ECO 344 - Applied Econometrics I
Chapter 5

Muhammad Salman Khalid

School of Economics & Social Sciences

October 2, 2026



Outline

1. Introduction
2. Consistency
3. Asymptotic Normality
4. The Lagrange Multiplier Statistic
5. Asymptotic Efficiency
6. Final Paper & Presentation

Introduction

Introduction

- So far, everything we proved about OLS holds for **any sample size** n. These are called finite sample (or exact) properties.

Introduction

- So far, everything we proved about OLS holds for **any sample size** n. These are called finite sample (or exact) properties.
  - Unbiasedness under M.L.R - I through M.L.R - IV.

Introduction

- So far, everything we proved about OLS holds for **any sample size** n. These are called finite sample (or exact) properties.
  - Unbiasedness under M.L.R - I through M.L.R - IV.
  - BLUE under M.L.R - I through M.L.R - V (Gauss-Markov).

Introduction

- So far, everything we proved about OLS holds for **any sample size** n. These are called finite sample (or exact) properties.
  - Unbiasedness under M.L.R - I through M.L.R - IV.
  - BLUE under M.L.R - I through M.L.R - V (Gauss-Markov).
  - Exact t and F distributions under M.L.R - I through M.L.R - VI.

Introduction

- So far, everything we proved about OLS holds for **any sample size** n. These are called finite sample (or exact) properties.
  - Unbiasedness under M.L.R - I through M.L.R - IV.
  - BLUE under M.L.R - I through M.L.R - V (Gauss-Markov).
  - Exact t and F distributions under M.L.R - I through M.L.R - VI.
- This chapter is all about relaxing our assumptions and making them a little more realistic.

Introduction

- So far, everything we proved about OLS holds for **any sample size** n. These are called finite sample (or exact) properties.
  - Unbiasedness under M.L.R - I through M.L.R - IV.
  - BLUE under M.L.R - I through M.L.R - V (Gauss-Markov).
  - Exact t and F distributions under M.L.R - I through M.L.R - VI.
- This chapter is all about relaxing our assumptions and making them a little more realistic.
- The price we pay is that the results only hold when the sample size is large.

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.
  1. M.L.R - IV - Zero Conditional Mean - $E[\mu|x_1, x_2, ..., x_k] = 0$

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.
  1. M.L.R - IV - Zero Conditional Mean - $E[\mu|x_1, x_2, ..., x_k] = 0$
  2. M.L.R - VI - Normality of the Error Term

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.
  1. M.L.R - IV - Zero Conditional Mean - $E[\mu|x_1, x_2, ..., x_k] = 0$
  2. M.L.R - VI - Normality of the Error Term
- In this chapter we will see that:

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.
  1. M.L.R - IV - Zero Conditional Mean - $E[\mu|x_1, x_2, ..., x_k] = 0$
  2. M.L.R - VI - Normality of the Error Term
- In this chapter we will see that:
  - M.L.R - IV can be replaced by a weaker assumption if we only want consistency.

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.
  1. M.L.R - IV - Zero Conditional Mean - $E[\mu|x_1, x_2, ..., x_k] = 0$
  2. M.L.R - VI - Normality of the Error Term
- In this chapter we will see that:
  - M.L.R - IV can be replaced by a weaker assumption if we only want consistency.
  - M.L.R - VI can be dropped completely if the sample size is large.

Which Assumptions Are We Worried About?

- We have assumptions that are not met in many applications.
  1. M.L.R - IV - Zero Conditional Mean - $E[\mu|x_1, x_2, ..., x_k] = 0$
  2. M.L.R - VI - Normality of the Error Term
- In this chapter we will see that:
  - M.L.R - IV can be replaced by a weaker assumption if we only want consistency.
  - M.L.R - VI can be dropped completely if the sample size is large.
- M.L.R - V (Homoskedasticity, $Var(\mu|x_1, ..., x_k) = \sigma^2$) is still needed for the usual inference. We relax it in Chapter 8.

Probability Limit

- Similar to limits in calculus, the **probability limit** tells us where an estimator settles down as the sample size n approaches ∞.

Probability Limit

- Similar to limits in calculus, the **probability limit** tells us where an estimator settles down as the sample size n approaches ∞.
- We write

$$\text{plim } \hat{\beta}_j = 3$$

Probability Limit

- Similar to limits in calculus, the **probability limit** tells us where an estimator settles down as the sample size n approaches ∞.
- We write

$$\text{plim } \hat{\beta}_j = 3$$

- This means that as n approaches ∞, the probability that $\hat{\beta}_j$ is very close to 3 approaches one.

Probability Limit

- Similar to limits in calculus, the **probability limit** tells us where an estimator settles down as the sample size n approaches ∞.
- We write

$$\text{plim } \hat{\beta}_j = 3$$

- This means that as n approaches ∞, the probability that $\hat{\beta}_j$ is very close to 3 approaches one.
- In other words, the distribution of $\hat{\beta}_j$ collapses around 3 as we collect more and more data.

Probability Limit

- Similar to limits in calculus, the **probability limit** tells us where an estimator settles down as the sample size n approaches ∞.
- We write

$$\text{plim } \hat{\beta}_j = 3$$

- This means that as n approaches ∞, the probability that $\hat{\beta}_j$ is very close to 3 approaches one.
- In other words, the distribution of $\hat{\beta}_j$ collapses around 3 as we collect more and more data.
- The properties of OLS estimators when n approaches ∞ are called **asymptotic properties** or **large sample properties**.

Consistency

Consistency

- The large sample equivalent of unbiasedness is known as **consistency**.

Consistency

- The large sample equivalent of unbiasedness is known as **consistency**.
- An estimator $\hat{\beta}_j$ is said to be consistent if and only if

$$\text{plim } \hat{\beta}_j = \beta_j$$

Consistency

- The large sample equivalent of unbiasedness is known as **consistency**.
- An estimator $\hat{\beta}_j$ is said to be consistent if and only if

$$\text{plim } \hat{\beta}_j = \beta_j$$

- If obtaining a larger and larger sample makes the estimator converge to the parameter, then the estimator is consistent.

Consistency

- The large sample equivalent of unbiasedness is known as **consistency**.
- An estimator $\hat{\beta}_j$ is said to be consistent if and only if

$$\text{plim } \hat{\beta}_j = \beta_j$$

- If obtaining a larger and larger sample makes the estimator converge to the parameter, then the estimator is consistent.
- An estimator can be biased but still consistent.

Consistency

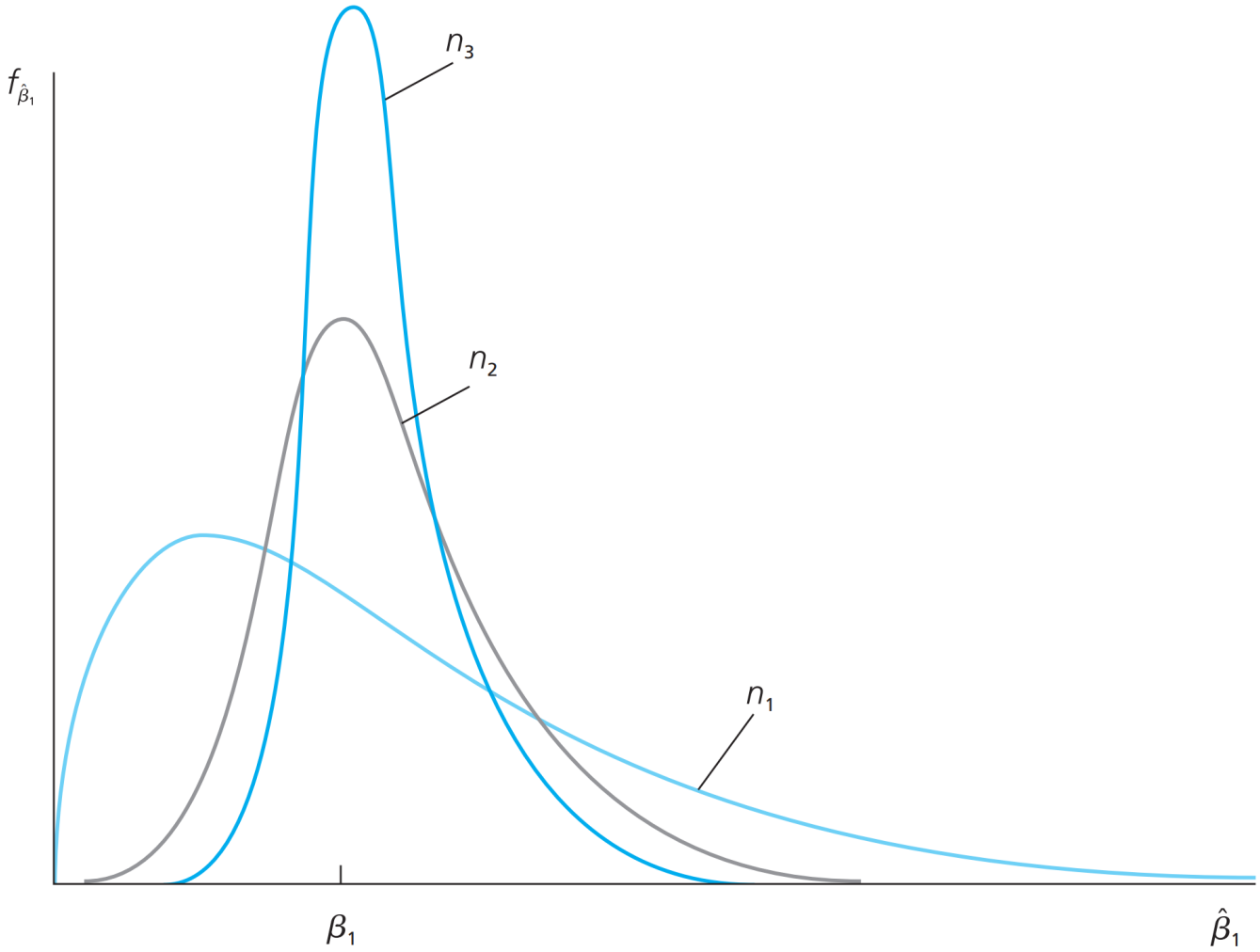
- The large sample equivalent of unbiasedness is known as **consistency**.
- An estimator $\hat{\beta}_j$ is said to be consistent if and only if

$$\text{plim } \hat{\beta}_j = \beta_j$$

- If obtaining a larger and larger sample makes the estimator converge to the parameter, then the estimator is consistent.
- An estimator can be biased but still consistent.
- Consistency is the minimum we ask of an estimator. If more data does not get us closer to the truth, the estimator is of little use.

Consistency

FIGURE 5.1 Sampling distributions of $\hat{\beta}_1$ for sample sizes $n_1 < n_2 < n_3$.



Reading the Figure

- For each sample size n, $\hat{\beta}_1$ has its own sampling distribution.

Reading the Figure

- For each sample size n, $\hat{\beta}_1$ has its own sampling distribution.
- Because OLS is unbiased under M.L.R - I through M.L.R - IV, each distribution is centred at β_1.

Reading the Figure

- For each sample size n, $\hat{\beta}_1$ has its own sampling distribution.
- Because OLS is unbiased under M.L.R - I through M.L.R - IV, each distribution is centred at β_1.
- As n increases from n_1 to n_2 to n_3, the distribution becomes more and more tightly packed around β_1.

Reading the Figure

- For each sample size n, $\hat{\beta}_1$ has its own sampling distribution.
- Because OLS is unbiased under M.L.R - I through M.L.R - IV, each distribution is centred at β_1.
- As n increases from n_1 to n_2 to n_3, the distribution becomes more and more tightly packed around β_1.
- As n approaches ∞, the distribution collapses to the single point β_1.

Reading the Figure

- For each sample size n, $\hat{\beta}_1$ has its own sampling distribution.
- Because OLS is unbiased under M.L.R - I through M.L.R - IV, each distribution is centred at β_1.
- As n increases from n_1 to n_2 to n_3, the distribution becomes more and more tightly packed around β_1.
- As n approaches ∞, the distribution collapses to the single point β_1.
- Recall from Chapter 3 that $Var(\hat{\beta}_j)$ falls as SST_j grows, and SST_j grows with n. This is why the distribution keeps shrinking.

Theorem 5.1 - Consistency of OLS

- **Theorem:**

  Under the assumptions M.L.R - I through M.L.R - IV, the OLS estimator $\hat{\beta}_j$ is consistent for β_j for all $j = 0, 1, 2, ..., k$.

Theorem 5.1 - Consistency of OLS

- **Theorem:**

  Under the assumptions M.L.R - I through M.L.R - IV, the OLS estimator $\hat{\beta}_j$ is consistent for β_j for all $j = 0, 1, 2, ..., k$.

- So the same four assumptions that give us unbiasedness also give us consistency.

Theorem 5.1 - Consistency of OLS

- **Theorem:**

  Under the assumptions M.L.R - I through M.L.R - IV, the OLS estimator $\hat{\beta}_j$ is consistent for β_j for all $j = 0, 1, 2, ..., k$.

- So the same four assumptions that give us unbiasedness also give us consistency.
- The interesting question is whether we can get consistency with **less** than M.L.R - IV.

Why is OLS Consistent?

- We know that in simple regression:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Why is OLS Consistent?

- We know that in simple regression:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- Substituting $y_i = \beta_0 + \beta_1 x_i + \mu_i$:

$$\hat{\beta}_1 = \beta_1 + \frac{n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})\mu_i}{n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Why is OLS Consistent?

- We know that in simple regression:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- Substituting $y_i = \beta_0 + \beta_1 x_i + \mu_i$:

$$\hat{\beta}_1 = \beta_1 + \frac{n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})\mu_i}{n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- By the law of large numbers, the numerator converges to $Cov(x, \mu)$ and the denominator converges to $Var(x)$.

$$\text{plim } \hat{\beta}_1 = \beta_1 + \frac{Cov(x, \mu)}{Var(x)}$$

M.L.R - IV′ Assumption

- We know that M.L.R - IV (Zero Conditional Mean) is required for the OLS estimator to be unbiased.

$$E[\mu|x_1, x_2, x_3, ..., x_k] = 0$$

M.L.R - IV′ Assumption

- We know that M.L.R - IV (Zero Conditional Mean) is required for the OLS estimator to be unbiased.

$$E[\mu|x_1, x_2, x_3, ..., x_k] = 0$$

- As we discussed, this is a very strict assumption. For consistency, we just need the following.

$$E[\mu] = 0 \ \ \& \ \ Cov(x_j, \mu) = 0 \ \ \forall \, j = 1, 2, 3, ..., k$$

M.L.R - IV′ Assumption

- We know that M.L.R - IV (Zero Conditional Mean) is required for the OLS estimator to be unbiased.

$$E[\mu|x_1, x_2, x_3, ..., x_k] = 0$$

- As we discussed, this is a very strict assumption. For consistency, we just need the following.

$$E[\mu] = 0 \ \ \& \ \ Cov(x_j, \mu) = 0 \ \ \forall \, j = 1, 2, 3, ..., k$$

- This is **M.L.R - IV′** (Zero Mean and Zero Correlation).

M.L.R - IV′ Assumption

- We know that M.L.R - IV (Zero Conditional Mean) is required for the OLS estimator to be unbiased.

$$E[\mu|x_1, x_2, x_3, ..., x_k] = 0$$

- As we discussed, this is a very strict assumption. For consistency, we just need the following.

$$E[\mu] = 0 \ \ \& \ \ Cov(x_j, \mu) = 0 \ \ \forall \, j = 1, 2, 3, ..., k$$

- This is **M.L.R - IV′** (Zero Mean and Zero Correlation).
- Now we do not need the error to be mean independent of the explanatory variables. We just need each x_j to be uncorrelated with the error.

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.
- Under M.L.R - IV, μ is unrelated to **any function** of the explanatory variables, such as x_1^2 or $x_1 x_2$.

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.
- Under M.L.R - IV, μ is unrelated to **any function** of the explanatory variables, such as x_1^2 or $x_1 x_2$.
- Under M.L.R - IV′, μ is only uncorrelated with the x_j themselves.

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.
- Under M.L.R - IV, μ is unrelated to **any function** of the explanatory variables, such as x_1^2 or $x_1 x_2$.
- Under M.L.R - IV′, μ is only uncorrelated with the x_j themselves.
- With only M.L.R - IV′:

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.
- Under M.L.R - IV, μ is unrelated to **any function** of the explanatory variables, such as x_1^2 or $x_1 x_2$.
- Under M.L.R - IV′, μ is only uncorrelated with the x_j themselves.
- With only M.L.R - IV′:
  1. OLS is consistent, but it can be biased in small samples.

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.
- Under M.L.R - IV, μ is unrelated to **any function** of the explanatory variables, such as x_1^2 or $x_1 x_2$.
- Under M.L.R - IV′, μ is only uncorrelated with the x_j themselves.
- With only M.L.R - IV′:
  1. OLS is consistent, but it can be biased in small samples.
  2. We may not have correctly modelled $E[y|x_1, ..., x_k]$, so we cannot always read β_j as a partial effect on the conditional mean.

M.L.R - IV versus M.L.R - IV′

- M.L.R - IV implies M.L.R - IV′, but not the other way around.
- Under M.L.R - IV, μ is unrelated to **any function** of the explanatory variables, such as x_1^2 or $x_1 x_2$.
- Under M.L.R - IV′, μ is only uncorrelated with the x_j themselves.
- With only M.L.R - IV′:
  1. OLS is consistent, but it can be biased in small samples.
  2. We may not have correctly modelled $E[y|x_1, ..., x_k]$, so we cannot always read β_j as a partial effect on the conditional mean.
- It is important to note that this weaker result only applies to large samples.

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.
- In simple regression, the **inconsistency** (or asymptotic bias) is

$$\text{plim}\ \hat{\beta}_1 - \beta_1 = \frac{Cov(x_1, \mu)}{Var(x_1)}$$

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.
- In simple regression, the **inconsistency** (or asymptotic bias) is

$$\text{plim } \hat{\beta}_1 - \beta_1 = \frac{Cov(x_1, \mu)}{Var(x_1)}$$

1. If $Cov(x_1, \mu) = 0$ - Consistent

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.
- In simple regression, the **inconsistency** (or asymptotic bias) is

$$\text{plim } \hat{\beta}_1 - \beta_1 = \frac{Cov(x_1, \mu)}{Var(x_1)}$$

1. If $Cov(x_1, \mu) = 0$ - Consistent
2. If $Cov(x_1, \mu) > 0$ - Positive inconsistency

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.
- In simple regression, the **inconsistency** (or asymptotic bias) is

$$\text{plim } \hat{\beta}_1 - \beta_1 = \frac{Cov(x_1, \mu)}{Var(x_1)}$$

1. If $Cov(x_1, \mu) = 0$ - Consistent
2. If $Cov(x_1, \mu) > 0$ - Positive inconsistency
3. If $Cov(x_1, \mu) < 0$ - Negative inconsistency

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.
- In simple regression, the **inconsistency** (or asymptotic bias) is

$$\text{plim } \hat{\beta}_1 - \beta_1 = \frac{Cov(x_1, \mu)}{Var(x_1)}$$

1. If $Cov(x_1, \mu) = 0$ - Consistent
2. If $Cov(x_1, \mu) > 0$ - Positive inconsistency
3. If $Cov(x_1, \mu) < 0$ - Negative inconsistency
4. If $Cov(x_1, \mu)$ is small relative to $Var(x_1)$ - Negligible inconsistency

Inconsistency

- If the error is correlated with **any** of the explanatory variables, OLS is both biased and inconsistent.
- In simple regression, the **inconsistency** (or asymptotic bias) is

$$\text{plim } \hat{\beta}_1 - \beta_1 = \frac{Cov(x_1, \mu)}{Var(x_1)}$$

1. If $Cov(x_1, \mu) = 0$ - Consistent
2. If $Cov(x_1, \mu) > 0$ - Positive inconsistency
3. If $Cov(x_1, \mu) < 0$ - Negative inconsistency
4. If $Cov(x_1, \mu)$ is small relative to $Var(x_1)$ - Negligible inconsistency

- We cannot compute this number because μ is unobserved, but we can often guess its sign.

Inconsistency and Omitted Variables

- Recall the omitted variable bias from Chapter 3. Suppose the true model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + v$$

Inconsistency and Omitted Variables

- Recall the omitted variable bias from Chapter 3. Suppose the true model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + v$$

- We omit x_2 and run the simple regression of y on x_1. Call the slope $\tilde{\beta}_1$. Then

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_2 \delta_1 \quad \text{where} \quad \delta_1 = \frac{Cov(x_1, x_2)}{Var(x_1)}$$

Inconsistency and Omitted Variables

- Recall the omitted variable bias from Chapter 3. Suppose the true model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + v$$

- We omit x_2 and run the simple regression of y on x_1. Call the slope $\tilde{\beta}_1$. Then

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_2 \delta_1 \quad \text{where} \quad \delta_1 = \frac{Cov(x_1, x_2)}{Var(x_1)}$$

- The direction of the inconsistency follows the same table we used for the bias.

Inconsistency and Omitted Variables

- Recall the omitted variable bias from Chapter 3. Suppose the true model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + v$$

- We omit x_2 and run the simple regression of y on x_1. Call the slope $\tilde{\beta}_1$. Then

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_2 \delta_1 \quad \text{where} \quad \delta_1 = \frac{Cov(x_1, x_2)}{Var(x_1)}$$

- The direction of the inconsistency follows the same table we used for the bias.
  - $\beta_2 > 0$ and $Cov(x_1, x_2) > 0$ - Positive inconsistency

Inconsistency and Omitted Variables

- Recall the omitted variable bias from Chapter 3. Suppose the true model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + v$$

- We omit x_2 and run the simple regression of y on x_1. Call the slope $\tilde{\beta}_1$. Then

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_2 \delta_1 \quad \text{where} \quad \delta_1 = \frac{Cov(x_1, x_2)}{Var(x_1)}$$

- The direction of the inconsistency follows the same table we used for the bias.
  - $\beta_2 > 0$ and $Cov(x_1, x_2) > 0$ - Positive inconsistency
  - $\beta_2 > 0$ and $Cov(x_1, x_2) < 0$ - Negative inconsistency

Inconsistency and Omitted Variables

- Recall the omitted variable bias from Chapter 3. Suppose the true model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + v$$

- We omit x_2 and run the simple regression of y on x_1. Call the slope $\tilde{\beta}_1$. Then

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_2 \delta_1 \quad \text{where} \quad \delta_1 = \frac{Cov(x_1, x_2)}{Var(x_1)}$$

- The direction of the inconsistency follows the same table we used for the bias.
  - $\beta_2 > 0$ and $Cov(x_1, x_2) > 0$ - Positive inconsistency
  - $\beta_2 > 0$ and $Cov(x_1, x_2) < 0$ - Negative inconsistency
- The difference is that inconsistency is expressed with population variance and covariance, whereas bias uses their sample counterparts.

Asymptotic Normality

M.L.R - VI - Normality Assumption

- Consistency tells us where the estimator ends up. For inference, we need the whole sampling distribution.

M.L.R - VI - Normality Assumption

- Consistency tells us where the estimator ends up. For inference, we need the whole sampling distribution.
- We know that under M.L.R - VI, we can do exact inference (t-test & F-test).

M.L.R - VI - Normality Assumption

- Consistency tells us where the estimator ends up. For inference, we need the whole sampling distribution.
- We know that under M.L.R - VI, we can do exact inference (t-test & F-test).
- However, as we discussed in Chapter 4, M.L.R - VI can be challenged in most cases.

M.L.R - VI - Normality Assumption

- Consistency tells us where the estimator ends up. For inference, we need the whole sampling distribution.
- We know that under M.L.R - VI, we can do exact inference (t-test & F-test).
- However, as we discussed in Chapter 4, M.L.R - VI can be challenged in most cases.
  - Wages, arrests, savings and many other outcomes are skewed or take only a few values.

M.L.R - VI - Normality Assumption

- Consistency tells us where the estimator ends up. For inference, we need the whole sampling distribution.
- We know that under M.L.R - VI, we can do exact inference (t-test & F-test).
- However, as we discussed in Chapter 4, M.L.R - VI can be challenged in most cases.
  - Wages, arrests, savings and many other outcomes are skewed or take only a few values.
- If μ is not normal, the t statistic does not have an exact t distribution and the F statistic does not have an exact F distribution.

M.L.R - VI - Normality Assumption

- Consistency tells us where the estimator ends up. For inference, we need the whole sampling distribution.
- We know that under M.L.R - VI, we can do exact inference (t-test & F-test).
- However, as we discussed in Chapter 4, M.L.R - VI can be challenged in most cases.
  - Wages, arrests, savings and many other outcomes are skewed or take only a few values.
- If μ is not normal, the t statistic does not have an exact t distribution and the F statistic does not have an exact F distribution.
- Note that unbiasedness and BLUE never required normality. Only exact inference did.

Central Limit Theorem to the Rescue

- When we apply the **central limit theorem**, we can conclude that as the sample size grows, the distribution of the OLS estimators approaches the normal distribution. This is called **asymptotic normality**.

Central Limit Theorem to the Rescue

- When we apply the **central limit theorem**, we can conclude that as the sample size grows, the distribution of the OLS estimators approaches the normal distribution. This is called **asymptotic normality**.
- It is important to note that the distribution of the error in the population has nothing to do with the sample size. A skewed error stays skewed.

Central Limit Theorem to the Rescue

- When we apply the **central limit theorem**, we can conclude that as the sample size grows, the distribution of the OLS estimators approaches the normal distribution. This is called **asymptotic normality**.
- It is important to note that the distribution of the error in the population has nothing to do with the sample size. A skewed error stays skewed.
- What changes with the sample size is the **sampling distribution of** $\hat{\beta}_j$, which approaches the normal distribution.

Central Limit Theorem to the Rescue

- When we apply the **central limit theorem**, we can conclude that as the sample size grows, the distribution of the OLS estimators approaches the normal distribution. This is called **asymptotic normality**.
- It is important to note that the distribution of the error in the population has nothing to do with the sample size. A skewed error stays skewed.
- What changes with the sample size is the **sampling distribution of** $\hat{\beta}_j$, which approaches the normal distribution.
- Therefore, we can drop M.L.R - VI for large sample inference as long as the error has a finite variance.

Theorem 5.2 - Asymptotic Normality of OLS

- Under the Gauss-Markov assumptions M.L.R - I through M.L.R - V:

Theorem 5.2 - Asymptotic Normality of OLS

- Under the Gauss-Markov assumptions M.L.R - I through M.L.R - V:
  1. $\hat{\beta}_j$ is asymptotically normally distributed.

Theorem 5.2 - Asymptotic Normality of OLS

- Under the Gauss-Markov assumptions M.L.R - I through M.L.R - V:
  1. $\hat{\beta}_j$ is asymptotically normally distributed.
  2. $\hat{\sigma}^2$ is a consistent estimator of $\sigma^2 = Var(\mu)$.

Theorem 5.2 - Asymptotic Normality of OLS

- Under the Gauss-Markov assumptions M.L.R - I through M.L.R - V:
  1. $\hat{\beta}_j$ is asymptotically normally distributed.
  2. $\hat{\sigma}^2$ is a consistent estimator of $\sigma^2 = Var(\mu)$.
  3. For each j,

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} Normal(0, 1)$$

Theorem 5.2 - Asymptotic Normality of OLS

- Under the Gauss-Markov assumptions M.L.R - I through M.L.R - V:
  1. $\hat{\beta}_j$ is asymptotically normally distributed.
  2. $\hat{\sigma}^2$ is a consistent estimator of $\sigma^2 = Var(\mu)$.
  3. For each j,

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} Normal(0, 1)$$

- The symbol $\overset{a}{\sim}$ reads "is approximately distributed as, in large samples".

Theorem 5.2 - Asymptotic Normality of OLS

- Under the Gauss-Markov assumptions M.L.R - I through M.L.R - V:
  1. $\hat{\beta}_j$ is asymptotically normally distributed.
  2. $\hat{\sigma}^2$ is a consistent estimator of $\sigma^2 = Var(\mu)$.
  3. For each j,

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} Normal(0, 1)$$

- The symbol $\overset{a}{\sim}$ reads "is approximately distributed as, in large samples".
- We still need homoskedasticity (M.L.R - V) here. If it fails, the usual standard errors are not valid even in large samples.

Asymptotic t-test

- As the degrees of freedom get large, the t_{n-k-1} distribution approaches the standard normal. So we can also write

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} t_{n-k-1}$$

Asymptotic t-test

- As the degrees of freedom get large, the t_{n-k-1} distribution approaches the standard normal. So we can also write

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} t_{n-k-1}$$

- Therefore, we can conduct the t-test and build confidence intervals exactly as in Chapter 4, as long as we have a large enough sample.

Asymptotic t-test

- As the degrees of freedom get large, the t_{n-k-1} distribution approaches the standard normal. So we can also write

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} t_{n-k-1}$$

- Therefore, we can conduct the t-test and build confidence intervals exactly as in Chapter 4, as long as we have a large enough sample.
- "Large enough" is not precisely defined. The larger the sample size, the better the approximation.

Asymptotic t-test

- As the degrees of freedom get large, the t_{n-k-1} distribution approaches the standard normal. So we can also write

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} t_{n-k-1}$$

- Therefore, we can conduct the t-test and build confidence intervals exactly as in Chapter 4, as long as we have a large enough sample.
- "Large enough" is not precisely defined. The larger the sample size, the better the approximation.
- Only the terminologies change if μ is not normally distributed.

Asymptotic t-test

- As the degrees of freedom get large, the t_{n-k-1} distribution approaches the standard normal. So we can also write

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} t_{n-k-1}$$

- Therefore, we can conduct the t-test and build confidence intervals exactly as in Chapter 4, as long as we have a large enough sample.
- "Large enough" is not precisely defined. The larger the sample size, the better the approximation.
- Only the terminologies change if μ is not normally distributed.
  - The standard error becomes the **asymptotic standard error**.

Asymptotic t-test

- As the degrees of freedom get large, the t_{n-k-1} distribution approaches the standard normal. So we can also write

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} t_{n-k-1}$$

- Therefore, we can conduct the t-test and build confidence intervals exactly as in Chapter 4, as long as we have a large enough sample.
- "Large enough" is not precisely defined. The larger the sample size, the better the approximation.
- Only the terminologies change if μ is not normally distributed.
  - The standard error becomes the **asymptotic standard error**.
  - The t statistic becomes the **asymptotic t statistic**.

Standard Errors and the Sample Size

- Recall the estimated variance of the OLS estimator.

$$\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1 - R_j^2)}$$

Standard Errors and the Sample Size

- Recall the estimated variance of the OLS estimator.

$$\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1 - R_j^2)}$$

- As n grows, $\hat{\sigma}^2$ settles at σ^2 and R_j^2 settles at a number between 0 and 1.

Standard Errors and the Sample Size

- Recall the estimated variance of the OLS estimator.

$$\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1 - R_j^2)}$$

- As n grows, $\hat{\sigma}^2$ settles at σ^2 and R_j^2 settles at a number between 0 and 1.
- But SST_j grows roughly in proportion to n. Therefore

$$se(\hat{\beta}_j) \approx \frac{c_j}{\sqrt{n}}$$

where c_j is a positive constant that does not depend on the sample size.

Standard Errors and the Sample Size

- Recall the estimated variance of the OLS estimator.

$$\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1 - R_j^2)}$$

- As n grows, $\hat{\sigma}^2$ settles at σ^2 and R_j^2 settles at a number between 0 and 1.
- But SST_j grows roughly in proportion to n. Therefore

$$se(\hat{\beta}_j) \approx \frac{c_j}{\sqrt{n}}$$

  where c_j is a positive constant that does not depend on the sample size.
- To cut the standard error in half, we need four times as much data.

Standard Errors and the Sample Size

- Recall the estimated variance of the OLS estimator.

$$\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1 - R_j^2)}$$

- As n grows, $\hat{\sigma}^2$ settles at σ^2 and R_j^2 settles at a number between 0 and 1.
- But SST_j grows roughly in proportion to n. Therefore

$$se(\hat{\beta}_j) \approx \frac{c_j}{\sqrt{n}}$$

  where c_j is a positive constant that does not depend on the sample size.
- To cut the standard error in half, we need four times as much data.
- Go over Example 5.2 in the book (Standard Errors in a Birth Weight Equation).

The Lagrange Multiplier Statistic

The Lagrange Multiplier Statistic

- Similar to the asymptotic t-test, we can also do the asymptotic F-test. Nothing changes except terminology.

The Lagrange Multiplier Statistic

- Similar to the asymptotic t-test, we can also do the asymptotic F-test. Nothing changes except terminology.
- But in large samples, we have another test for exclusion restrictions - the **LM Test** (Lagrange Multiplier Test), also called the n-R-squared statistic or the score statistic.

The Lagrange Multiplier Statistic

- Similar to the asymptotic t-test, we can also do the asymptotic F-test. Nothing changes except terminology.
- But in large samples, we have another test for exclusion restrictions - the **LM Test** (Lagrange Multiplier Test), also called the n-R-squared statistic or the score statistic.
- Consider the model with k explanatory variables and test whether the last q of them can be excluded.

$$H_0 : \beta_{k-q+1} = 0, ..., \beta_k = 0$$

$$H_A : \text{At least one of them is different from zero}$$

The Lagrange Multiplier Statistic

- Similar to the asymptotic t-test, we can also do the asymptotic F-test. Nothing changes except terminology.
- But in large samples, we have another test for exclusion restrictions - the **LM Test** (Lagrange Multiplier Test), also called the n-R-squared statistic or the score statistic.
- Consider the model with k explanatory variables and test whether the last q of them can be excluded.

$$H_0 : \beta_{k-q+1} = 0, ..., \beta_k = 0$$

$$H_A : \text{At least one of them is different from zero}$$

- We only need to estimate the **restricted model** to conduct this test.

The Idea Behind the LM Test

- Run the restricted model and save the residuals $\tilde{\mu}$.

The Idea Behind the LM Test

- Run the restricted model and save the residuals $\tilde{\mu}$.
- If the excluded variables $x_{k-q+1}, ..., x_k$ truly have zero coefficients, then $\tilde{\mu}$ should be uncorrelated with them.

The Idea Behind the LM Test

- Run the restricted model and save the residuals $\tilde{\mu}$.
- If the excluded variables $x_{k-q+1}, ..., x_k$ truly have zero coefficients, then $\tilde{\mu}$ should be uncorrelated with them.
- So we regress $\tilde{\mu}$ on **all** the explanatory variables and look at the R-squared.

The Idea Behind the LM Test

- Run the restricted model and save the residuals $\tilde{\mu}$.
- If the excluded variables $x_{k-q+1}, ..., x_k$ truly have zero coefficients, then $\tilde{\mu}$ should be uncorrelated with them.
- So we regress $\tilde{\mu}$ on **all** the explanatory variables and look at the R-squared.
- If H_0 is true, this R-squared should be close to zero.

The Idea Behind the LM Test

- Run the restricted model and save the residuals $\tilde{\mu}$.
- If the excluded variables $x_{k-q+1}, ..., x_k$ truly have zero coefficients, then $\tilde{\mu}$ should be uncorrelated with them.
- So we regress $\tilde{\mu}$ on **all** the explanatory variables and look at the R-squared.
- If H_0 is true, this R-squared should be close to zero.
- The question is how large it must be before we reject. That is what the LM statistic answers.

The LM Test

1. Run the restricted model and save the residuals $\tilde{\mu}$.

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_{k-q} x_{k-q} + \mu$$

The LM Test

1. Run the restricted model and save the residuals $\tilde{\mu}$.

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_{k-q} x_{k-q} + \mu$$

2. Regress the residuals on **all** the independent variables (**Auxiliary Regression**) and get the R-squared, R^2_{μ}.

$$\tilde{\mu} = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + error$$

The LM Test

1. Run the restricted model and save the residuals $\tilde{\mu}$.

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_{k-q} x_{k-q} + \mu$$

2. Regress the residuals on **all** the independent variables (**Auxiliary Regression**) and get the R-squared, R^2_μ.

$$\tilde{\mu} = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + error$$

3. Compute the LM statistic.

$$LM = nR^2_\mu$$

The LM Test

1. Run the restricted model and save the residuals $\tilde{\mu}$.

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_{k-q} x_{k-q} + \mu$$

2. Regress the residuals on **all** the independent variables (**Auxiliary Regression**) and get the R-squared, R^2_μ.

$$\tilde{\mu} = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + error$$

3. Compute the LM statistic.

$$LM = nR^2_\mu$$

4. Under H_0, $LM \overset{a}{\sim} \chi^2_q$. Get the critical value c from the chi-square table with q degrees of freedom. Reject the null hypothesis if $LM > c$.

Notes on the LM Test

- The degrees of freedom are q, the number of restrictions being tested. The degrees of freedom of the regressions play no role.

Notes on the LM Test

- The degrees of freedom are q, the number of restrictions being tested. The degrees of freedom of the regressions play no role.
- You must include **all** k explanatory variables in the auxiliary regression. Leaving out the ones from the restricted model is a common mistake.

Notes on the LM Test

- The degrees of freedom are q, the number of restrictions being tested. The degrees of freedom of the regressions play no role.
- You must include **all** k explanatory variables in the auxiliary regression. Leaving out the ones from the restricted model is a common mistake.
- The auxiliary regression is only a device to compute the statistic. We do not interpret its coefficients.

Notes on the LM Test

- The degrees of freedom are q, the number of restrictions being tested. The degrees of freedom of the regressions play no role.
- You must include **all** k explanatory variables in the auxiliary regression. Leaving out the ones from the restricted model is a common mistake.
- The auxiliary regression is only a device to compute the statistic. We do not interpret its coefficients.
- In large samples, the LM test and the F-test rarely lead to different conclusions.

Notes on the LM Test

- The degrees of freedom are q, the number of restrictions being tested. The degrees of freedom of the regressions play no role.
- You must include **all** k explanatory variables in the auxiliary regression. Leaving out the ones from the restricted model is a common mistake.
- The auxiliary regression is only a device to compute the statistic. We do not interpret its coefficients.
- In large samples, the LM test and the F-test rarely lead to different conclusions.
- Like the F-test, the usual LM test requires homoskedasticity (M.L.R - V).

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$
$$+ \beta_4 ptime86 + \beta_5 qemp86 + \mu$$

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$
$$+ \beta_4 ptime86 + \beta_5 qemp86 + \mu$$

- We want to test $H_0 : \beta_2 = 0, \beta_3 = 0$, so $q = 2$.

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$

$$+\beta_4 ptime86 + \beta_5 qemp86 + \mu$$

- We want to test $H_0 : \beta_2 = 0, \beta_3 = 0$, so $q = 2$.
  1. Regress *narr*86 on *pcnv*, *ptime*86 and *qemp*86 and save the residuals $\tilde{\mu}$.

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$

$$+\beta_4 ptime86 + \beta_5 qemp86 + \mu$$

- We want to test $H_0 : \beta_2 = 0, \beta_3 = 0$, so $q = 2$.
  1. Regress *narr*86 on *pcnv*, *ptime*86 and *qemp*86 and save the residuals $\tilde{\mu}$.
  2. Regress $\tilde{\mu}$ on *pcnv*, *ptime*86, *qemp*86, *avgsen* and *tottime*. We get $R^2_\mu = 0.0015$.

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$

$$+\beta_4 ptime86 + \beta_5 qemp86 + \mu$$

- We want to test $H_0 : \beta_2 = 0, \beta_3 = 0$, so $q = 2$.
  1. Regress *narr*86 on *pcnv*, *ptime*86 and *qemp*86 and save the residuals $\tilde{\mu}$.
  2. Regress $\tilde{\mu}$ on *pcnv*, *ptime*86, *qemp*86, *avgsen* and *tottime*. We get $R^2_{\mu} = 0.0015$.
  3. $LM = 2,725 \times 0.0015 \approx 4.09$

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$

$$+\beta_4 ptime86 + \beta_5 qemp86 + \mu$$

- We want to test $H_0 : \beta_2 = 0, \beta_3 = 0$, so $q = 2$.
  1. Regress *narr*86 on *pcnv*, *ptime*86 and *qemp*86 and save the residuals $\tilde{\mu}$.
  2. Regress $\tilde{\mu}$ on *pcnv*, *ptime*86, *qemp*86, *avgsen* and *tottime*. We get $R^2_\mu = 0.0015$.
  3. $LM = 2,725 \times 0.0015 \approx 4.09$
  4. The 10% critical value of χ^2_2 is 4.61. Since $4.09 < 4.61$, we fail to reject H_0 at the 10% level.

Example 5.3 - Economic Model of Crime

- Consider the following model from Wooldridge, with $n = 2,725$.

$$narr86 = \beta_0 + \beta_1 pcnv + \beta_2 avgsen + \beta_3 tottime$$

$$+\beta_4 ptime86 + \beta_5 qemp86 + \mu$$

- We want to test $H_0 : \beta_2 = 0, \beta_3 = 0$, so $q = 2$.
  1. Regress *narr*86 on *pcnv*, *ptime*86 and *qemp*86 and save the residuals $\tilde{\mu}$.
  2. Regress $\tilde{\mu}$ on *pcnv*, *ptime*86, *qemp*86, *avgsen* and *tottime*. We get $R^2_\mu = 0.0015$.
  3. $LM = 2,725 \times 0.0015 \approx 4.09$
  4. The 10% critical value of χ^2_2 is 4.61. Since $4.09 < 4.61$, we fail to reject H_0 at the 10% level.
- A tiny R-squared can still give a sizeable LM statistic because it is multiplied by n.

Asymptotic Efficiency

Asymptotic Efficiency of OLS

- Recall the Gauss-Markov theorem. Under M.L.R - I through M.L.R - V, OLS is BLUE.

Asymptotic Efficiency of OLS

- Recall the Gauss-Markov theorem. Under M.L.R - I through M.L.R - V, OLS is BLUE.
- There is a large sample version of this result.

Asymptotic Efficiency of OLS

- Recall the Gauss-Markov theorem. Under M.L.R - I through M.L.R - V, OLS is BLUE.
- There is a large sample version of this result.
- Under M.L.R - IV, μ is uncorrelated with any function of the explanatory variables. So there are many other consistent estimators of β_j besides OLS.

Asymptotic Efficiency of OLS

- Recall the Gauss-Markov theorem. Under M.L.R - I through M.L.R - V, OLS is BLUE.
- There is a large sample version of this result.
- Under M.L.R - IV, μ is uncorrelated with any function of the explanatory variables. So there are many other consistent estimators of β_j besides OLS.
- The natural question is whether any of them is better than OLS in large samples.

Theorem 5.3 - Asymptotic Efficiency of OLS

- **Theorem:**

  Under the Gauss-Markov assumptions, let $\hat{\beta}_j$ denote the OLS estimators and $\tilde{\beta}_j$ any other estimator in this class. Then for $j = 0, 1, 2, ..., k$, the OLS estimators have the smallest asymptotic variance.

$$Avar\ \sqrt{n}(\hat{\beta}_j - \beta_j) \leq Avar\ \sqrt{n}(\tilde{\beta}_j - \beta_j)$$

Theorem 5.3 - Asymptotic Efficiency of OLS

- **Theorem:**

  Under the Gauss-Markov assumptions, let $\hat{\beta}_j$ denote the OLS estimators and $\tilde{\beta}_j$ any other estimator in this class. Then for $j = 0, 1, 2, ..., k$, the OLS estimators have the smallest asymptotic variance.

$$Avar\ \sqrt{n}(\hat{\beta}_j - \beta_j) \leq Avar\ \sqrt{n}(\tilde{\beta}_j - \beta_j)$$

- We say that OLS is **asymptotically efficient**.

Theorem 5.3 - Asymptotic Efficiency of OLS

- **Theorem:**

  Under the Gauss-Markov assumptions, let $\hat{\beta}_j$ denote the OLS estimators and $\tilde{\beta}_j$ any other estimator in this class. Then for $j = 0, 1, 2, ..., k$, the OLS estimators have the smallest asymptotic variance.

$$Avar\ \sqrt{n}(\hat{\beta}_j - \beta_j) \leq Avar\ \sqrt{n}(\tilde{\beta}_j - \beta_j)$$

- We say that OLS is **asymptotically efficient**.
- All the other consistent estimators in this class have an asymptotic variance at least as large.

Theorem 5.3 - Asymptotic Efficiency of OLS

- **Theorem:**

  Under the Gauss-Markov assumptions, let $\hat{\beta}_j$ denote the OLS estimators and $\tilde{\beta}_j$ any other estimator in this class. Then for $j = 0, 1, 2, ..., k$, the OLS estimators have the smallest asymptotic variance.

$$Avar\ \sqrt{n}(\hat{\beta}_j - \beta_j) \leq Avar\ \sqrt{n}(\tilde{\beta}_j - \beta_j)$$

- We say that OLS is **asymptotically efficient**.
- All the other consistent estimators in this class have an asymptotic variance at least as large.
- Homoskedasticity is essential for this result. If M.L.R - V fails, there are estimators that beat OLS (Chapter 8).

Summary

- **Consistency:** Under M.L.R - I through M.L.R - IV (or with the weaker M.L.R - IV′), plim $\hat{\beta}_j = \beta_j$.

Summary

- **Consistency:** Under M.L.R - I through M.L.R - IV (or with the weaker M.L.R - IV′), plim $\hat{\beta}_j = \beta_j$.
- **Inconsistency:** If $Cov(x_j, \mu) \neq 0$, OLS is inconsistent and more data does not help.

Summary

- **Consistency:** Under M.L.R - I through M.L.R - IV (or with the weaker M.L.R - IV′), plim $\hat{\beta}_j = \beta_j$.
- **Inconsistency:** If $Cov(x_j, \mu) \neq 0$, OLS is inconsistent and more data does not help.
- **Asymptotic normality:** Under M.L.R - I through M.L.R - V, we can use t, F and LM tests in large samples without M.L.R - VI.

Summary

- **Consistency:** Under M.L.R - I through M.L.R - IV (or with the weaker M.L.R - IV′), plim $\hat{\beta}_j = \beta_j$.
- **Inconsistency:** If $Cov(x_j, \mu) \neq 0$, OLS is inconsistent and more data does not help.
- **Asymptotic normality:** Under M.L.R - I through M.L.R - V, we can use t, F and LM tests in large samples without M.L.R - VI.
- **LM statistic:** $LM = nR_{\mu}^2 \overset{a}{\sim} \chi_q^2$

Summary

- **Consistency:** Under M.L.R - I through M.L.R - IV (or with the weaker M.L.R - IV′), plim $\hat{\beta}_j = \beta_j$.
- **Inconsistency:** If $Cov(x_j, \mu) \neq 0$, OLS is inconsistent and more data does not help.
- **Asymptotic normality:** Under M.L.R - I through M.L.R - V, we can use t, F and LM tests in large samples without M.L.R - VI.
- **LM statistic:** $LM = nR_{\mu}^2 \overset{a}{\sim} \chi_q^2$
- **Asymptotic efficiency:** Under the Gauss-Markov assumptions, OLS has the smallest asymptotic variance.

Final Paper & Presentation

Final Paper

- Look at the datasets from following sources:

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets
  3. Opendata.com.pk

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets
  3. Opendata.com.pk
- Based on the dataset come up with the research question.

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets
  3. Opendata.com.pk
- Based on the dataset come up with the research question.
- Literature Review

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets
  3. Opendata.com.pk
- Based on the dataset come up with the research question.
- Literature Review
- Finalize the Model

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets
  3. Opendata.com.pk
- Based on the dataset come up with the research question.
- Literature Review
- Finalize the Model
- Run the Analysis

Final Paper

- Look at the datasets from following sources:
  1. STATA Data sets
  2. Wooldridge Data Sets
  3. Opendata.com.pk
- Based on the dataset come up with the research question.
- Literature Review
- Finalize the Model
- Run the Analysis
- Present your findings