Economics 326
Methods of Empirical Research in Economics

Lecture 12: Properties of OLS in the multiple regression model

Vadim Marmer
University of British Columbia

March 3, 2009

Multiple regression and OLS

- Consider the multiple regression model with k regressors:
  $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.

Multiple regression and OLS

- Consider the multiple regression model with k regressors: $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
- Let $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ be the OLS estimators

Multiple regression and OLS

- Consider the multiple regression model with k regressors: $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
- Let $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ be the OLS estimators: if

$$\hat{U}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1,i} - \hat{\beta}_2 X_{2,i} - \ldots - \hat{\beta}_k X_{k,i},$$

Multiple regression and OLS

- Consider the multiple regression model with k regressors: $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
- Let $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ be the OLS estimators: if

$$\hat{U}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1,i} - \hat{\beta}_2 X_{2,i} - \ldots - \hat{\beta}_k X_{k,i},$$

then

$$\sum_{i=1}^{n} \hat{U}_i = \sum_{i=1}^{n} X_{1,i} \hat{U}_i = \ldots = \sum_{i=1}^{n} X_{k,i} \hat{U}_i = 0.$$

Multiple regression and OLS

- As in Lecture 10, we can write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}, \text{ where}$$

Multiple regression and OLS

- As in Lecture 10, we can write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}, \text{ where}$$

  - $\tilde{X}_{1,i}$ are the fitted OLS residuals: $\tilde{X}_{1,i} = X_{1,i} - \hat{\gamma}_0 - \hat{\gamma}_2 X_{2,i} - \ldots - \hat{\gamma}_k X_{k,i}$.

Multiple regression and OLS

- As in Lecture 10, we can write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}, \text{ where}$$

  - $\tilde{X}_{1,i}$ are the fitted OLS residuals: $\tilde{X}_{1,i} = X_{1,i} - \hat{\gamma}_0 - \hat{\gamma}_2 X_{2,i} - \ldots - \hat{\gamma}_k X_{k,i}$.
  - $\hat{\gamma}_0, \hat{\gamma}_2, \ldots, \hat{\gamma}_k$ are the OLS coefficients: $\sum_{i=1}^{n} \tilde{X}_{1,i} = \sum_{i=1}^{n} \tilde{X}_{1,i} X_{2,i} = \ldots = \sum_{i=1}^{n} \tilde{X}_{1,i} X_{k,i} = 0$.

Multiple regression and OLS

- As in Lecture 10, we can write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}, \text{ where}$$

  - $\tilde{X}_{1,i}$ are the fitted OLS residuals: $\tilde{X}_{1,i} = X_{1,i} - \hat{\gamma}_0 - \hat{\gamma}_2 X_{2,i} - \ldots - \hat{\gamma}_k X_{k,i}$.
  - $\hat{\gamma}_0, \hat{\gamma}_2, \ldots, \hat{\gamma}_k$ are the OLS coefficients: $\sum_{i=1}^n \tilde{X}_{1,i} = \sum_{i=1}^n \tilde{X}_{1,i} X_{2,i} = \ldots = \sum_{i=1}^n \tilde{X}_{1,i} X_{k,i} = 0$.
- Similarly, we can write $\hat{\beta}_2$ as

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n \tilde{X}_{2,i} Y_i}{\sum_{i=1}^n \tilde{X}_{2,i}^2}, \text{ where}$$

Multiple regression and OLS

- As in Lecture 10, we can write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}, \text{ where}$$

  - $\tilde{X}_{1,i}$ are the fitted OLS residuals: $\tilde{X}_{1,i} = X_{1,i} - \hat{\gamma}_0 - \hat{\gamma}_2 X_{2,i} - \ldots - \hat{\gamma}_k X_{k,i}$.
  - $\hat{\gamma}_0, \hat{\gamma}_2, \ldots, \hat{\gamma}_k$ are the OLS coefficients: $\sum_{i=1}^{n} \tilde{X}_{1,i} = \sum_{i=1}^{n} \tilde{X}_{1,i} X_{2,i} = \ldots = \sum_{i=1}^{n} \tilde{X}_{1,i} X_{k,i} = 0$.
- Similarly, we can write $\hat{\beta}_2$ as

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} \tilde{X}_{2,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{2,i}^2}, \text{ where}$$

  - $\tilde{X}_{2,i}$ are the fitted OLS residuals: $\tilde{X}_{2,i} = X_{2,i} - \hat{\delta}_0 - \hat{\delta}_1 X_{1,i} - \hat{\delta}_3 X_{3,i} - \ldots - \hat{\delta}_k X_{k,i}$.

Multiple regression and OLS

- As in Lecture 10, we can write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}, \text{ where}$$

  - $\tilde{X}_{1,i}$ are the fitted OLS residuals: $\tilde{X}_{1,i} = X_{1,i} - \hat{\gamma}_0 - \hat{\gamma}_2 X_{2,i} - \ldots - \hat{\gamma}_k X_{k,i}$.
  - $\hat{\gamma}_0, \hat{\gamma}_2, \ldots, \hat{\gamma}_k$ are the OLS coefficients: $\sum_{i=1}^{n} \tilde{X}_{1,i} = \sum_{i=1}^{n} \tilde{X}_{1,i} X_{2,i} = \ldots = \sum_{i=1}^{n} \tilde{X}_{1,i} X_{k,i} = 0$.
- Similarly, we can write $\hat{\beta}_2$ as

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} \tilde{X}_{2,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{2,i}^2}, \text{ where}$$

  - $\tilde{X}_{2,i}$ are the fitted OLS residuals: $\tilde{X}_{2,i} = X_{2,i} - \hat{\delta}_0 - \hat{\delta}_1 X_{1,i} - \hat{\delta}_3 X_{3,i} - \ldots - \hat{\delta}_k X_{k,i}$.
  - $\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_3, \ldots, \hat{\delta}_k$ are the OLS coefficients: $\sum_{i=1}^{n} \tilde{X}_{2,i} = \sum_{i=1}^{n} \tilde{X}_{2,i} X_{1,i} = \sum_{i=1}^{n} \tilde{X}_{2,i} X_{3,i} = \ldots = \sum_{i=1}^{n} \tilde{X}_{2,i} X_{k,i} = 0$.

The OLS estimators are linear

- Consider $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$$

The OLS estimators are linear

- Consider $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \sum_{i=1}^{n} \frac{\tilde{X}_{1,i}}{\sum_{l=1}^{n} \tilde{X}_{1,l}^2} Y_i$$

The OLS estimators are linear

- Consider $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \sum_{i=1}^{n} \frac{\tilde{X}_{1,i}}{\sum_{l=1}^{n} \tilde{X}_{1,l}^2} Y_i = \sum_{i=1}^{n} w_{1,i} Y_i,$$

The OLS estimators are linear

- Consider $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} = \sum_{i=1}^n \frac{\tilde{X}_{1,i}}{\sum_{l=1}^n \tilde{X}_{1,l}^2} Y_i = \sum_{i=1}^n w_{1,i} Y_i,$$

where

$$w_{1,i} = \frac{\tilde{X}_{1,i}}{\sum_{l=1}^n \tilde{X}_{1,l}^2}.$$

The OLS estimators are linear

- Consider $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} = \sum_{i=1}^n \frac{\tilde{X}_{1,i}}{\sum_{l=1}^n \tilde{X}_{1,l}^2} Y_i = \sum_{i=1}^n w_{1,i} Y_i,$$

  where

$$w_{1,i} = \frac{\tilde{X}_{1,i}}{\sum_{l=1}^n \tilde{X}_{1,l}^2}.$$

- Recall that $\tilde{X}_1$ are the residuals from a regression of X_1 against $X_2, \ldots, X_k$ and a constant, and therefore $w_{1,i}$ depends only on X's.

Unbiasedness

- Suppose that
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E(U_i) = 0$ for all i's.

Unbiasedness

- Suppose that
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E(U_i) = 0$ for all i's.

  - Conditioning on X's means that we condition on $X_{1,1}, \ldots, X_{1,n}, X_{2,1}, \ldots, X_{2,n}, \ldots, X_{k,1}, \ldots, X_{k,n}$:

Unbiasedness

- Suppose that
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E(U_i) = 0$ for all i's.

  - Conditioning on X's means that we condition on $X_{1,1}, \ldots, X_{1,n}, X_{2,1}, \ldots, X_{2,n}, \ldots, X_{k,1}, \ldots, X_{k,n}$:

$$E(U_i | X_{1,1}, \ldots, X_{1,n}, X_{2,1}, \ldots, X_{2,n}, \ldots, X_{k,1}, \ldots, X_{k,n}) = 0.$$

Unbiasedness

- Suppose that
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E(U_i) = 0$ for all i's.

  - Conditioning on X's means that we condition on $X_{1,1}, \ldots, X_{1,n}, X_{2,1}, \ldots, X_{2,n}, \ldots, X_{k,1}, \ldots, X_{k,n}$:

$$E(U_i | X_{1,1}, \ldots, X_{1,n}, X_{2,1}, \ldots, X_{2,n}, \ldots, X_{k,1}, \ldots, X_{k,n}) = 0.$$

- Under the above assumptions:

$$\begin{aligned} E\hat{\beta}_0 &= \beta_0, \\ E\hat{\beta}_1 &= \beta_1, \\ &\vdots \\ E\hat{\beta}_k &= \beta_k. \end{aligned}$$

Proof of unbiasedness

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$$

Proof of unbiasedness

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} \quad = \quad \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} \left(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i\right)}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$$

Proof of unbiasedness

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} = \frac{\sum_{i=1}^n \tilde{X}_{1,i} \left(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i\right)}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$$

$$= \beta_0 \frac{\sum_{i=1}^n \tilde{X}_{1,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \beta_1 \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{1,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \beta_2 \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$$

$$+ \ldots + \beta_k \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{k,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

Proof of unbiasedness

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} = \frac{\sum_{i=1}^n \tilde{X}_{1,i} \left(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i\right)}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$$

$$= \beta_0 \frac{\sum_{i=1}^n \tilde{X}_{1,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \beta_1 \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{1,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \beta_2 \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$$

$$+ \ldots + \beta_k \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{k,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

Using the partitioned regression results from Lecture 10:

$$\sum_{i=1}^n \tilde{X}_{1,i} = \sum_{i=1}^n \tilde{X}_{1,i} X_{2,i} = \ldots = \sum_{i=1}^n \tilde{X}_{1,i} X_{k,i} = 0, \sum_{i=1}^n \tilde{X}_{1,i} X_{1,i} = \sum_{i=1}^n \tilde{X}_{1,i}^2.$$

Proof of unbiasedness

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} = \frac{\sum_{i=1}^n \tilde{X}_{1,i} (\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i)}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$$

$$= \beta_0 \frac{\sum_{i=1}^n \tilde{X}_{1,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \beta_1 \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{1,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \beta_2 \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \ldots + \beta_k \frac{\sum_{i=1}^n \tilde{X}_{1,i} X_{k,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2} + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

Using the partitioned regression results from Lecture 10:

$$\sum_{i=1}^n \tilde{X}_{1,i} = \sum_{i=1}^n \tilde{X}_{1,i} X_{2,i} = \ldots = \sum_{i=1}^n \tilde{X}_{1,i} X_{k,i} = 0, \sum_{i=1}^n \tilde{X}_{1,i} X_{1,i} = \sum_{i=1}^n \tilde{X}_{1,i}^2.$$

Therefore,

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

Proof of unbiasedness

- We have that

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

Proof of unbiasedness

- We have that
$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$
- Conditional on X's,
$$E\left(U_i\right) = 0.$$

Proof of unbiasedness

- We have that
$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- Conditional on X's,
$$E\left(U_i\right) = 0.$$

- Therefore, conditional on X's,
$$E\hat{\beta}_1$$

Proof of unbiasedness

- We have that
$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$
- Conditional on X's,
$$E\left(U_i\right) = 0.$$
- Therefore, conditional on X's,
$$E\hat{\beta}_1 = E\left(\beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right)$$

Proof of unbiasedness

- We have that

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- Conditional on X's,

$$E\left(U_i\right) = 0.$$

- Therefore, conditional on X's,

$$\begin{aligned} E\hat{\beta}_1 &= E\left(\beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right) \\ &= \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} E U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} \end{aligned}$$

Proof of unbiasedness

- We have that
$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- Conditional on X's,
$$E\left(U_i\right) = 0.$$

- Therefore, conditional on X's,
$$\begin{aligned} E\hat{\beta}_1 &= E\left(\beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right) \\ &= \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} E U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} \\ &= \beta_1. \end{aligned}$$

Conditional variance of the OLS estimators

- Suppose that:
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E\left(U_i\right) = 0$ for all i's.
  3. Conditional on X's, $E\left(U_i^2\right) = \sigma^2$ for all i's.
  4. Conditional on X's, $E\left(U_i U_j\right) = 0$ for all $i \neq j$.

Conditional variance of the OLS estimators

- Suppose that:
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E\left(U_i\right) = 0$ for all i's.
  3. Conditional on X's, $E\left(U_i^2\right) = \sigma^2$ for all i's.
  4. Conditional on X's, $E\left(U_i U_j\right) = 0$ for all $i \neq j$.
- The conditional variance of $\hat{\beta}_1$ given X's, is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

Conditional variance of the OLS estimators

- Suppose that:
  1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + U_i$.
  2. Conditional on X's, $E\left(U_i\right) = 0$ for all i's.
  3. Conditional on X's, $E\left(U_i^2\right) = \sigma^2$ for all i's.
  4. Conditional on X's, $E\left(U_i U_j\right) = 0$ for all $i \neq j$.
- The conditional variance of $\hat{\beta}_1$ given X's, is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- Gauss-Markov Theorem: Under Assumptions 1-4, the OLS estimators are BLUE.

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$.

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$Var\left(\hat{\beta}_1\right) = E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2$$

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$Var\left(\hat{\beta}_1\right) = E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2 = E\left(\frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right)^2$$

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$
\begin{aligned}
Var\left(\hat{\beta}_1\right) &= E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2 = E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i} U_i\right)^2
\end{aligned}
$$

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$
\begin{aligned}
Var\left(\hat{\beta}_1\right) &= E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2 = E\left(\frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^{n} \tilde{X}_{1,i} U_i\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^{n} \tilde{X}_{1,i}^2 U_i^2 + \sum_{i=1}^{n} \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} U_i U_j\right)
\end{aligned}
$$

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$
\begin{aligned}
Var\left(\hat{\beta}_1\right) &= E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2 = E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i} U_i\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i}^2 U_i^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} U_i U_j\right) \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \left(\sum_{i=1}^n \tilde{X}_{1,i}^2 \sigma^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} 0\right)
\end{aligned}
$$

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$
\begin{aligned}
Var\left(\hat{\beta}_1\right) &= E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2 = E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i} U_i\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i}^2 U_i^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} U_i U_j\right) \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \left(\sum_{i=1}^n \tilde{X}_{1,i}^2 \sigma^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} 0\right) \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \sigma^2 \sum_{i=1}^n \tilde{X}_{1,i}^2
\end{aligned}
$$

Derivation of the conditional variance of OLS

- We have $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$.
- Conditional on X's,

$$
\begin{aligned}
Var\left(\hat{\beta}_1\right) &= E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)^2 = E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i} U_i\right)^2 \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 E\left(\sum_{i=1}^n \tilde{X}_{1,i}^2 U_i^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} U_i U_j\right) \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \left(\sum_{i=1}^n \tilde{X}_{1,i}^2 \sigma^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i} \tilde{X}_{1,j} 0\right) \\
&= \left(\frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)^2 \sigma^2 \sum_{i=1}^n \tilde{X}_{1,i}^2 = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.
\end{aligned}
$$

Conditional covariance of the OLS estimators

- Consider $\hat{\beta}_1$ and $\hat{\beta}_2$

Conditional covariance of the OLS estimators

- Consider $\hat{\beta}_1$ and $\hat{\beta}_2$:

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2},$$
$$\hat{\beta}_2 = \beta_2 + \frac{\sum_{i=1}^n \tilde{X}_{2,i} U_i}{\sum_{i=1}^n \tilde{X}_{2,i}^2},$$

  where
  - $\tilde{X}_1$ are the fitted residuals from the regression of X_1 against a constant and $X_2, X_{3,} \ldots, X_k$.
  - $\tilde{X}_2$ are the fitted residuals from the regression of X_2 against a constant and $X_1, X_3, \ldots, X_k$.

Conditional covariance of the OLS estimators

- Consider $\hat{\beta}_1$ and $\hat{\beta}_2$:

$$\begin{aligned}\hat{\beta}_1 &= \beta_1 + \frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}, \\ \hat{\beta}_2 &= \beta_2 + \frac{\sum_{i=1}^n \tilde{X}_{2,i} U_i}{\sum_{i=1}^n \tilde{X}_{2,i}^2},\end{aligned}$$

  where
  - $\tilde{X}_1$ are the fitted residuals from the regression of X_1 against a constant and $X_2, X_3, \ldots, X_k$.
  - $\tilde{X}_2$ are the fitted residuals from the regression of X_2 against a constant and $X_1, X_3, \ldots, X_k$.
- We will show that given Assumptions 1-4, conditional on X's:

$$Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \sigma^2 \frac{\sum_{i=1}^n \tilde{X}_{1,i} \tilde{X}_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2}$$

Conditional covariance of the OLS estimators

Conditional on X's,

$$Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) =$$

Conditional covariance of the OLS estimators

Conditional on X's,

$$
\begin{aligned}
& Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \\
= \quad & E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)\left(\hat{\beta}_2 - E\hat{\beta}_2\right)
\end{aligned}
$$

Conditional covariance of the OLS estimators

Conditional on X's,

$$\begin{aligned} & Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \\ = \;& E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)\left(\hat{\beta}_2 - E\hat{\beta}_2\right) \\ = \;& E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)\left(\frac{\sum_{i=1}^n \tilde{X}_{2,i} U_i}{\sum_{i=1}^n \tilde{X}_{2,i}^2}\right) \end{aligned}$$

Conditional covariance of the OLS estimators

Conditional on X's,

$$
\begin{aligned}
& Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \\
= \; & E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)\left(\hat{\beta}_2 - E\hat{\beta}_2\right) \\
= \; & E\left(\frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}\right)\left(\frac{\sum_{i=1}^{n} \tilde{X}_{2,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{2,i}^2}\right) \\
= \; & \frac{1}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2 \sum_{i=1}^{n} \tilde{X}_{2,i}^2} E\left(\sum_{i=1}^{n} \tilde{X}_{1,i} U_i\right)\left(\sum_{i=1}^{n} \tilde{X}_{2,i} U_i\right)
\end{aligned}
$$

Conditional covariance of the OLS estimators

Conditional on X's,

$$
\begin{aligned}
& Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \\
= \; & E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)\left(\hat{\beta}_2 - E\hat{\beta}_2\right) \\
= \; & E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)\left(\frac{\sum_{i=1}^n \tilde{X}_{2,i} U_i}{\sum_{i=1}^n \tilde{X}_{2,i}^2}\right) \\
= \; & \frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2} E\left(\sum_{i=1}^n \tilde{X}_{1,i} U_i\right)\left(\sum_{i=1}^n \tilde{X}_{2,i} U_i\right) \\
= \; & \frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2} E\left(\sum_{i=1}^n \tilde{X}_{1,i}\tilde{X}_{2,i} U_i^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i}\tilde{X}_{2,j} U_i U_j\right)
\end{aligned}
$$

Conditional covariance of the OLS estimators

Conditional on X's,

$$
\begin{aligned}
& Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \\
= \; & E\left(\hat{\beta}_1 - E\hat{\beta}_1\right)\left(\hat{\beta}_2 - E\hat{\beta}_2\right) \\
= \; & E\left(\frac{\sum_{i=1}^n \tilde{X}_{1,i} U_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}\right)\left(\frac{\sum_{i=1}^n \tilde{X}_{2,i} U_i}{\sum_{i=1}^n \tilde{X}_{2,i}^2}\right) \\
= \; & \frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2} E\left(\sum_{i=1}^n \tilde{X}_{1,i} U_i\right)\left(\sum_{i=1}^n \tilde{X}_{2,i} U_i\right) \\
= \; & \frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2} E\left(\sum_{i=1}^n \tilde{X}_{1,i}\tilde{X}_{2,i} U_i^2 + \sum_{i=1}^n \sum_{j \neq i} \tilde{X}_{1,i}\tilde{X}_{2,j} U_i U_j\right) \\
= \; & \frac{1}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2} \sum_{i=1}^n \tilde{X}_{1,i}\tilde{X}_{2,i}\sigma^2.
\end{aligned}
$$

Normality of the OLS estimators

- In addition to Assumptions 1-4, assume that conditional on X's, U_i's are jointly normally distributed.

Normality of the OLS estimators

- In addition to Assumptions 1-4, assume that conditional on X's, U_i's are jointly normally distributed.
- $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ are linear estimators

Normality of the OLS estimators

- In addition to Assumptions 1-4, assume that conditional on X's, U_i's are jointly normally distributed.
- $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ are linear estimators:

$$\hat{\beta}_j = \sum_{i=1}^{n} w_{j,i} Y_i = \beta_j + \sum_{i=1}^{n} w_{j,i} U_i,$$

where

$$w_{j,i} = \frac{\tilde{X}_{j,i}}{\sum_{l=1}^{n} \tilde{X}_{j,i}^2},$$

and $\tilde{X}_{j,i}$ are the residuals from the regression of $X_{j,i}$ against the rest of the regressors.

Normality of the OLS estimators

- In addition to Assumptions 1-4, assume that conditional on X's, U_i's are jointly normally distributed.
- $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ are linear estimators:

$$\hat{\beta}_j = \sum_{i=1}^{n} w_{j,i} Y_i = \beta_j + \sum_{i=1}^{n} w_{j,i} U_i,$$

  where

$$w_{j,i} = \frac{\tilde{X}_{j,i}}{\sum_{l=1}^{n} \tilde{X}_{j,i}^2},$$

  and $\tilde{X}_{j,i}$ are the residuals from the regression of $X_{j,i}$ against the rest of the regressors.
- It follows that $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ are jointly normally distributed (conditional on X's).

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.
- We could estimate β_1 by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2}.$$

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.
- We could estimate β_1 by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2}.$$

- Suppose that instead we regress Y against a constant, X_1 and additional $k-1$ regressors $X_2, \ldots, X_k$

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.
- We could estimate β_1 by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2}.$$

- Suppose that instead we regress Y against a constant, X_1 and additional $k-1$ regressors $X_2, \ldots, X_k$, i.e. we estimate β_1 by

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.
- We could estimate β_1 by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2}.$$

- Suppose that instead we regress Y against a constant, X_1 and additional $k-1$ regressors $X_2, \ldots, X_k$, i.e. we estimate β_1 by

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- We have

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} \left(\beta_0 + \beta_1 X_{1,i} + U_i\right)}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$$

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.
- We could estimate β_1 by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2}.$$

- Suppose that instead we regress Y against a constant, X_1 and additional $k-1$ regressors $X_2, \ldots, X_k$, i.e. we estimate β_1 by

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- We have

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} \left(\beta_0 + \beta_1 X_{1,i} + U_i\right)}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

Inclusion of irrelevant regressors

- Suppose that the true model is $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$.
- We could estimate β_1 by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2}.$$

- Suppose that instead we regress Y against a constant, X_1 and additional $k-1$ regressors $X_2, \ldots, X_k$, i.e. we estimate β_1 by

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- We have

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} \left(\beta_0 + \beta_1 X_{1,i} + U_i\right)}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \beta_1 + \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} U_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- Since conditional on X's $E\left(U_i\right) = 0$, $\tilde{\beta}_1$ is unbiased !

Inclusion of irrelevant regressors

- When $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$,

  $\hat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right)^2}$ and $\tilde{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2}$ are both unbiased.

Inclusion of irrelevant regressors

- When $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right)^2} \text{ and } \tilde{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} \text{ are both unbiased.}$$

- Conditional on X's,

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right)^2} \text{ and } Var\left(\tilde{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

Inclusion of irrelevant regressors

- When $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right)^2} \text{ and } \tilde{\beta}_1 = \frac{\sum_{i=1}^n \tilde{X}_{1,i} Y_i}{\sum_{i=1}^n \tilde{X}_{1,i}^2} \text{ are both unbiased.}$$

- Conditional on X's,

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \left(X_{1,i} - \bar{X}_1\right)^2} \text{ and } Var\left(\tilde{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

- Since the true model has only X_1, by Gauss-Markov Theorem $\hat{\beta}_1$ is BLUE and

$$Var\left(\hat{\beta}_1\right) \leq Var\left(\tilde{\beta}_1\right).$$

Inclusion of irrelevant regressors

- When $Y_i = \beta_0 + \beta_1 X_{1,i} + U_i$,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2} \text{ and } \tilde{\beta}_1 = \frac{\sum_{i=1}^{n} \tilde{X}_{1,i} Y_i}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} \text{ are both unbiased.}$$

- Conditional on X's,

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2} \text{ and } Var\left(\tilde{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}.$$

- Since the true model has only X_1, by Gauss-Markov Theorem $\hat{\beta}_1$ is BLUE and

$$Var\left(\hat{\beta}_1\right) \leq Var\left(\tilde{\beta}_1\right).$$

- Without Gauss-Markov Theorem, one can show directly that $\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2 \geq \sum_{i=1}^{n} \tilde{X}_{1,i}^2$.

Proof of $\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2 \geq \sum_{i=1}^{n} \tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i}, \ldots, X_{k,i}$:

$$X_{1,i} = \hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} + \tilde{X}_{1,i}.$$

Proof of $\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2 \geq \sum_{i=1}^{n}\tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i},\ldots,X_{k,i}$:

$$X_{1,i} = \hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} + \tilde{X}_{1,i}.$$

- Consider the sums-of-squares for this regression:

$$SST_1 = \sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2,$$

Proof of $\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2 \geq \sum_{i=1}^{n}\tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i},\ldots,X_{k,i}$:

$$X_{1,i} = \hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} + \tilde{X}_{1,i}.$$

- Consider the sums-of-squares for this regression:

$$\begin{aligned} SST_1 &= \sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2, \\ SSE_1 &= \sum_{i=1}^{n}\left(\hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} - \bar{X}_1\right)^2, \end{aligned}$$

Proof of $\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2 \geq \sum_{i=1}^{n}\tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i},\ldots,X_{k,i}$:

$$X_{1,i} = \hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} + \tilde{X}_{1,i}.$$

- Consider the sums-of-squares for this regression:

$$\begin{aligned} SST_1 &= \sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2, \\ SSE_1 &= \sum_{i=1}^{n}\left(\hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} - \bar{X}_1\right)^2, \\ SSR_1 &= \sum_{i=1}^{n}\tilde{X}_{1,i}^2. \end{aligned}$$

Proof of $\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2 \geq \sum_{i=1}^{n}\tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i},\ldots,X_{k,i}$:

$$X_{1,i}=\hat{\gamma}_0+\hat{\gamma}_2 X_{2,i}+\ldots+\hat{\gamma}_k X_{k,i}+\tilde{X}_{1,i}.$$

- Consider the sums-of-squares for this regression:

$$\begin{aligned} SST_1 &= \sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2, \\ SSE_1 &= \sum_{i=1}^{n}\left(\hat{\gamma}_0+\hat{\gamma}_2 X_{2,i}+\ldots+\hat{\gamma}_k X_{k,i}-\bar{X}_1\right)^2, \\ SSR_1 &= \sum_{i=1}^{n}\tilde{X}_{1,i}^2. \end{aligned}$$

- Thus,

$$\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2-\sum_{i=1}^{n}\tilde{X}_{1,i}^2$$

Proof of $\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2 \geq \sum_{i=1}^{n} \tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i}, \ldots, X_{k,i}$:

$$X_{1,i} = \hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} + \tilde{X}_{1,i}.$$

- Consider the sums-of-squares for this regression:

$$\begin{aligned} SST_1 &= \sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2, \\ SSE_1 &= \sum_{i=1}^{n} \left(\hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} - \bar{X}_1\right)^2, \\ SSR_1 &= \sum_{i=1}^{n} \tilde{X}_{1,i}^2. \end{aligned}$$

- Thus,

$$\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_1\right)^2 - \sum_{i=1}^{n} \tilde{X}_{1,i}^2 = SST_1 - SSR_1$$

Proof of $\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2 \geq \sum_{i=1}^{n}\tilde{X}_{1,i}^2$

- $\tilde{X}_{1,i}$ are the fitted residuals from regressing $X_{1,i}$ against a constant, $X_{2,i},\ldots,X_{k,i}$:

$$X_{1,i} = \hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} + \tilde{X}_{1,i}.$$

- Consider the sums-of-squares for this regression:

$$\begin{aligned} SST_1 &= \sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2, \\ SSE_1 &= \sum_{i=1}^{n}\left(\hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \ldots + \hat{\gamma}_k X_{k,i} - \bar{X}_1\right)^2, \\ SSR_1 &= \sum_{i=1}^{n}\tilde{X}_{1,i}^2. \end{aligned}$$

- Thus,

$$\sum_{i=1}^{n}\left(X_{1,i}-\bar{X}_1\right)^2 - \sum_{i=1}^{n}\tilde{X}_{1,i}^2 = SST_1 - SSR_1 = SSE_1 \geq 0.$$

$Var\left(\hat{\beta}_1\right)$ and the number of regressors k

- In $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$, the variance of the OLS estimator $\hat{\beta}_1$ is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2}$$

$Var\left(\hat{\beta}_1\right)$ and the number of regressors k

- In $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$, the variance of the OLS estimator $\hat{\beta}_1$ is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \frac{\sigma^2}{SSR_1},$$

where SSR_1 is the residual sum-of-squares from the regression of X_1 against a constant and the rest of the regressors.

$Var\left(\hat{\beta}_1\right)$ and the number of regressors k

- In $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$, the variance of the OLS estimator $\hat{\beta}_1$ is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \frac{\sigma^2}{SSR_1},$$

  where SSR_1 is the residual sum-of-squares from the regression of X_1 against a constant and the rest of the regressors.
- Since SSR_1 can only decrease when we add more regressors, $Var\left(\hat{\beta}_1\right)$ increases with k, if the added regressors are irrelevant but correlated with the included regressors.

$Var\left(\hat{\beta}_1\right)$ and the number of regressors k

- In $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$, the variance of the OLS estimator $\hat{\beta}_1$ is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \frac{\sigma^2}{SSR_1},$$

  where SSR_1 is the residual sum-of-squares from the regression of X_1 against a constant and the rest of the regressors.
- Since SSR_1 can only decrease when we add more regressors, $Var\left(\hat{\beta}_1\right)$ increases with k, if the added regressors are irrelevant but correlated with the included regressors.
- If the added regressors are uncorrelated with X_1, inclusion of such regressors will not affect SSR_1 (in large samples) or the variance of $\hat{\beta}_1$.

$Var\left(\hat{\beta}_1\right)$ and the number of regressors k

- In $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$, the variance of the OLS estimator $\hat{\beta}_1$ is

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^{n} \tilde{X}_{1,i}^2} = \frac{\sigma^2}{SSR_1},$$

  where SSR_1 is the residual sum-of-squares from the regression of X_1 against a constant and the rest of the regressors.
- Since SSR_1 can only decrease when we add more regressors, $Var\left(\hat{\beta}_1\right)$ increases with k, if the added regressors are irrelevant but correlated with the included regressors.
- If the added regressors are uncorrelated with X_1, inclusion of such regressors will not affect SSR_1 (in large samples) or the variance of $\hat{\beta}_1$.
- If the added regressors are uncorrelated with X_1 and affect Y, their inclusion will reduce σ^2 without affecting SSR_1 and will reduce the variance of $\hat{\beta}_1$.

Estimation of variances and covariances

- In n $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$,

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2} \text{ and } Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \sigma^2 \frac{\sum_{i=1}^n \tilde{X}_{1,i}\tilde{X}_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2}.$$

- Variances and covariances can be estimated by replacing σ^2 with

$$s^2 = \frac{1}{n-k-1} \sum_{i=1}^n \hat{U}_i^2.$$

Estimation of variances and covariances

- In n $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \ldots + \hat{\beta}_k X_{k,i} + \hat{U}_i$,

$$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2} \text{ and } Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \sigma^2 \frac{\sum_{i=1}^n \tilde{X}_{1,i}\tilde{X}_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2}.$$

- Variances and covariances can be estimated by replacing σ^2 with

$$s^2 = \frac{1}{n-k-1} \sum_{i=1}^n \hat{U}_i^2.$$

- Estimated variance and covariance:

$$\widehat{Var}\left(\hat{\beta}_1\right) = \frac{s^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2} \text{ and } \widehat{Cov}\left(\hat{\beta}_1, \hat{\beta}_2\right) = s^2 \frac{\sum_{i=1}^n \tilde{X}_{1,i}\tilde{X}_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2}.$$