1 Multiple Regression Analysis: Estimation

Three-Variable Model: Notation and Assumptions

- $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

- Two new classical assumptions:
  - No perfect multicollinearity
  - Model is correctly specified

- Gauss-Markov Theorem: OLS is B.L.U.E.

Partial Regression Coefficients

- Interpretation of $\beta_2$ and $\beta_3$

- Terminology:
  - “Holding constant...”
  - “Controlling for...”
  - “Accounting for the influence of...”
  - “Filtering out the effect of...”

- Examples:
  - Housing price hedonics
  - UW enrollment
  - Resource curse
  - Empirics of economic growth

- Partial regressions and graphs

OLS Estimation

- The problem & normal equations
• OLS formulae:

\[ \hat{\beta}_1 = Y - \hat{\beta}_2 \bar{x}_2 - \hat{\beta}_3 \bar{x}_3 \]

\[ \hat{\beta}_2 = \frac{\left( \sum y_i x_{2i} \right) \left( \sum x_{3i}^2 \right) - \left( \sum y_i x_{3i} \right) \left( \sum x_{2i} x_{3i} \right)}{\left( \sum x_{3i}^2 \right) \left( \sum x_{3i}^2 \right) - \left( \sum x_{2i} x_{3i} \right)^2} \]

\[ \hat{\beta}_3 = \frac{\left( \sum y_i x_{3i} \right) \left( \sum x_{2i}^2 \right) - \left( \sum y_i x_{2i} \right) \left( \sum x_{2i} x_{3i} \right)}{\left( \sum x_{2i}^2 \right) \left( \sum x_{3i}^2 \right) - \left( \sum x_{2i} x_{3i} \right)^2} \]

• Standard errors formulae:

\[ se(\hat{\beta}_2) = \sqrt{var(\hat{\beta}_2)} = \sqrt{\frac{\sigma^2}{(1 - r_{23}^2) \sum x_{2i}^2}} \]

\[ se(\hat{\beta}_3) = \sqrt{var(\hat{\beta}_3)} = \sqrt{\frac{\sigma^2}{(1 - r_{23}^2) \sum x_{3i}^2}} \]

Coefficient of Determination, \( R^2 \)

• Formula:

\[ R^2 = \frac{ESS}{TSS} = \frac{\hat{\beta}_2 \left( \sum y_i x_{2i} \right) + \hat{\beta}_3 \left( \sum y_i x_{3i} \right)}{\sum y_i^2} \]

or

\[ R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum \hat{u}_i^2}{\sum y_i^2} \]

• Adjusted \( R^2 \) or \( \bar{R}^2 \):

\[ \bar{R}^2 = 1 - \frac{RSS/(n - k)}{TSS/(n - 1)} \]