1 Multiple Regression

- A consideration of functional form:

<table>
<thead>
<tr>
<th>Model</th>
<th>Form</th>
<th>Slope = $\frac{dY}{dX}$</th>
<th>Elasticity = $\frac{dY}{dX} \cdot \frac{X}{Y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$Y = a + bX$</td>
<td>$b$</td>
<td>$b \left( \frac{X}{Y} \right)$</td>
</tr>
<tr>
<td>Log-linear</td>
<td>$\ln Y = a + b \ln X$</td>
<td>$b \left( \frac{X}{Y} \right)$</td>
<td>$b \left( \frac{X}{Y} \right)^*$</td>
</tr>
<tr>
<td>Semi-log</td>
<td>$\ln Y = a + b(Y)$</td>
<td>$b(Y)$</td>
<td>$b(X)^*$</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>$Y = a + b \frac{1}{X}$</td>
<td>$-b \left( \frac{1}{X} \right)$</td>
<td>$-b \left( \frac{1}{X^2} \right)^*$</td>
</tr>
</tbody>
</table>

- Where * indicates that the elasticity coefficient is variable depending of the value taken by either $Y$ or $X$. In most instances, use the sample means: $\bar{Y}$ and $\bar{X}$.

- Examine functional forms. Have already seen linear (general case) and semi-log linear (returns to education equation).

- Look at the log-linear model: Cobb-Douglas production function: $Y = AL^aK^\beta$, where $A = e^a$. Taking logs: $\ln Y = a + a \ln L + \beta \ln K$. Also look at the reciprocal model: Phillips curve: $Y = a + b \frac{1}{X}$.

- Inference in the multivariate regression line:

  $$Y_i = a + b_1X_{1i} + b_2X_{2i} + \varepsilon_i$$

- Testing the significance of individual regression coefficients: (dropping $i$ subscripts)

- Intercept term ($a$):

  $$\hat{\sigma}_a^2 = \left[ \frac{1}{n} + \frac{\bar{X}_1^2 \sum x_2^2 + \bar{X}_2 \sum x_1^2 - 2\bar{X}_1 \bar{X}_2 \sum x_1 x_2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \right] \sigma_{YX}^2$$

  and the standard error

  $$\hat{\sigma}_a = \sqrt{\hat{\sigma}_a^2}$$

- Slope coefficients ($b_1 + b_2$):

  $$\hat{\sigma}_{b_1}^2 = \left[ \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \right] \sigma_{YX}^2$$

  $$\hat{\sigma}_{b_1} = \sqrt{\hat{\sigma}_{b_1}^2}$$

  $$\hat{\sigma}_{b_2}^2 = \left[ \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \right] \sigma_{YX}^2$$

  $$\hat{\sigma}_{b_2} = \sqrt{\hat{\sigma}_{b_2}^2}$$

  where $\sigma_{YX}^2 = \frac{\sum \varepsilon_i^2}{n-k}$, $k$ = number of parameters in the model; $n$ is the sample size.
• Sum of squares identity is now:

\[ \sum y^2 = (b_1 \sum x_1 y + b_2 \sum x_2 y) + \sum \varepsilon^2 \]

• Coefficient of determination:

\[ R^2 = \frac{\hat{b}_1 \sum x_1 y + \hat{b}_2 \sum x_2 y}{\sum y^2} \]

• F-test for the joint test \( H_0 : b_1 = b_2 = 0 \):

\[ F^* = \frac{(\hat{b}_1 \sum x_1 y + \hat{b}_2 \sum x_2 y) / (k - 1)}{\sum \varepsilon^2 / (n - k)} \]

\( F^* \sim F_{(k-1),(n-k)} \).

• Relationship between \( R^2 \) and the F-test:

\[ F^* = \frac{R^2 / (k - 1)}{(1 - R^2) / (n - k)} \]

• Leads to an adjustment on \( R^2 \):

\[ R^2 = 1 - (1 - R^2) \left( \frac{n - 1}{n - k} \right) \]