^{\beta_1 \Delta x_1} \\
\frac{y_1}{y_0} &= e^{\beta_1 \Delta x_1} \\
\frac{y_1}{y_0} - 1 &= e^{\beta_1 \Delta x_1} - 1 \\
\frac{y_1 - y_0}{y_0} &= e^{\beta_1 \Delta x_1} - 1
\end{aligned}$$

This change refers to the secant slope.

Functional form, logarithmic in the variable

The main implications of the logarithmic transformation for applied work are the following. In the log-linear model

$$\ln(y) = \beta_1 x_1 + u,$$

we have

$$\frac{\Delta \ln(y)}{\Delta x} = \beta_1$$

which gives a proportionate change interpretation. If we multiply by 100, we have

$$100 * \frac{\Delta \ln(y)}{\Delta x} = 100 * \beta_1$$

which gives a percentage change interpretation. That is,

$$\frac{\%\Delta y}{\Delta x} = 100 * \beta_1.$$

Functional form, logarithmic in the variable

In the linear-log model

$$y = \beta_1 \ln(x_1) + u$$

we have

$$1/100 * \Delta y / \Delta \ln(x) = 1/100 * \beta_1$$

and

$$\Delta y / \% \Delta x = \beta_1 / 100$$

Functional form, logarithmic in the variable

In the log-log model

$$\ln(y) = \beta_1 \ln((x_1) + u,$$

we have

$$100/100 * \Delta \ln(y) / \Delta \ln(x) = \beta_1$$

and

$$\%\Delta y / \%\Delta x = \beta_1.$$

This is constant elasticity. It is often used in applied work. In all cases, the proportionate change is approximated by the logarithmic change.

Functional form, interaction effects

Suppose that the model of interest is

$$\mathrm{E}(\textit{wage} \mid \textit{educ}, \textit{female}) = \beta_0 + \beta_1 \textit{educ} + \beta_2 \textit{female} + \beta_3 \textit{educ} * \textit{female}.$$

educ is a continuous variable.

female is a dummy variable.

$\textit{educ} * \textit{female}$ is an interaction variable which indicates *educ* with respect to the groups of *female*.

Functional form, interaction effects

We can see the motivation behind an interaction variable if we change either of the two interacting variables.

Functional form, interaction effects

First, consider a unit change in *educ*:

$$\frac{\Delta \mathrm{E}(\mathit{wage} \mid \mathit{educ}, \mathit{female})}{\Delta \mathit{educ}} = \beta_1 + \beta_3 \mathit{female}$$

This equation indicates that the effect of *educ* on *wage* depends on the group of *female*.

Functional form, interaction effects

Second, consider a change in *female*.

If $female = 1$, we obtain the following model:

$$E(wage \mid educ, female = 1) = \beta_0 + \beta_2 + (\beta_1 + \beta_3)educ.$$

If $female = 0$, we obtain another model:

$$E(wage \mid educ, female = 0) = \beta_0 + \beta_1 educ.$$

The two models differ in the constant β_2 and the slope β_3. We get the same interpretation: The effect of *educ* on *wage* depends on the *female*.

Functional form, interaction effects

$educ * female$ is also called the *slope dummy variable* because it allows a change in the slope.