Introductory Econometrics: A modern approach (Wooldridge) Chapter 2

The Simple Regression Model

Dr. Lê Văn Chơn – FTU, 2011

2.1 Definition of Simple Regression Model

2

- Applied econometric analysis often begins with 2 variables y and x. We are interested in "studying how y varies with changes in x".
  E.g., x is years of education, y is hourly wage.
  x is number of police officers, y is a community crime rate.
- In the simple linear regression model:
- $y = \beta_0 + \beta_1 x + u$ (2.1)
  y is called the *dependent variable*, the *explained variable*, or the *regressand*.
  x is called the *independent variable*, the *explanatory variable*, or the *regressor*.
  u, called *error term* or *disturbance*, represents factors other than x that affect y. u stands for "unobserved".

Dr. Lê Văn Chơn – FTU, 2011

2.1 Definition of Simple Regression Model

3

- If the other factors in u are held fixed, $\Delta u = 0$, then x has a linear effect on y: $\Delta y = \beta_1 \Delta x$
- β_1 is the slope parameter. This is of primary interest in applied economics.
- One-unit change in x has the same effect on y, regardless of the initial value of x. → Unrealistic.
- E.g., wage-education example, we might want to allow for *increasing* returns.

Dr. Lê Văn Chơn – FTU, 2011

2.1 Definition of Simple Regression Model

4

- **An assumption**: the average value of u in the population is zero.

  E(u) = 0 (2.5)

  This assumption is not restrictive since we can always use β_0 to normalize E(u) to 0.

- Because u and x are random variables, we can define conditional distribution of u given any value of x.
- **Crucial assumption**: average value of u does not depend on x.

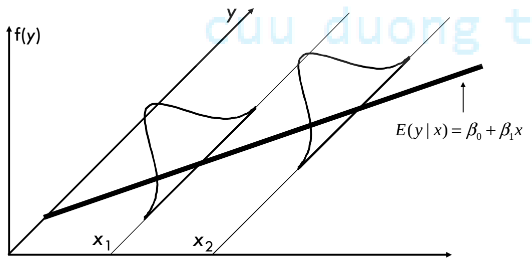
  E(u|x) = E(u) (2.6)

- (2.5) + (2.6) → the **zero conditional mean assumption.**
- This implies E(y|x) = $\beta_0 + \beta_1 x$

2.1 Definition of Simple Regression Model

5

- Population regression function (PRF): E(y|x) is a linear function of x. For any value of x, the distribution of y is centered about E(y|x).



2.2 Ordinary Least Squares

6

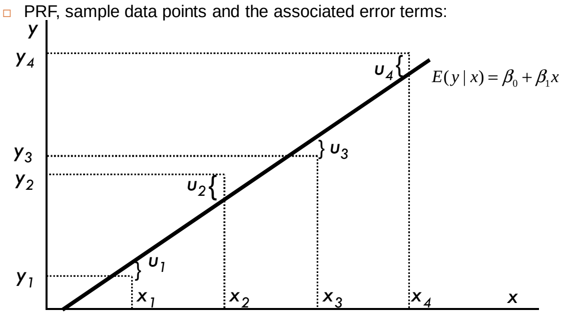
- How to estimate population parameters β_0 and β_1 from a sample?
- Let {(x_i,y_i): i = 1, 2, …, n} denote a random sample of size n from the population.
- For each observation in this sample, it will be the case that

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

2.2 Ordinary Least Squares

7

- PRF, sample data points and the associated error terms:



Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

8

- To derive the OLS estimates, we need to realize that our main assumption of E(u|x) = E(u) = 0 also implies that

  Cov(x,u) = E(xu) = 0 (2.11)

  Why? Cov(x,u) = E(xu) – E(x)E(u) = E_x[E(xu|x)] = E_x[xE(u|x)] = 0.

- We can write 2 restrictions (2.5) and (2.11) in terms of x, y, β_0 and β_1

$$E(y - \beta_0 - \beta_1 x) = 0 \quad (2.12)$$

$$E[x(y - \beta_0 - \beta_1 x)] = 0 \quad (2.13)$$

- (2.12) and (2.13) are 2 moment restrictions with 2 unknown parameters. → They can be used to obtain good estimators of β_0 and β_1.

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

9

- **Method of moments** approach to estimation implies imposing the population moment restrictions on the sample moments.

- Given a sample, we choose estimates β_0 and β_1 to solve the sample versions:

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad (2.14)$$

$$\frac{1}{n}\sum_{i=1}^{n}x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad (2.15)$$

- Given the properties of summation, (2.14) can be rewritten as

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} \quad (2.16)$$

or

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (2.17)$$

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

10

- Drop 1/n in (2.15) and plug (2.17) into (2.15):

$$\sum_{i=1}^{n} x_i (y_i - [\bar{y} - \hat{\beta}_1 \bar{x}] - \hat{\beta}_1 x_i) = 0$$

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

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

- Provided that $\sum_{i=1}^{n} (x_i - \bar{x})^2 > 0$ (2.18)

the estimated slope is $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$ (2.19)

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

11

- Summary of OLS slope estimate:
  - Slope estimate is the sample covariance between x and y divided by the sample variance of x.
  - If x and y are positively correlated, the slope will be positive.
  - If x and y are negatively correlated, the slope will be negative.
  -Only need x to vary in the sample.

- $\hat{\beta}_0$ and $\hat{\beta}_1$ given in (2.17) and (2.19) are called the **ordinary least squares (OLS)** estimates of β_0 and β_1.

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

12

- To justify this name, for any $\hat{\beta}_0$ and $\hat{\beta}_1$, define a fitted value for y given x = x_i: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ (2.20)

- The residual for observation i is the difference between the actual y_i and its fitted value: $\hat{u}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$

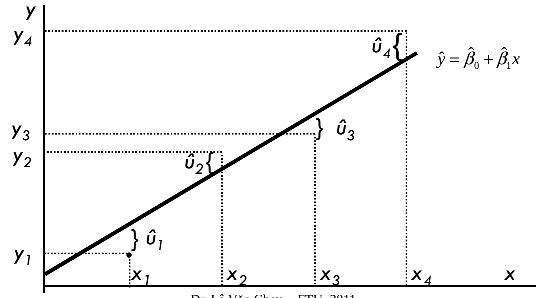
- Intuitively, OLS is fitting a line through the sample points such that the sum of squared residuals is as small as possible → term "ordinary least squares".
- Formal minimization problem:
- $\min_{\hat{\beta}_0, \hat{\beta}_1} \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$ (2.22)

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

13

- Sample regression line, sample data points and residuals:



Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

14

- To solve (2.22), we obtain 2 first order conditions, which are the same as (2.14) and (2.15), multiplied by n.
- Once we have determined the OLS $\hat{\beta}_0$ and $\hat{\beta}_1$, we have the **OLS regression line**: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ (2.23)
- (2.23) is also called the sample regression function (SRF) because it is the estimated version of the population regression function (PRF) $E(y\,|\,x) = \beta_0 + \beta_1 x$.
- Remember that PRF is fixed but unknown.
  Different samples generate different SRFs.

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

15

- Slope estimate $\hat{\beta}_1$ is of primary interest. It tells us the amount by which $\hat{y}$ changes when x increases by 1 unit.

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x$$

- E.g., we study the relationship between firm performance and CEO compensation.

  $salary = \beta_0 + \beta_1 roe + u$

  salary = CEO's annual salary in thousands of dollars,

  roe = average return (%) on the firm's equity for previous 3 years.
- Because a higher *roe* is good for the firm, we think β_1 > 0.
- Data set CEOSAL1 contains information on 209 CEOs in 1990.
- OLS regression line: $\widehat{salary}$ = 963.191 + 18.501*roe* (2.26)

Dr. Lê Văn Chơn – FTU, 2011

2.2 Ordinary Least Squares

16

- E.g., for the population of the workforce in 1976, let
  y = *wage*, $ per hour,
  x = *educ*, years of schooling.
- Using data in WAGE1 with 526 observations, we obtain the OLS regression line:
  $\hat{wage} = -0.90 + 0.54educ$ (2.27)
- Implication of the intercept? Why?
  Only 18 people in the sample have less than 8 years of education.
  → the regression line does poorly at very low levels.
- Implication of the slope?

Dr. Lê Văn Chơn – FTU, 2011

2.3 Mechanics of OLS

17

Fitted Values and Residuals

- Given $\hat{\beta}_0$ and $\hat{\beta}_1$, we can obtain the fitted value $\hat{y}_i$ for each observation. Each $\hat{y}_i$ is on the OLS regression line.
- OLS residual associated with observation i, $\hat{u}_i$, is the difference between y_i and its fitted value.
  If $\hat{u}_i$ is positive, the line underpredicts y_i.
  If $\hat{u}_i$ is negative, the line overpredicts y_i.
- In most cases, every $\hat{u}_i \neq 0$, none of the data points lie on the OLS line.

Dr. Lê Văn Chơn – FTU, 2011

2.3 Mechanics of OLS

18

Algebraic Properties of OLS Statistics

(1) The sum and thus the sample average of the OLS residuals is zero.

$$\sum_{i=1}^{n} \hat{u}_i = 0 \text{ and thus } \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i = 0$$

(2) The sample covariance between the regressors and the OLS residuals is zero.

$$\sum_{i=1}^{n} x_i \hat{u}_i = 0$$

(3) The OLS regression line always goes through the mean of the sample.

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

Dr. Lê Văn Chơn – FTU, 2011

2.3 Mechanics of OLS

19

- We can think of each observation i as being made up of an explained part and an unexplained part $y_i = \hat{y}_i + \hat{u}_i$.
- We define the following:
  $\sum_{i=1}^{n}(y_i - \bar{y})^2$ is the total sum of squares (SST),
  $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ is the explained sum of squares (SSE),
  $\sum_{i=1}^{n}\hat{u}_i^2$ is the residual sum of squares (SSR).
- Then SST = SSE + SSR (2.36)

Dr. Lê Văn Chơn – FTU, 2011

2.3 Mechanics of OLS

20

- Proof:

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}[(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})]^2 = \sum_{i=1}^{n}[\hat{u}_i + (\hat{y}_i - \bar{y})]^2$$
$$= \sum_{i=1}^{n}\hat{u}_i^2 + 2\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i - \bar{y}) + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$$
$$= SSR + 2\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i - \bar{y}) + SSE$$

and we know that $\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i - \bar{y}) = 0$

Dr. Lê Văn Chơn – FTU, 2011

2.3 Mechanics of OLS

21

Goodness-of-Fit

- How well the OLS regression line fits the data?
- Divide (2.36) by SST to get:
  1 = SSE/SST + SSR/SST
- The **R-squared** of the regression or the **coefficient of determination**

$$R^2 \equiv \frac{SSE}{SST} = 1 - \frac{SSR}{SST} \quad (2.38)$$

It implies the fraction of the sample variation in y that is explained by the model.

$$0 \le R^2 \le 1$$

Dr. Lê Văn Chơn – FTU, 2011

2.3 Mechanics of OLS

22

- E.g., CEOSAL1. roe explains only about 1.3% of the variation in salaries for this sample.
- → 98.7% of the salary variations for these CEOs is left unexplained!
- Notice that a seemingly low R^2 does not mean that an OLS regression equation is useless.
- It is still possible that (2.26) is a good estimate of the ceteris paribus relationship between *salary* and *roe*.

Dr. Lê Văn Chơn – FTU, 2011

2.4 Units of Measurement

23

- OLS estimates change when the units of measurement of the dependent and independent variables change.
- E.g., CEOSAL1. Rather than measuring salary in $'000, we measure it in $, *salardol* = 1,000.*salary*.

  Without regression, we know that

  $$\widehat{salardol} = 963{,}191 + 18{,}501 roe. \qquad (2.40)$$
- Multiply the intercept and the slope in (2.26) by 1,000 → (2.26) and (2.40) have the same interpretations.
- Define *roedec* = *roe*/100 where *roedec* is a decimal.

  $$\widehat{salary} = 963.191 + 1850.1 roedec. \qquad (2.41)$$

Dr. Lê Văn Chơn – FTU, 2011

2.4 Units of Measurement

24

- What happens to R^2 when units of measurement change?

  Nothing.

Dr. Lê Văn Chơn – FTU, 2011

2.4 Nonlinearities in Simple Regression

25

- It is rather easy to incorporate many nonlinearities into simple regression analysis by appropriately defining y and x.
- E.g., WAGE1. $\hat{\beta}_1$ of 0.54 means that each additional year of education increases wage by 54 cents. → maybe not reasonable.
- Suppose that the *percentage* increase in wage is the same given one more year of education.
  (2.27) does not imply a constant percentage increase.
- New model: log(*wage*) = $\beta_0 + \beta_1$ *educ* + *u* (2.42)
  where log(.) denotes the natural logarithm.

Dr. Lê Văn Chơn – FTU, 2011

2.4 Nonlinearities in Simple Regression

26

- For each additional year of education, the *percentage* change in *wage* is the same. → the change in *wage increases*.
- (2.42) implies an *increasing* return to education.
- Estimating this model and the mechanics of OLS are the same:
- $\widehat{\log}$(*wage*) = 0.584 + 0.083*educ* (2.44)
- *wage* increases by 8.3 percent for every additional year of *educ*.

Dr. Lê Văn Chơn – FTU, 2011

2.4 Nonlinearities in Simple Regression

27

- Another important use of the natural log is in obtaining a **constant elasticity model**.
- E.g., CEOSAL1. We can estimate a constant elasticity model relating CEO salary ($'000) to firm sales ($ mil):
  log(*salary*) = $\beta_0 + \beta_1$ log(*sales*) + *u* (2.45)
  where β_1 is the elasticity of *salary* with respect to *sales*.
- If we change the units of measurement of *y*, what happens to β_1?
  Nothing.

Dr. Lê Văn Chơn – FTU, 2011

2.4 Meaning of Linear Regression

28

- We have seen a model that allows for *nonlinear* relationships. So what does "linear" mean?
- An equation $y = \beta_0 + \beta_1 x + u$ is linear in parameters, β_0 and β_1.
  There are no restrictions on how *y* and *x* relate to the original dependent and independent variables.
- Plenty of models cannot be cast as linear regression models because they are not linear in their parameters.
  E.g., *cons* = 1/($\beta_0 + \beta_1$ inc) + *u*

Dr. Lê Văn Chơn – FTU, 2011

2.5 Unbiasedness of OLS

29

Unbiasedness of OLS is established under a set of assumptions:

- **Assumption SLR.1** (Linear in Parameters)
  The population model is linear in parameters as
  $$y = \beta_0 + \beta_1 x + u \quad (2.47)$$
  where β_0 and β_1 are the population intercept and slope parameters.
- Realistically, *y*, *x*, *u* are all viewed as random variables.
- **Assumption SLR.2** (Random Sampling)
  We can use a random sample of size *n*, {(x_i,y_i): *i* = 1, 2, …, *n*}, from the population model.

Dr. Lê Văn Chơn – FTU, 2011

2.5 Unbiasedness of OLS

30

- Not all cross-sectional samples can be viewed as random samples, but many may be.
- We can write (2.47) in terms of the random sample as
  $$y_i = \beta_0 + \beta_1 x_i + u_i \,, \quad i = 1, 2, \ldots, n \quad (2.48)$$
- To obtain unbiased estimators of β_0 and β_1, we need to impose
- **Assumption SLR.3** (Zero Conditional Mean)
  E(u|x) = 0

  This assumption implies $E(u_i|x_i) = 0$ for all i = 1, 2, …, n.

Dr. Lê Văn Chơn – FTU, 2011

2.5 Unbiasedness of OLS

31

- **Assumption SLR.4** (Sample Variation in the Independent Variable)
  In the sample, x_i, i = 1, 2, …, n are not all equal to a constant.

  This assumption is equivalent to $\sum_{i=1}^{n}(x_i-\bar{x})^2>0$
- From (2.19) :
$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})y_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$
- Plug (2.48) into this:
$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(\beta_0+\beta_1x_i+u_i)}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\beta_1+\frac{\sum_{i=1}^{n}(x_i-\bar{x})u_i}{SST_x}$$

Dr. Lê Văn Chơn – FTU, 2011

2.5 Unbiasedness of OLS

32

- Errors u_i's are generally different from 0. → $\hat{\beta}_1$ differs from β_1 .
- The first important statistical property of OLS:
  Theorem 2.1 (Unbiasedness of OLS)
  Using Assumptions SLR.1 through SLR.4,
  $$E(\hat{\beta}_0)=\beta_0\text{, and }E(\hat{\beta}_1)=\beta_1 \qquad (2.53)$$
  The OLS estimates of β_0 and β_1 are unbiased.
- Proof: $E(\hat{\beta}_1)=\beta_1+E[(1/SST_x)\sum_{i=1}^{n}(x_i-\bar{x})u_i]$
$$=\beta_1+(1/SST_x)\sum_{i=1}^{n}(x_i-\bar{x})E(u_i)=\beta_1$$

Dr. Lê Văn Chơn – FTU, 2011

2.5 Unbiasedness of OLS

33

- (2.17) implies
$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=\beta_0+\beta_1\bar{x}+\bar{u}-\hat{\beta}_1\bar{x}=\beta_0+(\beta_1-\hat{\beta}_1)\bar{x}+\bar{u}$$
$$E(\hat{\beta}_0)=\beta_0+E[(\beta_1-\hat{\beta}_1)\bar{x}]=\beta_0$$
- Remember unbiasedness is a feature of the sampling distributions of $\hat{\beta}_0$ and $\hat{\beta}_1$ .It says nothing about the estimate we obtain for a given sample.
- If any of four assumptions fails, then OLS is not necessarily unbiased.
- When u contains factors affecting y that are also correlated with x can result in *spurious correlation*.

Dr. Lê Văn Chơn – FTU, 2011

2.5 Unbiasedness of OLS

34

- E.g., let *math10* denote % of tenth graders at a high school receiving a passing score on a standardized math exam.
  Let *lnchprg* denote % of students eligible for the federally funded school lunch program.
- We expect the lunch program has a positive effect on performance:
  $$math10 = \beta_0 + \beta_1 \ln chprg + u$$
- MEAP93 has data on 408 Michigan high school for the 1992-1993 school year.
  $$\hat{math10} = 32.14 - 0.319 lnchprg$$
- Why? *u* contains such as the poverty rate of children attending school, which affects student performance and is highly correlated with eligibility in the lunch program.

Dr. Lê Văn Chơn – FTU, 2011

2.5 Variances of the OLS Estimators

35

- Now we know that the sampling distribution of our estimate is centered about the true parameter.
  How spread out is this distribution? → the variance.
- We need to add an assumption.
  Assumption SLR.5 (Homoskedasticity)
  $$\text{Var}(u|x) = \sigma^2$$
- This assumption is distinct from Assumption SLR.3: E(u|x) = 0.
- This assumption simplifies the variance calculations for $\hat{\beta}_0$ and $\hat{\beta}_1$ and it implies OLS has certain efficiency properties.

Dr. Lê Văn Chơn – FTU, 2011

2.5 Variances of the OLS Estimators

36

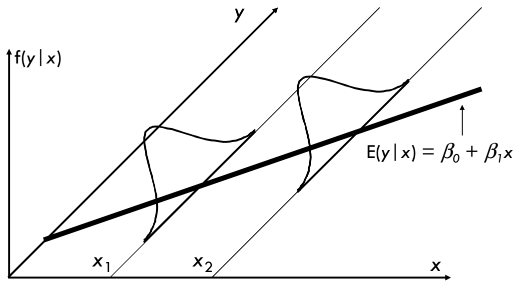
- $\text{Var}(u|x) = E(u^2|x) - [E(u|x)]^2 = E(u^2|x) = \sigma^2$ → $\text{Var}(u) = E(u^2) = \sigma^2$
- σ^2 is often called the error variance.
- σ, the square root of the error variance, is called the standard deviation of the error.
- We can say that
  $$E(y|x) = \beta_0 + \beta_1 x \qquad (2.55)$$
  $$Var(y|x) = \sigma^2 \qquad (2.56)$$

Dr. Lê Văn Chơn – FTU, 2011

2.5 Variances of the OLS Estimators

37

- Homoskedastic case:

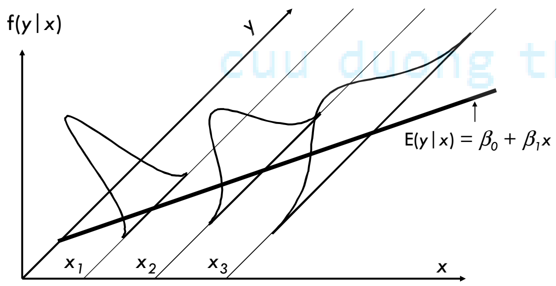


Dr. Lê Văn Chơn – FTU, 2011

2.5 Variances of the OLS Estimators

38

- Heteroskedastic case:



Dr. Lê Văn Chơn – FTU, 2011

2.5 Variances of the OLS Estimators

39

- **Theorem 2.2** (Sampling variances of the OLS estimators)
  Under Assumptions SLR.1 through SLR.5,

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{SST_x} \qquad (2.57)$$

and

$$Var(\hat{\beta}_0) = \frac{\sigma^2 \frac{1}{n}\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (2.58)$$

- Proof: $Var(\hat{\beta}_1) = \frac{1}{SST_x^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 Var(u_i) = \frac{SST_x}{SST_x^2}\sigma^2 = \frac{\sigma^2}{SST_x}$

Dr. Lê Văn Chơn – FTU, 2011

2.5 Variances of the OLS Estimators

40

- (2.57) and (2.58) are invalid in the presence of heteroskedasticity.
- (2.57) and (2.58) imply that:

(i) The larger the error variance, the larger are Var($\hat{\beta}_j$) .
(ii) The larger the variability in the x_i, the smaller are Var($\hat{\beta}_j$).

- Problem: the error variance σ^2 is unknown because we don't observe the errors, u_i.

Dr. Lê Văn Chơn – FTU, 2011

2.5 Estimating the Error Variance

41

- What we observe are the residuals, $\hat{u}_i$. We can use the residuals to form an estimate of the error variance.
- We write the residuals as a function of the errors:

$$\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i = (\beta_0 + \beta_1 x_i + u_i) - \hat{\beta}_0 - \hat{\beta}_1 x_i$$

$$\hat{u}_i = u_i - (\hat{\beta}_0 - \beta_0) - (\hat{\beta}_1 - \beta_1) x_i \qquad (2.59)$$

- An unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \hat{u}_i^2 = \frac{SSR}{n-2} \qquad (2.61)$$

Dr. Lê Văn Chơn – FTU, 2011

2.5 Estimating the Error Variance

42

- $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$ = standard error of the regression (SER).
- Recall that $sd(\hat{\beta}_1) = \frac{\sigma}{\sqrt{SST_x}}$, if we substitute $\hat{\sigma}^2$ for σ^2, then we have the standard error of $\hat{\beta}_1$:

$$se(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{SST_x}} = \frac{\hat{\sigma}}{\sqrt{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)}}$$

Dr. Lê Văn Chơn – FTU, 2011

2.6 Regression through the Origin

43

- In rare cases, we impose the restriction that when x = 0, E(y|0) = 0. E.g., if income (x) is zero, income tax revenues (y) must also be zero.
- Equation $y = \tilde{\beta}_1 x + \tilde{u}$ (2.63)
  Obtaining (2.63) is called **regression through the origin**.
- We still use OLS method with the corresponding first order condition

$$\sum_{i=1}^{n} x_i(y_i - \hat{\tilde{\beta}}_1 x_i) = 0 \quad \rightarrow \quad \hat{\tilde{\beta}}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \qquad (2.66)$$

If $\beta_0 \neq 0$, then $\hat{\tilde{\beta}}_1$ is a biased estimator of β_1.

Dr. Lê Văn Chơn – FTU, 2011