

ASYMPTOTIC PROPERTIES of OLS ESTIMATORS

Hüseyin Taştan¹

¹Yıldız Technical University
Department of Economics

These presentation notes are based on
Introductory Econometrics: A Modern Approach (2nd ed.)
by J. Wooldridge.

31 Ekim 2012

Finite Sample Properties of OLS Estimators

- ▶ Finite sample properties: unbiasedness, efficiency
- ▶ These properties are valid for any sample size n .
- ▶ OLS estimators are
 - ▶ Unbiased (under assumptions MLR.1-MLR.4)
 - ▶ BLUE (under assumptions MLR.1-MLR.5) (Best = the most efficient)
- ▶ Assumption MLR.6: Normality of the error term, u , independence from explanatory variables.
- ▶ Normality assumption is used to derive the sampling distributions of OLS estimators for any n .
- ▶ Under normality usual t and F test statistics follow standard distributions for any sample size.

Asymptotic Properties of OLS Estimators

- ▶ What are the asymptotic properties of OLS estimators?
- ▶ Asymptotics: "as the sample size, n , increases without limit"
- ▶ These properties are: consistency and asymptotic normality
- ▶ We will be able to make weaker MLR assumptions.

Asymptotic Properties: Consistency

Definition

Let W_n be an estimator for the unknown population parameter θ based on a random sample of $\{Y_1, Y_2, \dots, Y_n\}$. For an arbitrary small number $\epsilon > 0$, as n becomes larger and larger if the following condition is satisfied

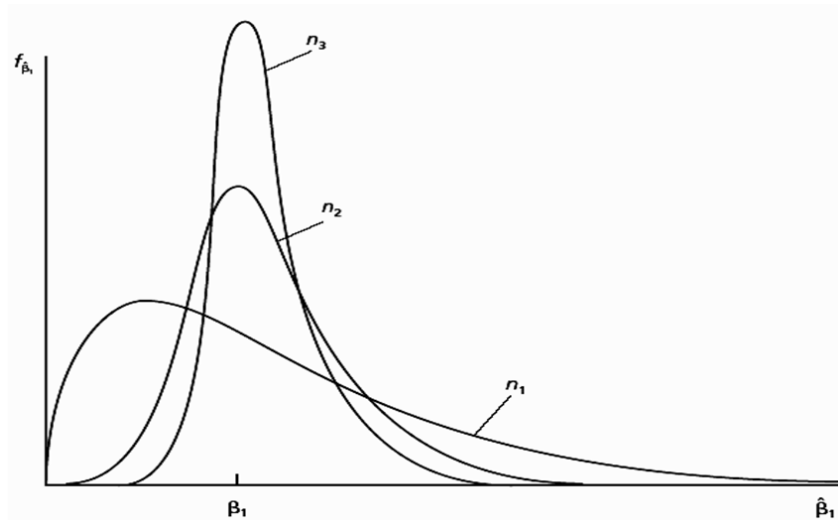
$$P(|W_n - \theta| > \epsilon) \rightarrow 0, \quad n \rightarrow \infty$$

then, W_n is a consistent estimator for θ : $\text{plim}(W_n) = \theta$.

- ▶ Example: arithmetic mean \bar{x} is a consistent estimator for the population mean μ
- ▶ Law of Large Numbers.

Sampling Distributions for Different Sample Sizes:

$$n_1 < n_2 < n_3$$



6

Asymptotic Properties: Consistency

- ▶ For each sample size n , $\hat{\beta}_j$ has a sampling distribution. In other words, sampling distributions of OLS estimators are indexed by the sample size.
- ▶ If $\hat{\beta}_j$ is consistent, as n gets larger and larger the sampling distributions become more concentrated around the true value β_j .
- ▶ When $n \rightarrow \infty$, i.e. obtaining more and more data gets us closer to the parameter of interest, β_j . In the limit, the sampling distribution collapses on a single point.
- ▶ This means that if we can collect more and more data we can make our estimator arbitrarily close to the true value.

7

Asymptotic Properties: Consistency

- ▶ If, as n goes to infinity, the sampling distribution does not get closer to the true value β_j then the estimator is said to be **inconsistent**.
- ▶ The same set of assumptions for unbiasedness also implies the consistency of OLS estimators (MLR.1-MLR.4).
- ▶ But for consistency alone, the assumption MLR.3 can be replaced by a weaker assumption: '**error term is uncorrelated with each explanatory variable**'.

8

Asymptotic Properties: Consistency

Assumption MLR.3': Zero Mean and Zero Correlation

$$E(u) = 0$$

$$\text{Cov}(x_j, u) = 0, \quad j = 1, 2, \dots, k$$

- ▶ Replacing the MLR.3 with MLR.3' it can be shown that OLS estimators are consistent: $\text{plim}(\hat{\beta}_j) = \beta_j$.
- ▶ This assumption says that the error term is uncorrelated with each x_j .
- ▶ The assumption MLR.3' is weaker than the assumption MLR.3: the MLR.3 implies the MLR.3' but the reverse may not be true. The MLR.3' is not enough to show unbiasedness.

Asymptotic Properties: Consistency

- ▶ In the simple regression model the OLS estimator of the slope parameter can be written as:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_1) y_i}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}$$

- ▶ Substituting $y = \beta_0 + \beta_1 x_1 + u$ and rearranging:

$$\hat{\beta}_1 = \beta_1 + \frac{n^{-1} \sum_{i=1}^n (x_{i1} - \bar{x}_1) u_i}{n^{-1} \sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}$$

- ▶ Taking plim:

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

- ▶ By the assumption MLR.3': $\text{Cov}(x_1, u) = 0$, we obtain

$$\text{plim}(\hat{\beta}_1) = \beta_1$$

Deriving Asymptotic Bias

- ▶ Asymptotic bias:

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

- ▶ The sign of the asymptotic bias depends on the sign of the covariance between x_1 and u . Notice that it is not possible to estimate $\text{Cov}(x_1, u)$ since we cannot observe u .
- ▶ If we omit an important variable from the model then OLS estimators will be inconsistent.
- ▶ For example, suppose that the true population model is given below and the first four MLR assumptions are satisfied:

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

- ▶ If we omit x_2 then the error term will be $u = \beta_2 x_2 + \nu$.

Deriving Asymptotic Bias

- ▶ In this case the probability limit will be

$$\text{plim}(\hat{\beta}_1) = \beta_1 + \beta_2 \delta_1$$

$$\delta_1 = \frac{\text{Cov}(x_1, x_2)}{\text{Var}(x_1)}$$

- ▶ This is derived using:

$$\hat{\beta}_1 = \beta_1 + \beta_2 \left(\frac{n^{-1} \sum (x_{i1} - \bar{x}_1) x_{i2}}{n^{-1} \sum (x_{i1} - \bar{x}_1)^2} \right)$$

$$\text{plim} \hat{\beta}_1 = \beta_1 + \beta_2 \underbrace{\frac{\text{plim } n^{-1} \sum (x_{i1} - \bar{x}_1) x_{i2}}{\text{plim } n^{-1} \sum (x_{i1} - \bar{x}_1)^2}}_{=\delta_1}$$

Consistency of OLS Estimators

- ▶ Increasing the sample size will not cure inconsistency. In fact, as n gets larger and larger the OLS estimator will get closer to $\beta_1 + \beta_2 \delta_1$ instead of β_1 .
- ▶ If only one of the x variables (say, x_1) is correlated with u then this will cause all OLS estimators to be inconsistent (in general).
- ▶ If there are x variables that are uncorrelated with x_1 , then their OLS estimators will be consistent.

Inconsistency: Example

- ▶ y : house price
- ▶ x_1 : distance from a trash incinerator
- ▶ x_2 : quality index of a house (reflects the overall effects of the factors such as size, number of rooms, the attractiveness of neighborhood, etc.)
- ▶ Suppose that x_2 is omitted from the model.
- ▶ $\beta_1 > 0$, $\beta_2 > 0$
- ▶ If on average, quality houses are built far from the trash incinerator then: $\delta_1 > 0$
- ▶ This implies that

$$\beta_1 + \beta_2 \delta_1 > \beta_1$$

- ▶ The OLS estimator of β_1 , $\hat{\beta}_1$ will get closer to $\beta_1 + \beta_2 \delta_1$ in the limit.

Asymptotic Normality and Large Sample Inference

- ▶ The assumption MLR.6: conditional x s the error term has a normal distribution. This implies that the conditional distribution of y is also normal.
- ▶ We do not need the normality assumption for the unbiasedness of OLS estimators. The normality assumption is required to derive the exact (finite sample) sampling distributions of OLS estimators (which are also normal).
- ▶ If the normality assumption fails, does this imply that we cannot carry out t and F tests?
- ▶ The answer is NO! If the sample size is large enough, we can use the Central Limit Theorem to conclude that OLS estimators are **asymptotically normal**.

Asymptotic Normality: Theorem 5.2

- ▶ Under the Gauss-Markov assumptions:

$$\sqrt{n}(\hat{\beta}_j - \beta_j) \sim^a \text{Normal} \left(0, \frac{\sigma^2}{a_j^2} \right)$$

- ▶ Asymptotic variance is

$$\frac{\sigma^2}{a_j^2} > 0$$

$$a_j^2 = \text{plim} \left(n^{-1} \sum_{i=1}^n \hat{r}_{ij}^2 \right)$$

- ▶ Where \hat{r}_{ij}^2 denotes the residuals from the regression x_j on all other x s.
- ▶ Standardizing we obtain:

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

Asymptotic Normality

- ▶ As n gets larger, the t statistic converges to the standard normal distribution (asymptotic t statistic):

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

- ▶ If Theorem 5.2 is valid then we do not need the assumption MLR.6 Normality.
- ▶ The only requirement is that the error variance is finite: $\sigma^2 > 0$.
- ▶ Theorem 5.2: constant variance and zero conditional mean assumptions are required.
- ▶ The standard errors of the OLS estimators ($se(\hat{\beta}_j)$) shrink to zero at the rate $1/\sqrt{n}$.

Asymptotic Efficiency

- ▶ We know that OLS estimators are BLUE under the Gauss-Markov assumptions.
- ▶ OLS estimators are also asymptotically efficient under the Gauss-Markov assumptions.

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

$Avar$ denotes asymptotic variance. $\tilde{\beta}_j$ denotes another estimator (other than OLS).

A Large Sample Test: The Lagrange Multiplier (LM) Test Statistic

- ▶ In large samples, we can use the Lagrange Multiplier (LM, or score test) statistic to test linear restrictions.
- ▶ The LM statistic relies only on the estimation of the restricted model. After the restricted model is estimated an auxiliary regression is run to get the LM statistic.
- ▶ Example: Exclusion restrictions - let the unrestricted model be:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + u$$

- ▶ $H_0 : \beta_3 = \beta_4 = 0$, H_1 : at least one of them is not zero.
- ▶ The LM test statistic is computed by multiplying the sample size n by R^2 which is obtained from the regression of the residuals from the restricted model on all explanatory variables.

LM Test

- ▶ Under $H_0 : \beta_3 = \beta_4 = 0$ the unrestricted model is

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

- ▶ Let \tilde{u} be the residuals from this regression. Solve the following auxiliary regression by running the regression of these residuals on all x s:

$$\tilde{u} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \text{error}$$

- ▶ Let $R_{\tilde{u}}^2$ be the coefficient of determination of this regression. Then the LM test statistic is:

$$LM = nR_{\tilde{u}}^2 \sim \chi_q^2$$

- ▶ Under the null hypothesis the LM statistic follows a chi-squared distribution with q degrees of freedom.
- ▶ Decision Rule: If $LM > c$ then reject H_0 .

LM Test: Example

Newborn weight and parents' education: `bwght.gdt`

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + u$$

- ▶ Dependent variable: y = weights of newly born babies (ounces)
- ▶ Explanatory variables:
 - ▶ x_1 : average number of cigarettes smoked per day during pregnancy
 - ▶ x_2 : parity (birth order)
 - ▶ x_3 : family income
 - ▶ x_4 : mother's education level, year
 - ▶ x_5 : father's education level, year.
- ▶ We are interested in: $H_0 : \beta_4 = 0, \beta_5 = 0$, parents' education level does not have any effect on weights of newborns.

LM Test: Example

Step 1: Unrestricted Model

$$\widehat{\text{bwght}} = \underset{(1.656)}{115.470} - \underset{(0.1088)}{0.5979} \text{cigs} + \underset{(0.6575)}{1.8323} \text{parity} + \underset{(0.0324)}{0.0671} \text{faminc}$$

$$n = 1191 \quad R^2 = 0.036 \quad F(3, 1187) = 14.953 \quad \hat{\sigma} = 19.796$$

Save the residuals from this regression: **uhat**

Step 2: Run the regression of uhat on all xs

$$\widehat{\text{uhat}} = \underset{(3.729)}{-0.9456} + \underset{(0.110)}{0.0019} \text{cigs} - \underset{(0.659)}{0.0447} \text{parity} - \underset{(0.0366)}{0.011} \text{faminc}$$

$$- \underset{(0.319)}{0.370} \text{motheduc} + \underset{(0.283)}{0.472} \text{fatheduc}$$

$$n = 1191 \quad R^2 = \mathbf{0.00242} \quad F(5, 1185) = 0.57491 \quad \hat{\sigma} = 19.789$$

LM Test: Example

Step 3: calculate the LM test statistic

$$LM = nR_u^2 \sim \chi_q^2$$

$$LM = (1191)(0.00242) \sim \chi_2^2$$

$$LM = 2.88$$

5% critical value at 2 degrees of freedom is $c = 5.99$. Thus, we fail to reject H_0 . Also note that $p\text{-value} = 0.24$