

Introductory Econometrics

Lecture 12: Properties of OLS in the multiple regression model

Instructor: Ma, Jun

Renmin University of China

April 26, 2023

Multiple regression and OLS

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

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

then

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

- ▶ 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:
 $\tilde{X}_{1,i} = X_{1,i} - \hat{\gamma}_0 - \hat{\gamma}_2 X_{2,i} - \dots - \hat{\gamma}_k X_{k,i}$.
- ▶ $\hat{\gamma}_0, \hat{\gamma}_2, \dots, \hat{\gamma}_k$ are the OLS coefficients:
 $\sum_{i=1}^n \tilde{X}_{1,i} = \sum_{i=1}^n \tilde{X}_{1,i} X_{2,i} = \dots = \sum_{i=1}^n \tilde{X}_{1,i} X_{k,i} = 0$.
- ▶ 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:
 $\tilde{X}_{2,i} = X_{2,i} - \hat{\delta}_0 - \hat{\delta}_1 X_{1,i} - \hat{\delta}_3 X_{3,i} - \dots - \hat{\delta}_k X_{k,i}$.
- ▶ $\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_3, \dots, \hat{\delta}_k$ are the OLS coefficients: $\sum_{i=1}^n \tilde{X}_{2,i} = \sum_{i=1}^n \tilde{X}_{2,i} X_{1,i} = \sum_{i=1}^n \tilde{X}_{2,i} X_{3,i} = \dots = \sum_{i=1}^n \tilde{X}_{2,i} X_{k,i} = 0$.

The OLS estimators are linear

- ▶ Consider $\hat{\beta}_1$:

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

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, \dots, X_k and a constant, and therefore $w_{1,i}$ depends only on X 's.

Unbiasedness

► Suppose that

1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \dots + \beta_k X_{k,i} + U_i.$

2. Conditional on X 's, $E[U_i] = 0$ for all i 's.

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

$$E[U_i | X_{1,1}, \dots, X_{1,n}, X_{2,1}, \dots, X_{2,n}, \dots, X_{k,1}, \dots, X_{k,n}] = 0.$$

► Under the above assumptions:

$$E[\hat{\beta}_0] = \beta_0,$$

$$E[\hat{\beta}_1] = \beta_1,$$

$$\vdots \quad \vdots \quad \vdots$$

$$E[\hat{\beta}_k] = \beta_k.$$

Proof of unbiasedness

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

Using the partitioned regression results from Lecture 10:

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

Therefore,

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

- ▶ 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,

$$E[U_i] = 0.$$

- ▶ Therefore, conditional on X 's,

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

Conditional variance of the OLS estimators

► Suppose that:

1. $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \dots + \beta_k X_{k,i} + U_i$.
2. Conditional on X 's, $E[U_i] = 0$ for all i 's.
3. Conditional on X 's, $E[U_i^2] = \sigma^2$ for all i 's.
4. Conditional on X 's, $E[U_i U_j] = 0$ for all $i \neq j$.

► The conditional variance of $\hat{\beta}_1$ given X 's, is

$$\text{Var} [\hat{\beta}_1] = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2}.$$

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

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

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

Conditional covariance of the OLS estimators

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

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

where

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

$$\text{Cov} [\hat{\beta}_1, \hat{\beta}_2] = \sigma^2 \frac{\sum_{i=1}^n \tilde{X}_{1,i} \tilde{X}_{2,i}}{\sum_{i=1}^n \tilde{X}_{1,i}^2 \sum_{i=1}^n \tilde{X}_{2,i}^2}$$

Conditional on X 's,

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

Normality of the OLS estimators

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

Inclusion of irrelevant regressors

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

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

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

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

- ▶ Since conditional on X 's $E[U_i] = 0$, $\tilde{\beta}_1$ is unbiased !

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

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

- ▶ Conditional on X 's,

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

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

$$\text{Var} [\hat{\beta}_1] \leq \text{Var} [\tilde{\beta}_1].$$

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

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

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

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

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

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

$$SSE_1 = \sum_{i=1}^n (\hat{\gamma}_0 + \hat{\gamma}_2 X_{2,i} + \dots + \hat{\gamma}_k X_{k,i} - \bar{X}_1)^2,$$

$$SSR_1 = \sum_{i=1}^n \tilde{X}_{1,i}^2.$$

- ▶ Thus,

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

Var $[\hat{\beta}_1]$ and the number of regressors k

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

$$\text{Var} [\hat{\beta}_1] = \frac{\sigma^2}{\sum_{i=1}^n \tilde{X}_{1,i}^2} = \frac{\sigma^2}{SSR_1},$$

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, $\text{Var} [\hat{\beta}_1]$ 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 $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1,i} + \hat{\beta}_2 X_{2,i} + \dots + \hat{\beta}_k X_{k,i} + \hat{U}_i$,

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

- ▶ 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:

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