II.8 More about functional forms

- Models with quadratic terms
- Interaction terms
- Dummy variables with interaction terms

Models with quadratic terms

Interest in modeling increasing or decreasing marginal effects of certain variables - capture such effects by including quadratic terms:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u. \tag{71}$$

- OLS estimation of β_0, β_1, and β_2 proceeds as discussed above, based on the predictors $X_1 = X$ and $X_2 = X^2$.

- Although the relationship between X_1 and X_2 is deterministic (note that $X_2 = X_1^2$), the predictors X_1 and X_2 are not linearly dependent. Hence, OLS estimation is feasible.

- Note that the relationship between X and Y is non-linear.

Models with quadratic terms

The parabola corresponding to the quadratic function (71) opens up, if $\beta_2 > 0$, and opens down, if $\beta_2 < 0$

Often, only part for the parabola is used to describe a monotonic behavior over a certain range of X

Model (71) reduces to a model linear in X, if $\beta_2 = 0 \Rightarrow$ test the null hypothesis $H_0 : \beta_2 = 0$ to test for the presence of non-linear effects.

If $\beta_2 \neq 0$, then non-linearity is present in model (71). In this case β_1 does not measure the expected change in Y with respect to X, since $X_2 = X^2$ cannot be held constant, while $X_1 = X$ changes. Changing X changes both predictors X_1 and X_2.

Models with quadratic terms

The instantaneous change is equal to the first derivative of $\mathrm{E}(Y)$ with respect to X:

$$\frac{\partial \mathrm{E}(Y)}{\partial X} = \beta_1 + 2\beta_2 X.$$

If X is changed by ΔX, then the expected change of Y is approximately given by:

$$\Delta \mathrm{E}(Y) \approx \frac{\partial \mathrm{E}(Y)}{\partial X} \cdot \Delta X = (\beta_1 + 2\beta_2 X)\Delta X. \tag{72}$$

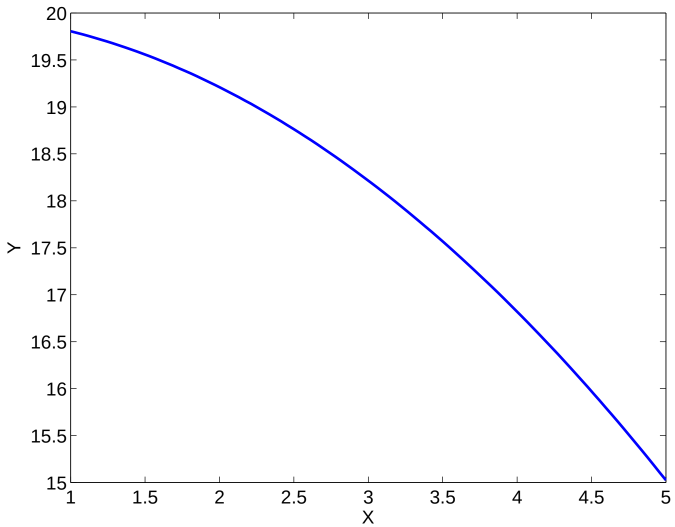
Models with quadratic terms

- The expected change of Y is equal to β_1 only, if $X = 0$.
- For $X \neq 0$, the expected change of Y depends not only on β_1, but also on β_2 and X.
- The expected change of Y switches sign at the turning point $X = -\beta_1/(2\beta_2)$.
- The model describes a monotonic behavior, if only value of X are considered, which lie on one side of the turning point.

Models with quadratic terms

Example: $Y = 20 + 0.005 \cdot X - 0.2 \cdot X^2, \quad 1 \leq X \leq 5$

Parabola opens down, because $\beta_2 = -0.2 < 0$; turning point: $0.005 - 0.4X_0 = 0 \Rightarrow X_0 = 0.0125$; range of X restricted to the right hand side $\Rightarrow$ monotonically decreasing function



Models with quadratic terms

Suppose that β_1 is positive, while β_2 is negative. Then according to the first term in (72), increasing X will **increase** $\mathrm{E}(Y)$, however, this positive effect becomes smaller with increasing X. It remains positive as long X is smaller than the turning point:

$$X < \frac{\beta_1}{2|\beta_2|},$$

However, if X is larger than the turning point, then there is a **negative** effect of increasing X, which gets larger with increasing X.

Models with quadratic terms

Suppose that β_1 is negative, while β_2 is positive. Then according to the first term in (72), increasing X will **decrease** $\mathrm{E}(Y)$, however, this negative effect becomes smaller with increasing X. It remains positive as long X is smaller than the turning point:

$$X < \frac{|\beta_1|}{2\beta_2}.$$

then the expected change is equal to 0. If X is larger than the turning point, i.e.

However, if X is larger than the turning point, then there is a **positive** effect of increasing X, which gets larger with increasing X.

Models with quadratic terms

If β_1 and β_2 have the same sign, then the turning point is negative. The interpretation is easy, if X allows only positive values:

- A negative effect (i.e. both coefficients have a negative sign) becomes more pronounced with increasing X.

- A positive effect (i.e. both coefficients have a positive sign) becomes more pronounced with increasing X.

Case Study Chicken

Estimate the model:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 P_1 + \beta_3 P_1^2 + \beta_4 P_2 + \beta_5 P_2^2 + u,$$

with X_1 income, P_1 price of chicken and P_2 the price of pork. This model outperforms a model without quadratic terms according to AIC and the Schwarz criterion.

β_2 is negative, but the negative effect decreases as the price increases, since β_3 is positive. The turning point is equal to

$$\frac{|\beta_2|}{2\beta_3} = \frac{1.69}{2 \cdot 0.014^2} = 4311.$$

Case Study Chicken

This value is much larger than the largest observed price, hence the price effect is negative over the whole range of observations.

β_4 is positive, but the positive effect decreases as the price of pork increases, since β_5 is negative. The turning point is equal to

$$\frac{\beta_4}{|2\beta_5|} = \frac{0.542}{2 \cdot 0.0024^2} = 47048.6$$

This value is much larger than the largest observed price, hence the price effect is positive over the whole range of observations.

Interaction terms

- In some cases it makes sense to make the effect of a variable X_1 on Y dependent on another regressor X_2.

- One way to capture such effects is including interactions terms:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 \cdot X_2 + u. \tag{73}$$

- OLS estimation of $\beta_0, \ldots, \beta_3$ proceeds as discussed above, based on the predictors X_1, X_2, and $X_3 = X_1 \cdot X_2$.

- Note that the relationship between X_1 and Y is non-linear as is the relationship between X_2 and Y.

Interaction terms

In this framework, β_1 does not measure the expected change in Y with respect to X_1, since $X_3 = X_1 X_2$ cannot be held constant, if X_1 changes. Hence, if X_1 is changed by ΔX_1, then the expected change of Y is approximately given by:

$$\Delta \mathrm{E}(Y) \approx \Delta X_1 \frac{\partial Y}{\partial X_1}$$

where the first derivative of $\mathrm{E}(Y)$ with respect to X_1 is given by:

$$\frac{\partial Y}{\partial X_1} = \beta_1 + \beta_3 X_2.$$

Interaction terms

Therefore, β_1 is the effect of X_1 on Y only for $X_2 = 0$, which is not necessarily a reasonable value of X_2. The **average** effect of X_1 on Y can be evaluated at the sample mean of X_2:

$$\delta_1 = \beta_1 + \beta_3 \overline{X_2}.$$

An alternative parameterization of the model is

$$Y = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 (X_1 - \overline{X_1})(X_2 - \overline{X_2}) + u,$$

where δ_1 is directly the partial effect of X_1 on Y at the mean of X_2.

Case Study Chicken

Estimate a model with income X_1, price of chicken P_1 and price of pork P_2:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 P_1 + \beta_3 P_2 + \beta_4 X_1 P_2 + u.$$

This model outperforms a model without the interaction term according to AIC and the Schwarz criterion.

β_3 is positive, but the positive effect of increasing the price of pork decreases as the income X_1 increases, since β_4 is negative:

$$\frac{\partial Y}{\partial P_2} = \beta_3 + \beta_4 X_1.$$

Case Study Chicken

The average income is equal to $\overline{X_1} = 1035.065$, hence the average effect of the price of pork is equal to:

$$\delta_1 = \beta_3 + \beta_4 \overline{X_1} = 0.16 + 1035.065 \cdot (-8.62E - 05) = 0.0708.$$

This value is considerably smaller than the effect obtained from the model without interaction term.

The average effect of the price of pork is obtained immediately from OLS estimation, if following model is fitted to the data:

$$Y = \gamma_0 + \gamma_1 X_1 + \beta_2 P_1 + \gamma_3 P_2 + \gamma_4 (X_1 - \overline{X_1})(P_2 - \overline{P_2}) + u,$$

and is equal to γ_3.

Interaction with Dummy Variables

Consider interacting a dummy variable, D, with a continuous variable, X:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 X \cdot D + u. \tag{74}$$

If $D = 0$, then:

$$Y = \beta_0 + \beta_2 X + u.$$

If $D = 1$, then:

$$Y = (\beta_0 + \beta_1) + (\beta_2 + \beta_3) X + u.$$

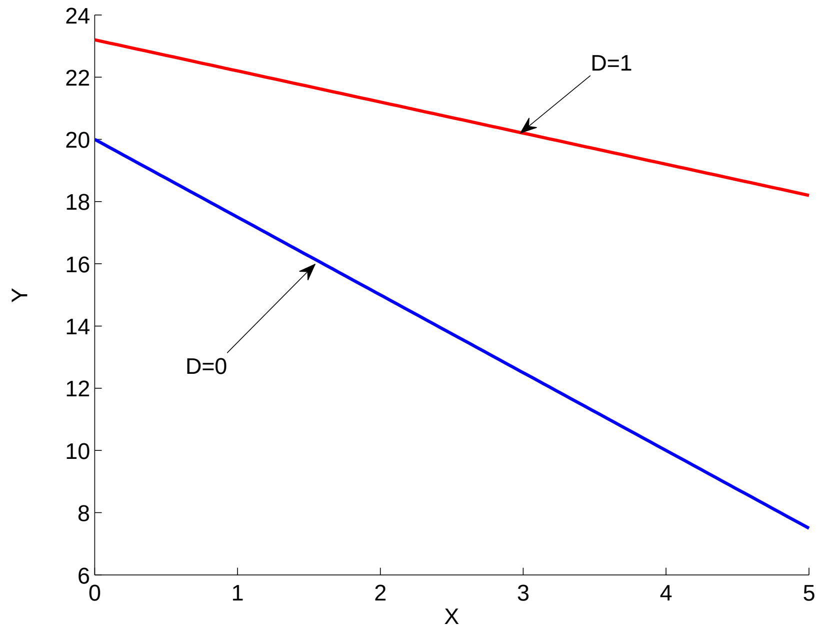
Interaction with Dummy Variables

Interpretation:

- The observed units are split into 2 groups according to D (e.g. into men and women).
- Then β_1 quantifies the average difference of the dependent variable Y between the two groups, while holding X fixed.
- The coefficient β_3 models a change in the slope of the regression model.
- The joint null hypothesis $\beta_1 = 0$, $\beta_3 = 0$ corresponds to the assumption that the regression model is the same for both groups.

Interaction with Dummy Variables

Example: $Y = 20 + 3.2 \cdot D - 2.5 \cdot X + 1.5 \cdot D \cdot X$



Case Study Marketing

Estimate a model with a specific brand $(D = 1)$ and price P:

$$Y = \beta_0 + \beta_1 D + \beta_2 P + \beta_3 PD + u,$$

- There is a very significant price effect for the specific brand.
- Increasing the price for an ordinary brand by one unit leads to an expected decrease in the rating by β_2, i.e. nearly 2 points.
- For the specific brand, the price effect is equal to $\beta_2 + \beta_3$, i.e. increasing the price for the specific brand by one unit leads to an expected decrease in the rating only by 0.38 points.