University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2. Multiple Regression Model

Prof. Dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Ljubljana, October 2022

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2.1 Population Regression Model and Sample Regression Model

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Models and Variables

MULTIPLE REGRESSION MODEL
MULTIVARIATE REGRESSION MODEL

Single equation models

dependent variable y	explanatory variables x_j; $j = 1, \dots, k$
explained variable	unexplained variables
regressand	regressors
predictand	predictors
response variable	control variables

Multiple equation models

endogenous variable(s)	exogenous variables

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

Notation of Variables

$y_i, (y_t)$ → odvisna spremenljivka

– *value of the dependent variable at i-th (t-th) observation*

$x_{ji}, (x_{jt})$ → neodvisna spremenljivka

– *value of the j-th explanatory variable at the i-th (t-th) observation*

We distinguish between:
1) Time series (t);
2) Cross sections (i),
3) Panels (it).

$$i = 1, 2, \ldots, n \quad (t = 1, 2, \ldots, T)$$
$$j = 1, 2, \ldots, k$$

n (T) – number of observations

k – number of parameters

AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

Population Regression Model

↗ you are examining all units in the population

POPULATION REGRESSION MODEL - old term

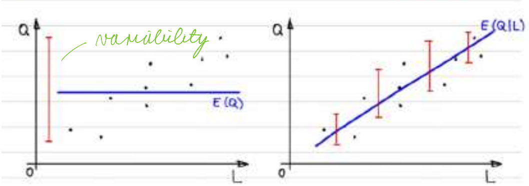
new term — **DATA GENERATING PROCESS (DGP)**

↳ mathematical structure of a certain phenomenom

Example: *PRODUCT = f (LABOUR, CAPITAL)*

Conditioning: let us assume that $E(Q|L) = f(L)$.

Slide 4



Conditioning "reduces variability".

$$E(Q|L,K) = f(L,K)$$

Conditional expected value of the dependent variable (y) is a function of explanatory variables (x)

Expected value = mathematical expectation = average value of a stochastic variable given an infinite number of measurements or realisations

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBAS ACCREDITED

Population Regression Model

Linear population regression model

(regression model = regression function = regression line = regression)

β - standarized regression coefficients

$$E(Q_i | L_i, K_i) = \beta_1 + \beta_2 L_i + \beta_3 K_i$$

these β tell us how L and K affect the value of E

$$Q_i = E(Q_i | L_i, K_i) + u_i$$

or alternatively

$$u_i = Q_i - E(Q_i | L_i, K_i)$$

$$Q_i = \beta_1 + \beta_2 L_i + \beta_3 K_i + u_i$$

u_i **= disturbances (stochastic variable)**

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Population Regression Model

Reasons for including u_i in the model:

- First, it is a **substitute** for all those variables that affect the dependent variable, but are *not included among the explanatory variables* (indicate some for our example!). Often this is caused by an incomplete or insufficient economic theory.
- Even if some variables are recognized as important explanatory variables, they often can not be included in the model specification, as they are *difficult to express numerically or obtain data* (e.g. preferences and habits, or in our case idiosyncratic knowledge).
- When there is a high probability that the total effect of omitted explanatory variables is small and insignificant, but most of all *non-systematic*, then these effects can be treated as random and captured in the model by a stochastic variable u.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Population Regression Model

- Even if one could capture all relevant explanatory variables when specifying the regression model, some random elements of the dependent variable would still be left unexplained – *pure stochastic effects or shocks*.
- Even though the classical linear regression model assumes that the values of (dependent and explanatory) *variables are measured without error*, this is often not the case in practice.
- When specifying the regression model, we should stick to the well-known rule, stating that *simpler model should be preferred* to a more complicated one, until proven that it is unsuitable due to its simplicity (Occam's razor principle).

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAs ACCREDITED

Population Regression Model

GENERAL POPULATION REGRESSION MODEL (PRM):

$$E(y_i | x_{1i}, \ldots, x_{ki}) = f(x_{1i}, \ldots, x_{ki})$$

or alternatively

$$y_i = E(y_i | x_{1i}, \ldots, x_{ki}) + u_i$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs ACCREDITED

Population Regression Model

Linear population regresion model:

$$E(y_i \mid x_{1i},\ldots,x_{ki}) = \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki}$$

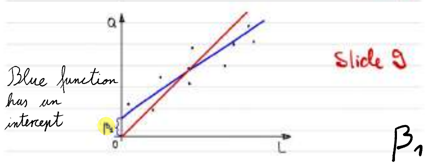
$$E(y_i \mid x_{1i},\ldots,x_{ki}) = \beta_1 + \beta_2 x_{2i} \ldots + \beta_k x_{ki}$$

or alternatively

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i$$

$$y_i = \beta_1 + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i$$

Including a constant term / an intercept:



β_1

the cost of not including an intercept is higher than including it when it is not necessary.

β_j – ***partial regression coefficient*** **of the *j*-th explanatory variable**

u_i – ***disturbances (stochastic variable)*** **at the *i*-th observation**

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

→ Regresija na vzorcu

Sample Regression Model

LINEAR SAMPLE REGRESSION MODEL (SRM):

estimator is a formula, estimate is the number you get by putting values into the formula.

$$y_i = b_1 + b_2 x_{2i} + \ldots + b_k x_{ki} + e_i$$

$$\hat{y}_i = b_1 + b_2 x_{2i} + \ldots + b_k x_{ki}$$

$$y_i = \hat{y}_i + e_i$$

$\hat{y}_i$ – estimator of the conditional expectation $E(y|x_{1i}, \ldots, x_{ki})$

b is an estimator of β

$b_1, \ldots, b_k$ – estimators of regression coefficients $\beta_1, \ldots, \beta_k$

e_i – residuals of the sample regression model

The term estimator is used in statistical literature for sample statistics. These are formulas that tell us how to calculate the parameters for population, based on (stochastic) sample data.

A given, specific numerical value of a parameter, calculated with the estimator, is then called an estimate of a parameter.

Slide 10

PRM	$E(y\|X)$	β	u
vs. SRM	$\hat{y}$	b or $\hat{\beta}$	e or $\hat{u}$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED

Sample Regression Model

Linear sample regression model:

$$\hat{y}_i = b_1 x_{1i} + b_2 x_{2i} + \ldots + b_k x_{ki}$$

$$\hat{y}_i = b_1 + b_2 x_{2i} + \ldots + b_k x_{ki}$$

or alternatively

$$y_i = b_1 x_{1i} + b_2 x_{2i} + \ldots + b_k x_{ki} + e_i$$

$$y_i = b_1 + b_2 x_{2i} + \ldots + b_k x_{ki} + e_i$$

Keep in mind an example of a SRM, such as:

Slide 11

$$Q_i = \underbrace{b_1 + b_2 L_i + b_3 K_i}_{\hat{Q}_i} + e_i$$

(fitted values)

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2.2 Estimation of the Population Regression Model and the Method of Least Squares

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Motivation

"The only way to a position in which our science might give positive advice in a large scale to politicians and business men, leads through quantitative work. For as long as we are unable to put our arguments into figures, the voice of our science, although occasionally it may help to dispel gross errors, will never be heard by practical men. They are, by instinct, econometricians all of them, in their distrust of anything not amenable to exact proof."

Joseph A. Schumpeter: "The Common Sense of Econometrics", *Econometrica*, 1, 1933, p. 12.

Estimation of the PRM

How should the estimator minimise the residuals of the SRM:

glej naslednjo stran

- **should sum of the residuals be minimal?** No
- **should sum of the absolute values of residuals be minimal?** No
- **should sum of the squared residuals be minimal?** yes

Least Squares Estimator or Method (OLS, LS)

"Discovered" by the German mathematician C. F. Gauss (at age 16). First published in 1809 and finalised in 1823.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs AMBA ACCREDITED

An estimator of β, i.e. the formula to calculate b, should somehow minimize "e".

$$e_i = y_i - \hat{y}_i$$

Infinite number of possibilities:

1) $\min\limits_{b} \sum\limits_{i} e_i$ ✗

2) $\min\limits_{b} \sum\limits_{i} |e_i|$ ✗

3) $\min\limits_{b} \sum\limits_{i} e_i^2$ ✓ $\longmapsto$ Least squares estimator

etc.

Estimation of the PRM

Carl Friederich Gauss (1777–1855)

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Estimation of the PRM

$$y_i = b_1 + b_2 x_{2i} + \ldots + b_k x_{ki} + e_i = \hat{y}_i + e_i$$

Residuals of the sample regression model:

$$e_i = y_i - \hat{y}_i =$$
$$= y_i - b_1 - b_2 x_{2i} \ldots - b_k x_{ki}$$

$$\sum e_i^2 = \sum (y_i - b_1 - b_2 x_{2i} \ldots - b_k x_{ki})^2$$

It is obvious that this sum is a function (*S*) of estimators of regression coefficients:

$$\boxed{\sum e_i^2 = S(b_1, \ldots, b_k)}$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
AMBA ACCREDITED

Method of Least Squares

Linear sample regression model (SRM):

$$y_i = b_1 + b_2 x_{2i} + \ldots + b_k x_{ki} + e_i$$

if you have 5000 companies the model has 5000 lines.

$$\begin{cases} i=1: & y_1 = b_1 + b_2 x_{21} + \ldots + b_k x_{k1} + e_1 \\ i=2: & y_2 = b_1 + b_2 x_{22} + \ldots + b_k x_{k2} + e_2 \\ \vdots & \vdots \qquad \qquad \vdots \\ i=n: & y_n = b_1 + b_2 x_{2n} + \ldots + b_k x_{kn} + e_n \end{cases}$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS AMBA ACCREDITED

It is convenient to employ matrix notation: Slide 17

$$\underset{n\times 1}{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\underbrace{\underset{n\times 1}{x_1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \underset{n\times 1}{x_2} = \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} \quad \cdots \quad \underset{n\times 1}{x_k} = \begin{bmatrix} x_{k1} \\ x_{k2} \\ \vdots \\ x_{kn} \end{bmatrix}}$$

$$\underset{n\times k}{X} = \begin{bmatrix} \boxed{\begin{matrix} 1 \\ 1 \\ \vdots \\ 1 \end{matrix}} & \boxed{\begin{matrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} x_{k1} \\ x_{k2} \\ \vdots \\ x_{kn} \end{matrix}} \end{bmatrix}$$

$$\begin{matrix} x_1 & x_2 & & x_k \end{matrix}$$

$$\underset{n\times 1}{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$$\underset{k\times 1}{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

$$\underset{n\times 1}{y} = \underbrace{\underset{(n\times k)}{X} \cdot \underset{(k\times 1)}{b}}_{n\times 1} + \underset{n\times 1}{e} \quad \text{and} \quad e = y - X\cdot b$$

$\mapsto$ Slide 18

Remember that: Slide 18

$$e^T e = [e_1 \; e_2 \cdots e_n] \cdot \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} =$$

$$= e_1^2 + e_2^2 + \ldots + e_n^2 = \sum_i e_i^2$$

Method of Least Squares

Ordinary least squares

Least squares algebra and derivation of OLS:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b}$$

$$\mathbf{e}^{T}\mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{b})^{T}(\mathbf{y} - \mathbf{X}\mathbf{b})$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs ACCREDITED

Method of Least Squares

$$\frac{\partial \mathbf{e}^{\mathrm{T}}\mathbf{e}}{\partial \mathbf{b}} = -2\mathbf{X}^{\mathrm{T}}\mathbf{y} + 2\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b} = \mathbf{0}$$

$$\mathbf{X}^{\mathrm{T}}\mathbf{y} = \mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b}$$

Least squares estimator of regression coefficients:

$$\boxed{\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}}$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

System of normal equations:

$$(X^TX)^{-1} / \quad X^T y = (X^T X) b \qquad \textbf{Slide 13}$$

$$(X^TX)^{-1} X^T y = \underbrace{(X^TX)^{-1} X^T X}_{I \text{ - identity matrix}} b$$

$$b = (X^TX)^{-1} X^T y$$

THE LEAST SQUARES ESTIMATOR

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 1 \end{bmatrix}$$

Properties of Least Squares Estimator

1. $\mathbf{X}^{\mathrm{T}}\mathbf{e} = \mathbf{0}$

2. $\hat{\mathbf{y}}^{\mathrm{T}}\mathbf{e} = \mathbf{0}$

always valid

3. $\overline{\hat{y}} = \overline{y}$

4. $\sum_i e_i = 0$

valid only if there is an intercept (constant) β_1

↳ vrednost ko $x=0$

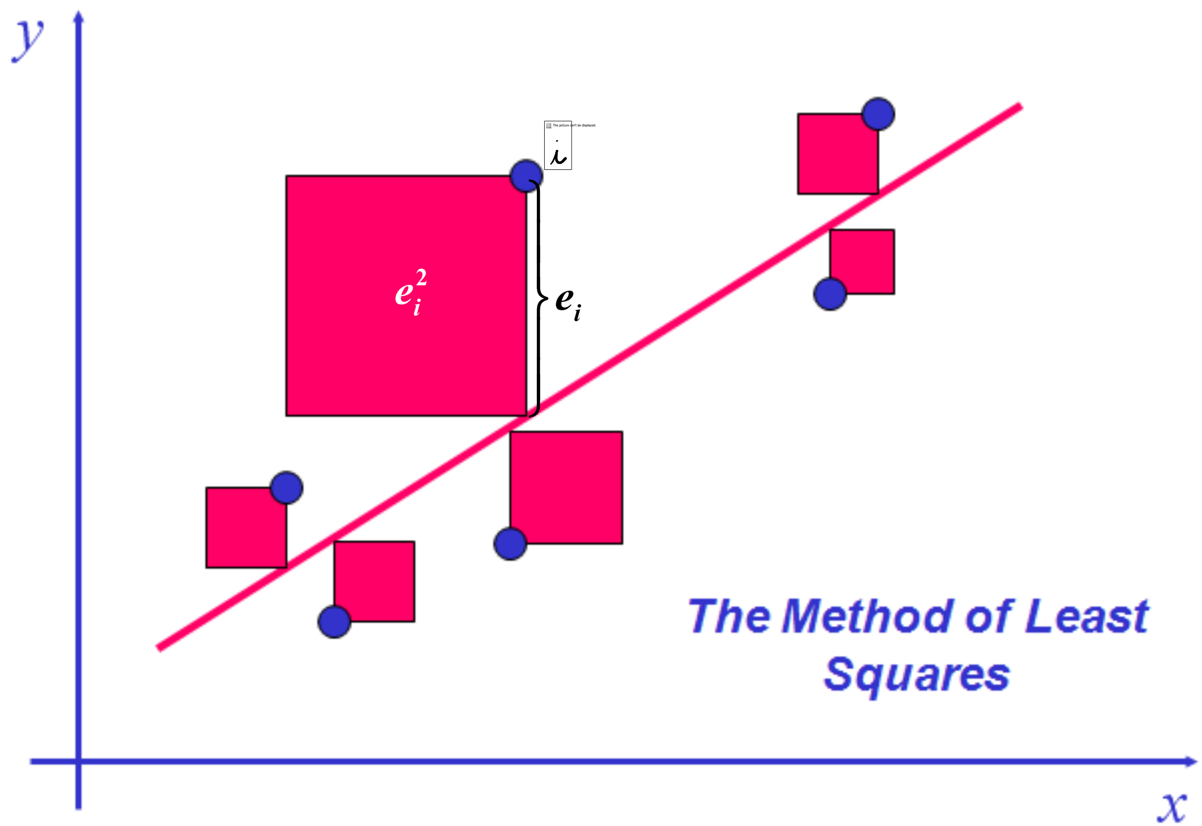
EFMD EQUIS

Remember that: Slide 20

$X^T e = 0 \iff \sum_i x_{ji} \cdot e_i = 0, \forall j = 1, \dots, k$

$\hat{y}^T e = 0 \iff \sum_i \hat{y}_i \cdot e_i = 0$

Graphic Illustration of OLS



2.3 Assumptions of the Method of Least Squares

if certain assumptions are correct this method gives you satisfactory results ⇐ optimal estimator

⇓

LS is an optimal estimator

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 1

Linearity of regression model:

Assumptions of the LS estimator

Non-linearity: Slide 23

a) In variables: x and y can be linearized

$$y_i = \beta_1 + \beta_2 \underline{x_{2i}^2} + u_i$$

$$\big\downarrow \text{linearization}$$

$$x_{2i}^2 = x_{2i}' \quad ✓$$

b) In parameters: β cannot be linearized

$$y_i = \beta_1 + \underbrace{\beta_2^2}_{\text{unknown!}} x_{2i} + u_i \quad ✗$$

$$y_i = \beta_1 + \beta_2 x_i + u_i$$

Assumption 2

Fixed (non-stochastic) values of explanatory variables at sampling (drawing samples from a distribution):

CONDITIONAL REGRESSION ANALYSIS

Fixed explanatory variables:

$E(X) = X$

$E[(X^T X)^{-1} X^T] = (X^T X)^{-1} X^T$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs AMBA ACCREDITED

Assumption 3

Zero conditional mathematical expectation of u_i:

Sum of all disturbances is 0 (some are positive, while others are negative)

$$E(u_i | x_{2i}, \ldots, x_{ki}) = 0 \text{ or in short } E(u_i) = 0;$$

which also means:

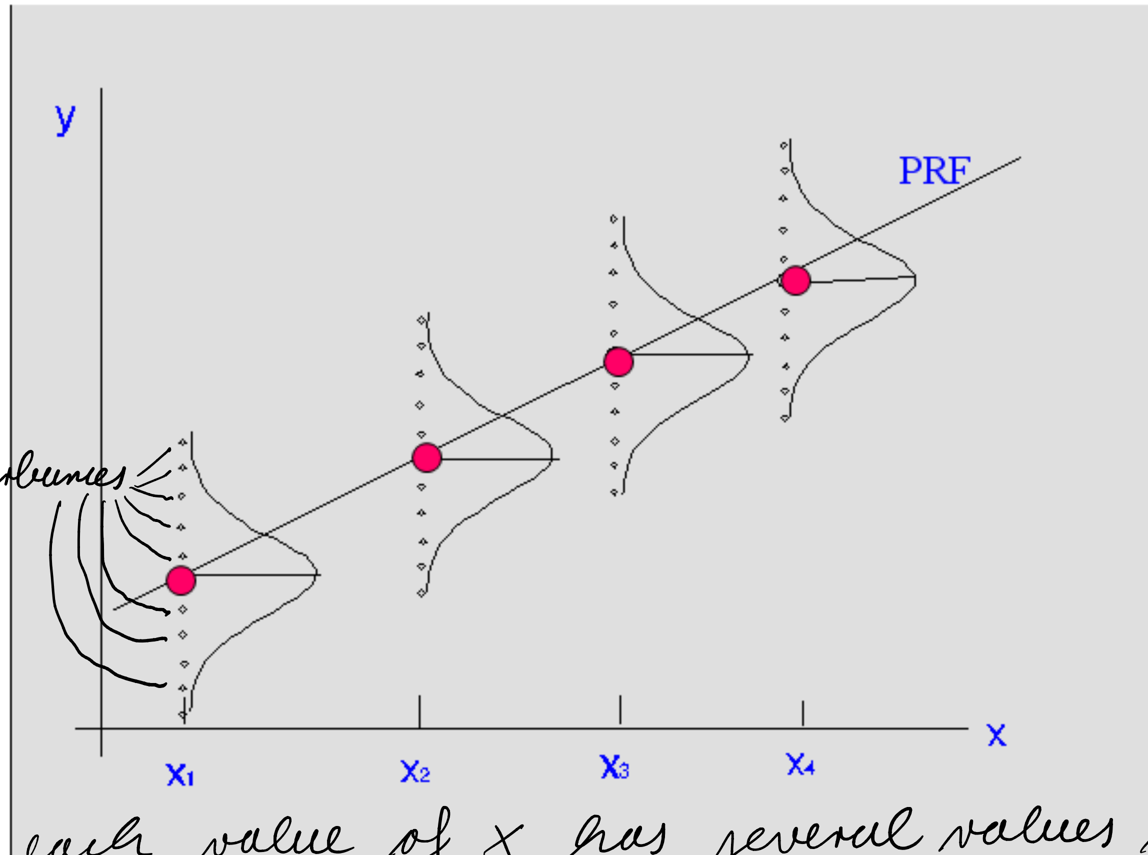
$$E(y_i | x_{2i}, \ldots, x_{ki}) = \beta_1 + \beta_2 x_{2i} + \ldots + \beta_k x_{ki}.$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 3

Distribution of stochastic variables y and u:

disturbances



each value of x has several values of y.
(for example if 10 students were given 500€
not everyone would spend the same amount of money.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 4

→ good assumption

Homoscedasticity:

$$Var(u_i|x_i) = E\left[\left(u_i - E(u_i|x_i)\right)^2 \middle| x_i\right] = E\left[u_i^2 | x_i\right] = E(u_i^2) = \sigma^2$$

or alternatively:

$$Var(u_i) = E(u_i^2) = \sigma^2$$

→ bad assumption

Heteroscedasticity: Slide 27

→ variance depends on an observation

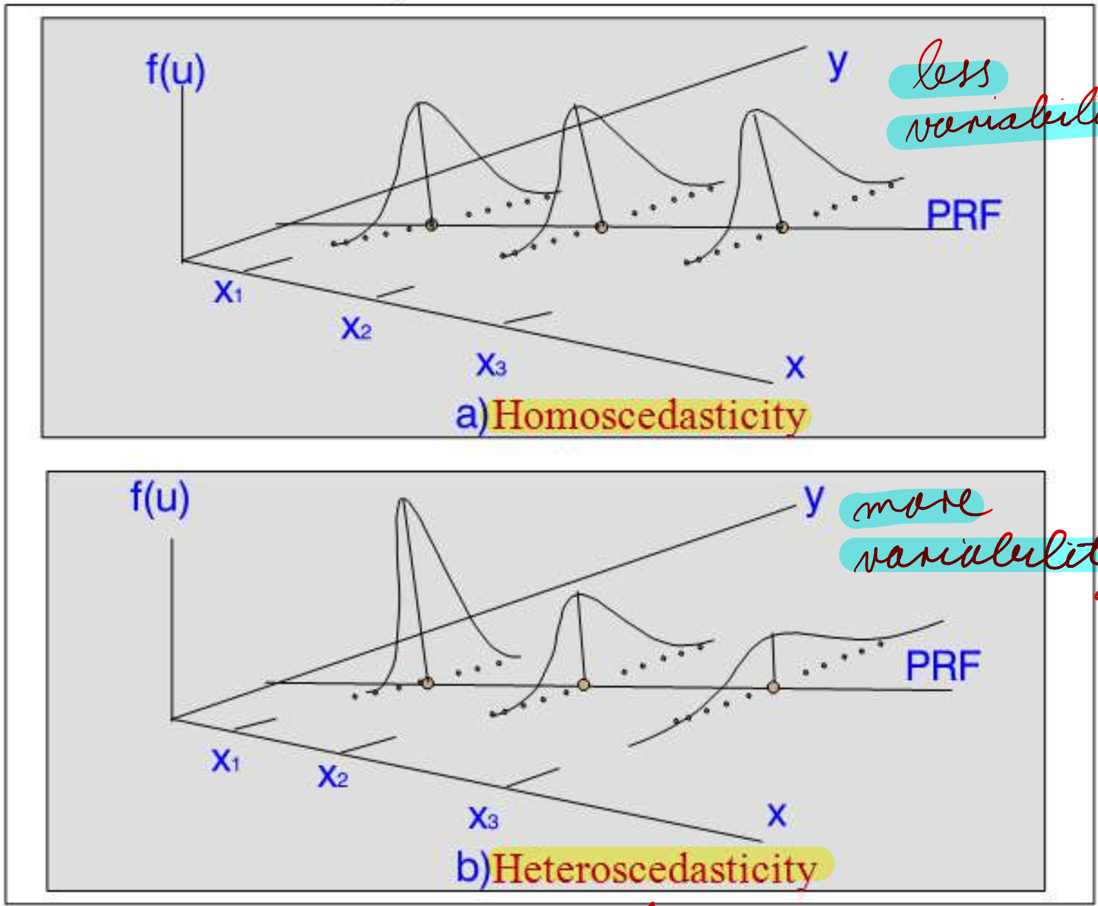
$Var(u_i) = \sigma_i^2$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 4

The distribution is the same for every value of x.

Variability of stochastic variable *u*:



The distribution is not the same for every value of x.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 5

j is a lag w.r.t. i.

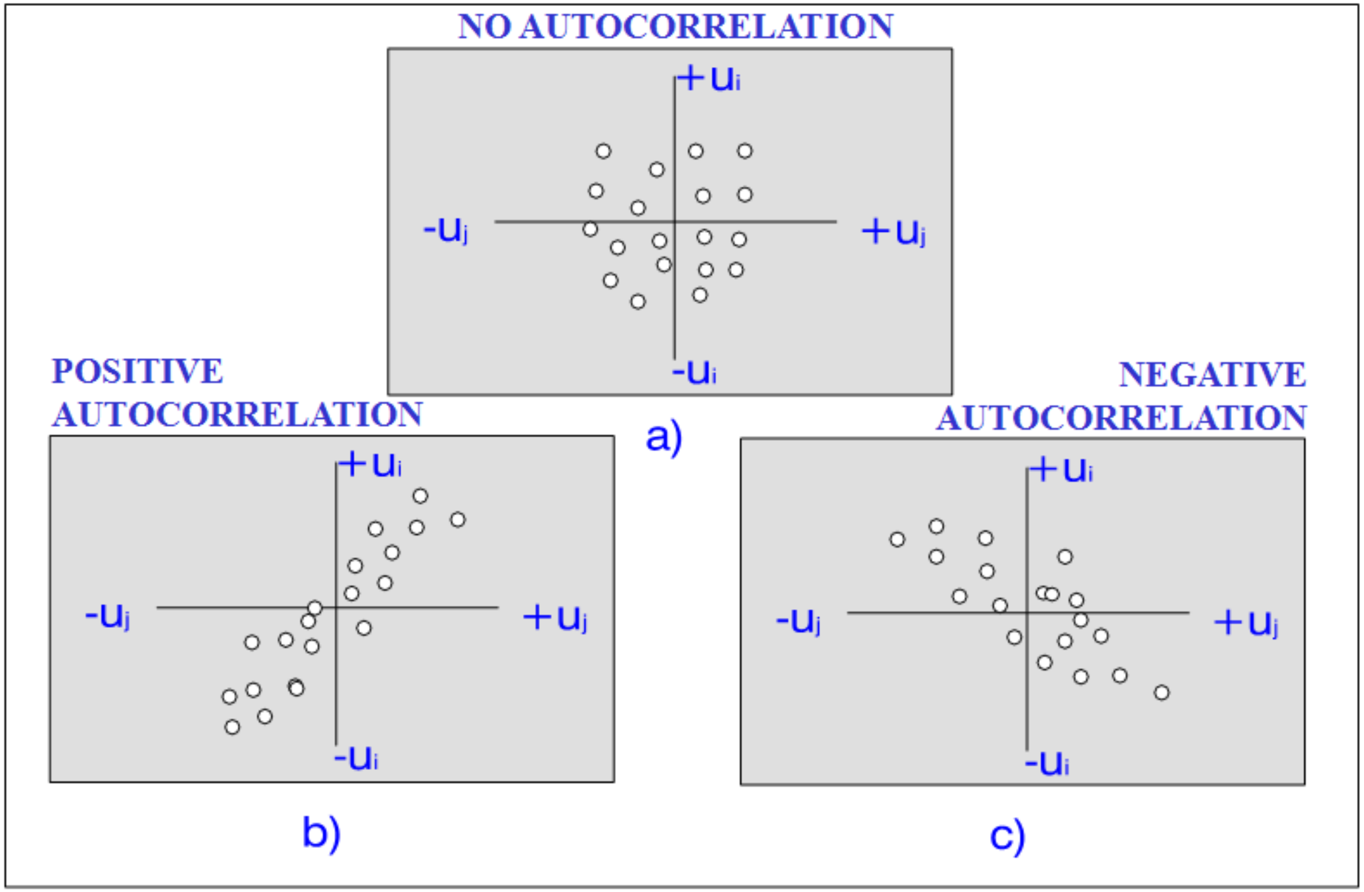
No autocorrelation:

$$Cov(u_i, u_j \mid x_i, x_j) = 0; \quad i \neq j$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
AMBA ACCREDITED

Assumption 5

Scatter plots of (lagged) stochastic variable u:



EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 6

Assumption number 3 and 5 give this assumption.

No correlation between explanatory variables and the stochastic variable *u*:

$$Cov(x_2, u) = Cov(x_3, u) = \dots = Cov(x_k, u) = 0$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 7

Important assumption for identification of the model.

Number of observations should be greater than the number of estimated parameters (explanatory variables):

$$n > k$$

Assumption 8

We are explaining variability of y with x.

Variability in values of explanatory variables:

var(*X*) is a positive finite value

Assumption 8, example: Slide 33

$y_i = b_1 + b_2 x_{2i} + e_i$, $\operatorname{var}(x_2) = 0$

Then $x_{2i} = \bar{x}_2$:

$b_2 = \frac{\sum_i (x_{2i} - \bar{x}_2)(y_i - \bar{y})}{\sum_i \underbrace{(x_{2i} - \bar{x}_2)^2}_{=0}}$ not defined

and

$b_1 = \bar{y} - \underbrace{b_2}_{\text{not defined}} \bar{x}_2$ not defined.

No b can be calculated if there is no variability.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 9

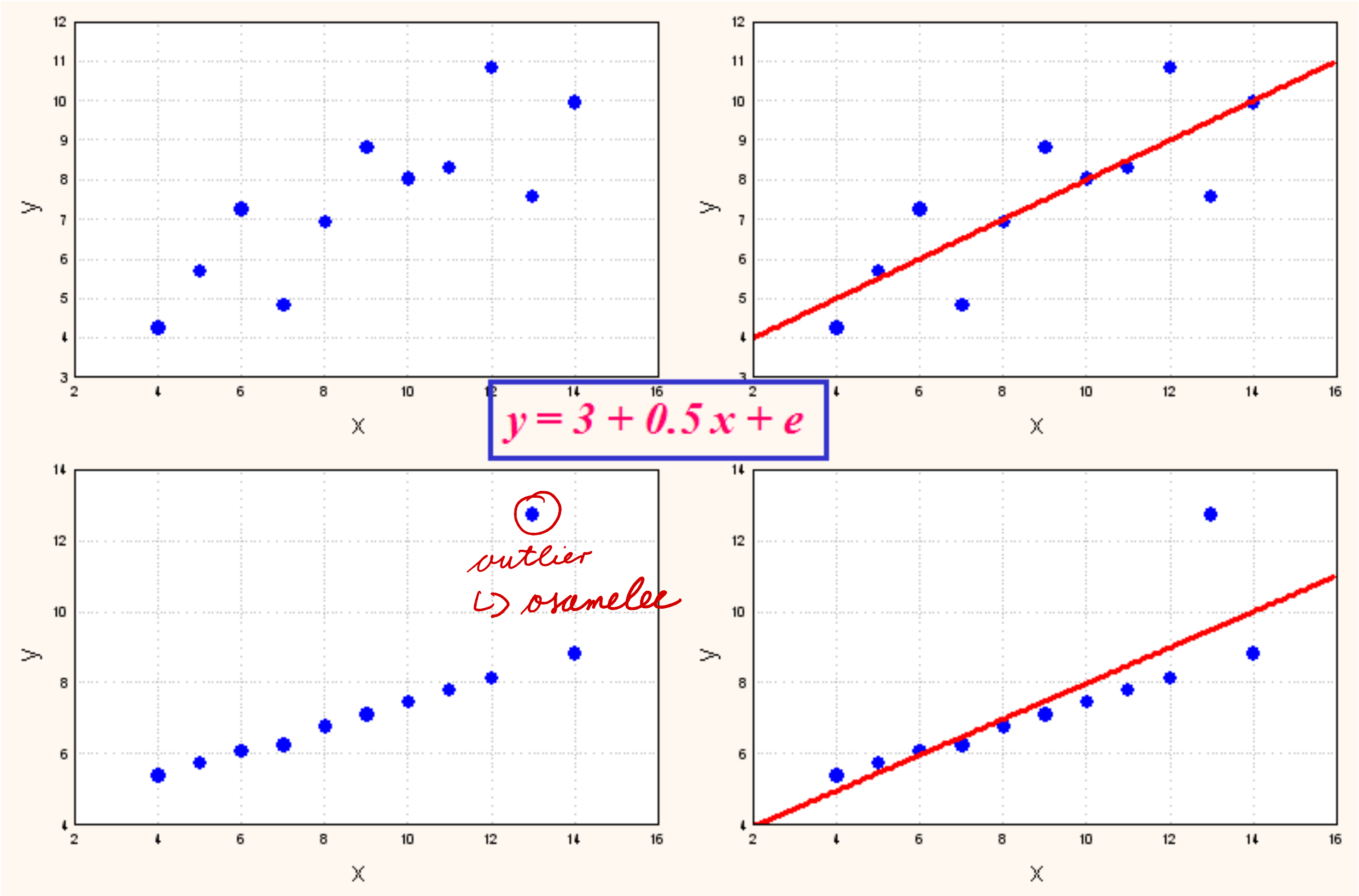
Regression model is correctly specified:

- **All relevant explanatory variables are included**
- **The right functional form of the model is chosen**

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs ACCREDITED

Assumption 9

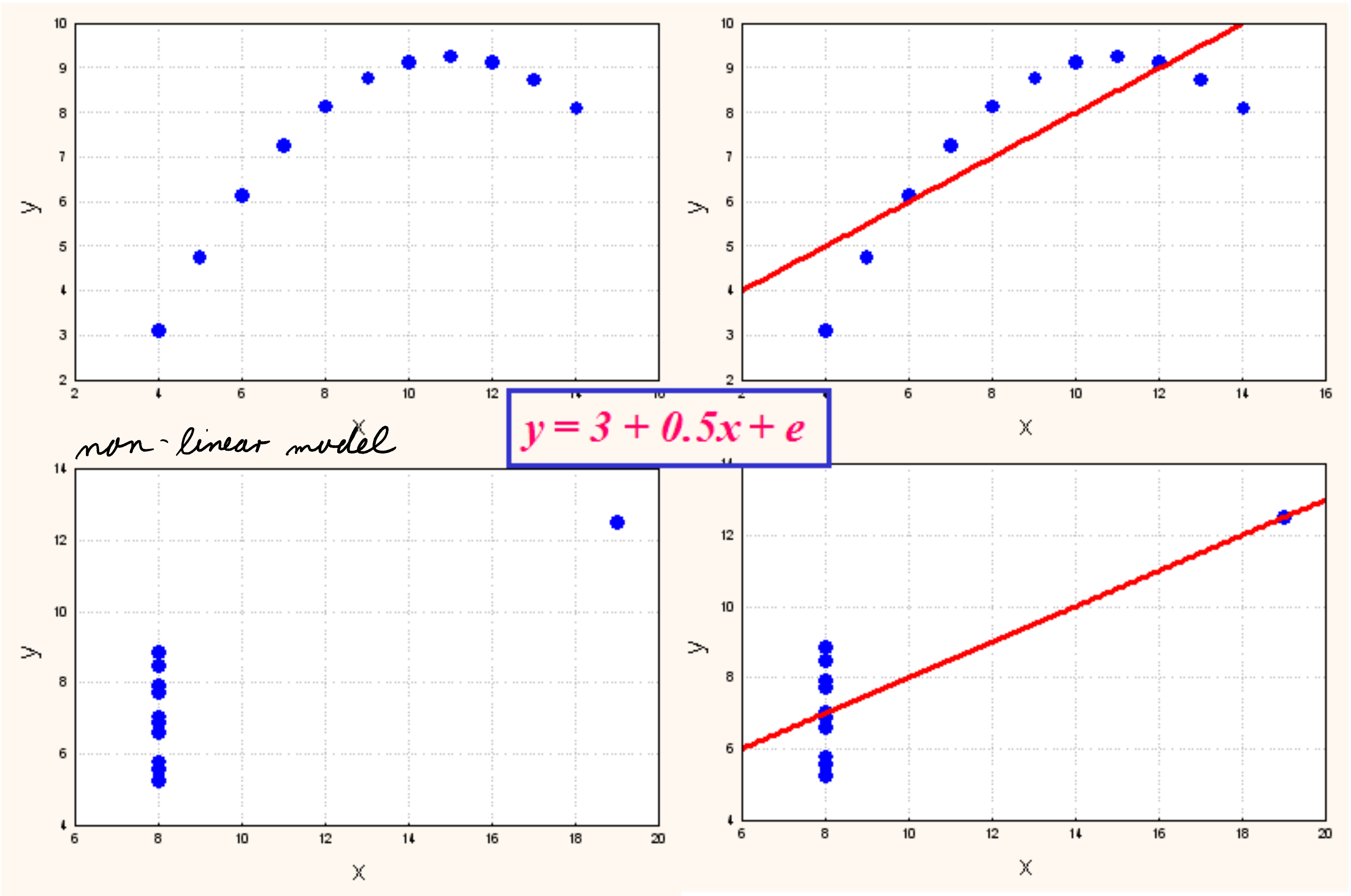
linear model



EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBAS ACCREDITED

Assumption 9

non-linear model



$\boxed{y = 3 + 0.5x + e}$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Assumption 10

Absence of perfect multicollinearity:

There exists no perfect *linear* relationship among the explanatory variables of the form:

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \ldots + \lambda_k \mathbf{x}_k = \mathbf{0}$$

Consequences of perfect (multi)collinearity

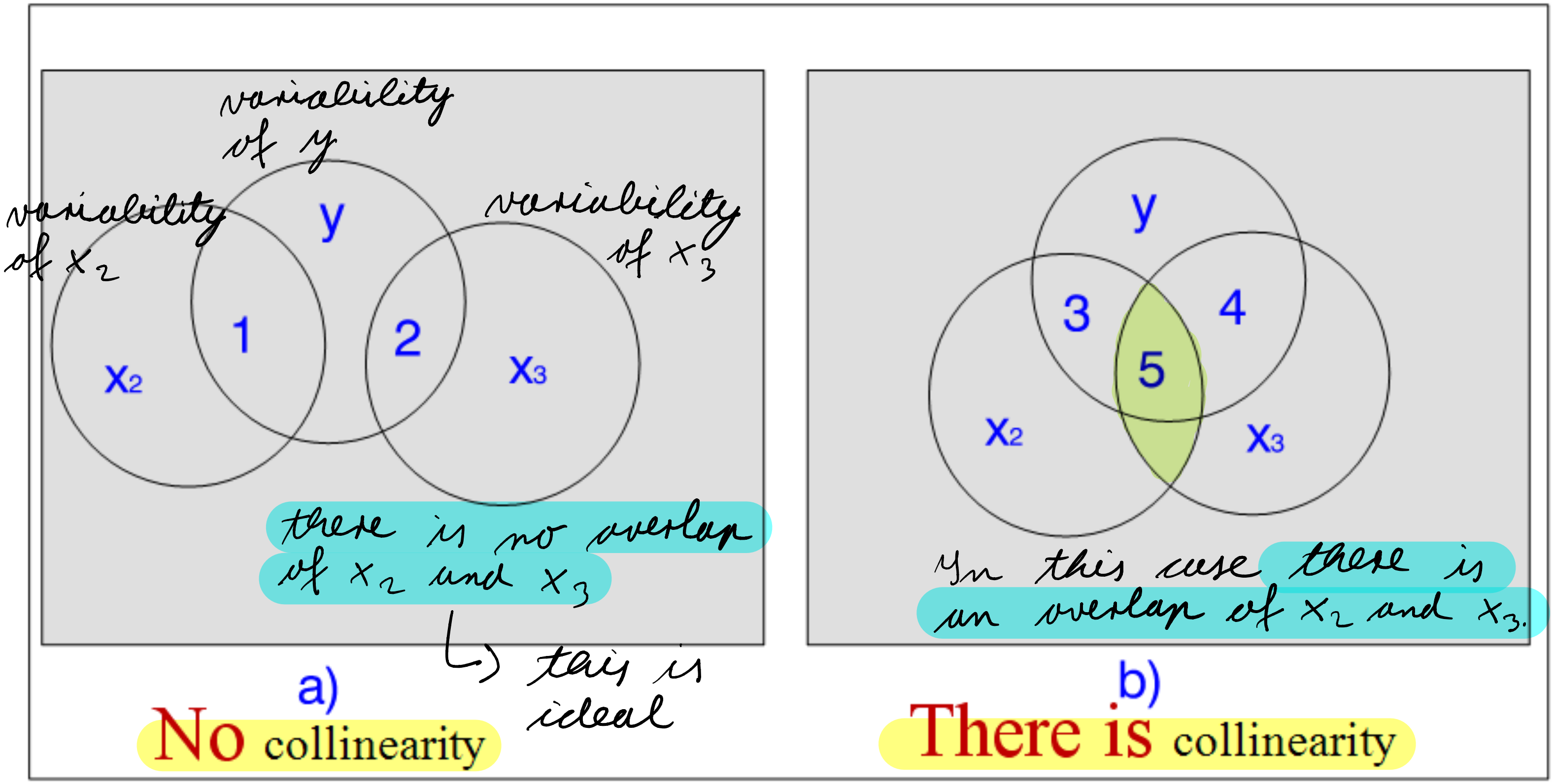
Assumption 10: Slide 37

1) $\det(X^T X) = 0$;
2) matrix $X^T X$ is singular and thus non-invertible;
3) $b = (X^T X)^{-1} X^T y$ is not defined.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAs ACCREDITED

Assumption 10

Analogy with Venn diagrams:



EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Operationalisation of the PRM

normalno porazdeliteo

Stochastic variable u is normally distributed:

$$u_i \sim N(0, \sigma_u^2)$$

NORMAL DISTRIBUTION
it has 2 parameters.
- mean
- δ^2 => variance

Consequently, the dependent variable y is also a normally distributed stochastic variable:

$$y_i \sim N(\beta_1 x_{1i} + \ldots + \beta_k x_{ki}, \sigma_u^2)$$

explanatory variables are not stochastic, since they are fixed.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2.4 Sample Properties of the Method of Least Squares

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Motivation

- The estimates of regression coefficients, obtained by the least squares estimator, are *stochastic variables*. How to establish their mean, variance, covariances, and probability distribution?

- The method of least squares is only one of possible estimators of regression coefficients. How *successful is the chosen estimator* compared to other possible stimators?

- What is the *reliability of estimates* of the regression coefficients? How good are the estimates, obtained based on one sample only?

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Sample Properties of OLS

PROPERTY A: LINEARITY

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Estimates of regression coefficients based on sample data are linear combinations of sample values of the dependent variable *y*.

PROPERTY B: CONSISTENCY

$$\mathbf{b} \xrightarrow{n \to \infty} \beta$$

when you increase the sample size b tend to be true values of parameters

estimators of regression coefficients

Estimates of regression coefficients tend to the true values of parameters, when the sample size increases.

very important

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Sample Properties of OLS

PROPERTY C: UNBIASEDNESS

$$E(\mathbf{b}) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u})$$

$E(b) = \beta$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{0} = \boldsymbol{\beta}$$

Estimates of regression coefficients are unbiased estimates of their population values.

Sample properties of the LS estimator

Unbiasedness:

Slide 43

$b = (X^T X)^{-1} X^T y = (X^T X)^{-1} X^T (X\beta + u)$

$b = \underbrace{(X^T X)^{-1} X^T X}_{I} \beta + (X^T X)^{-1} X^T u$

$b = \beta + (X^T X)^{-1} X^T u$

$E(b) = E(\beta) + E[(X^T X)^{-1} X^T u]$

$E(b) = \beta + (X^T X)^{-1} X^T \underbrace{E(u)}_{=0} = \underline{\underline{\beta}}$

To get unbiasedness you need assumption number 3 $(E(u) = 0)$ and assumption number 2 $(E(X) = X)$.

Sample Properties of OLS

an estimator is efficient if variability is low and accuracy is high

PROPERTY D: EFFICIENCY

Sample properties of the LS estimator

Unbiasedness: Slide 43

$$b = (X^TX)^{-1}X^Ty = (X^TX)^{-1}X^T(X\beta + u)$$
$$b = \underbrace{(X^TX)^{-1}X^TX}_{I}\beta + (X^TX)^{-1}X^Tu$$
$$b = \beta + (X^TX)^{-1}X^Tu$$
$$E(b) = E(\beta) + E[(X^TX)^{-1}X^Tu]$$
$$E(b) = \beta + \underbrace{(X^TX)^{-1}X^T}_{\text{fixed regressors}}\underbrace{E(u)}_{=0} = \underline{\underline{\beta}}$$

Efficiency: Slide 44

$$b - \beta = (X^TX)^{-1}X^Tu$$
$$(b-\beta)^T = u^TX\underbrace{(X^TX)^{-1}}_{\text{symmetry}}$$

transposed

$$Var-Cov(b) = E[(b-\beta)(b-\beta)^T] =$$
$$= E[(X^TX)^{-1}X^Tu \cdot u^TX(X^TX)^{-1}] =$$
$$= \underbrace{(X^TX)^{-1}X^T}_{\text{fixed regressors}}\underbrace{E(uu^T)}_{\delta^2 \cdot I}X(X^TX)^{-1} =$$
$$= \delta^2 \cdot (X^TX)^{-1}\underbrace{X^TX(X^TX)^{-1}}_{=I} =$$
$$= \underline{\underline{\delta^2 \cdot (X^TX)^{-1}}}$$

$$Var - Cov(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

Estimates of regression coefficients are efficient.

$$Var - Cov(\mathbf{b}) = \begin{bmatrix} Var(b_1) & Cov(b_1, b_2) & \cdots & Cov(b_1, b_k) \\ Cov(b_2, b_1) & Var(b_2) & \cdots & Cov(b_2, b_k) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(b_k, b_1) & Cov(b_k, b_2) & \cdots & Var(b_k) \end{bmatrix}$$

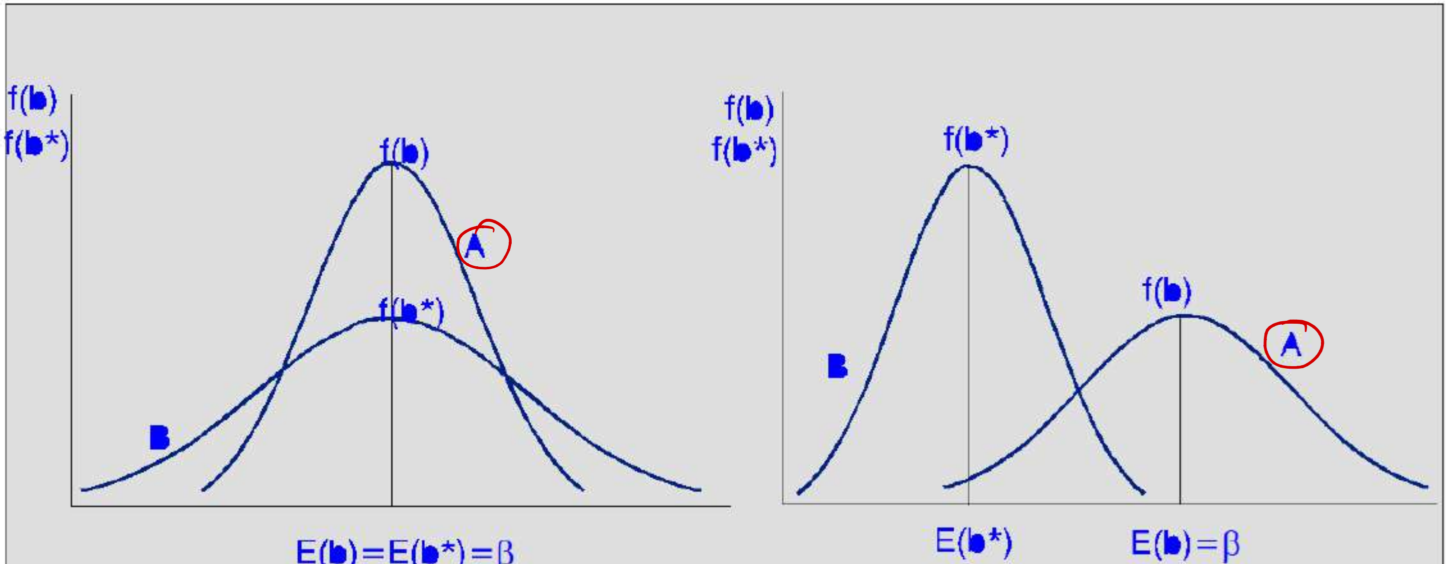
This model has the lowest variance of all linear models.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Sample Properties of OLS

Distribution of estimates of regression coefficients based on a large number of repeated samples is normal:

$$\mathbf{b} \sim N\left[\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}\right]$$



a) Both estimators are unbiased, but A is more efficient.

b) Estimator B is more efficient, but biased. Estimator A is unbiased, but less efficient.

Unbiasness > Efficiency > Consistency

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Sample Properties of OLS

You need to make sure that these properties apply for the sample
⇓
then you can use the OLS method.

Variance of the regression coefficient estimates:

$$Var(b_2) = \frac{\sigma^2}{nVar(x_2)} \cdot \frac{1}{(1 - r^2_{x_2 x_3})}$$

Variance decreases, when:

1. **the sample size (n) increases, i.e. when the calculation of a regression coefficient estimate is based on more observations;**
2. **the variability (variance) of an explanatory variable x increases;**
3. **the variability (variance) of the stochastic variable u decreases;**
4. **the linear dependence among the explanatory variables decreases.**

the higher the variability of x_2, the lower variability of y

δ^2

there should be no relationship between the explanatory variables.

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs AMBA ACCREDITED

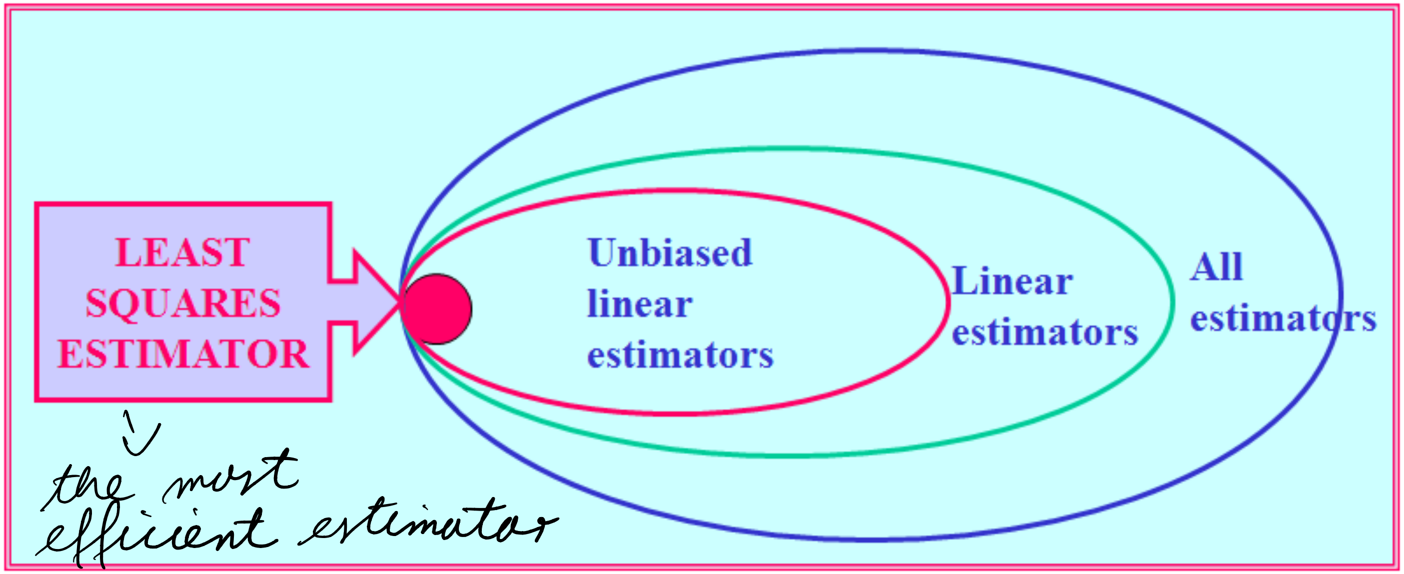
Gauss–Markov Theorem

The method of least squares is

Best **L**inear **U**nbiased **E**stimator.

most efficient

BLUE



Monte Carlo Experiments

In simulation you choose the true values and then you do the sampling

Monte Carlo experiments
(Los Alamos – Manhattan)

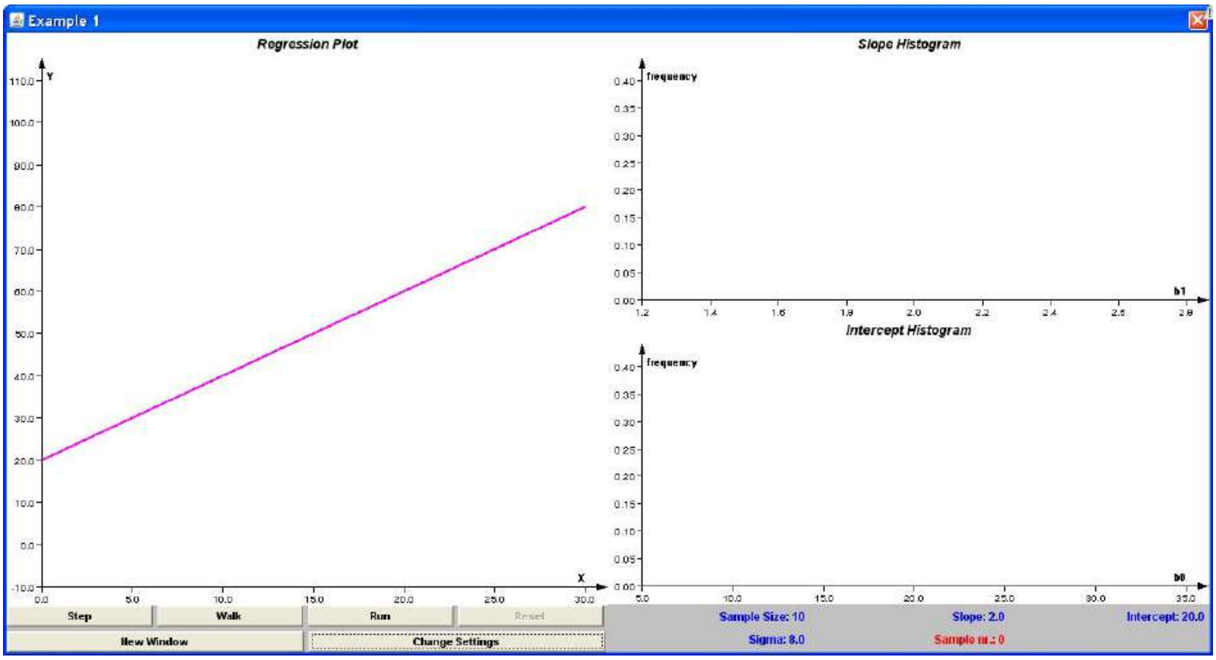
population regression model

* We define a specific form for the PRM;
* We set the values of regression coefficients, values of explanatory values, and calculate the values of the dependent variable;
* We prepare the "mended" values of the dependent variable for a large number of samples, such that at any given sample we add a random value (drawing from a distribution) of the stochastic variable u;
* For each of the prepared samples, we calculate the estimates of regression coefficients by applying OLS;
* We prepare the distributions of estimates of regression coefficients and we analyze them.

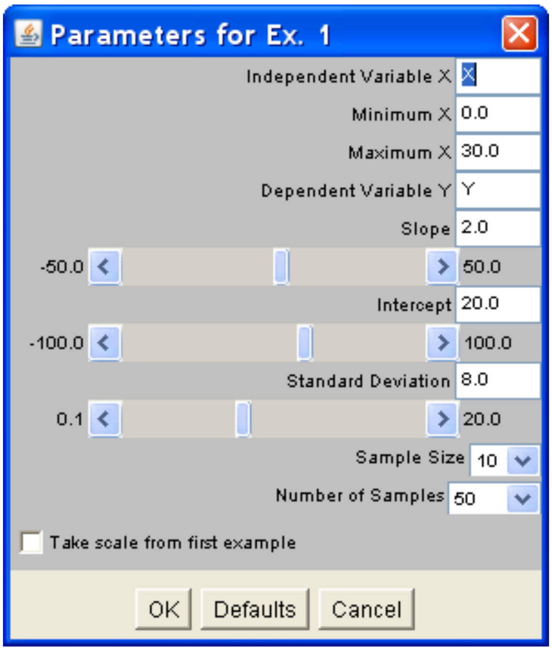
Monte Carlo experiments: http://lstat.kuleuven.be/newjava/vestac

Regression: Histograms on simple linear regression

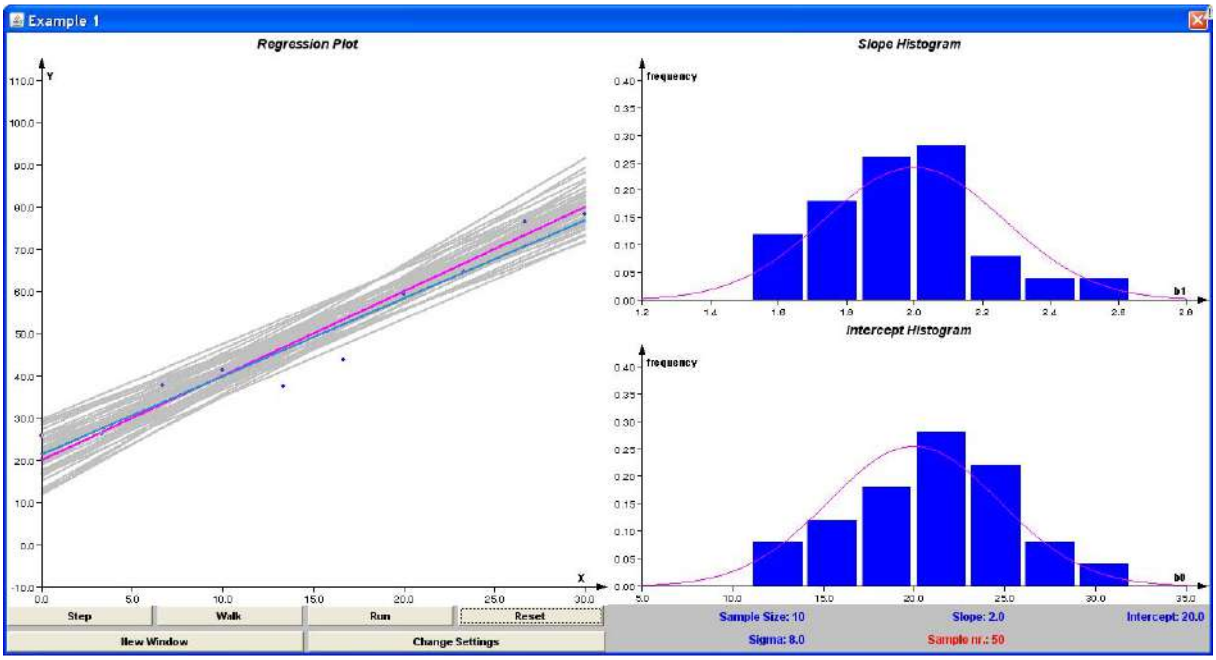
MAIN WINDOW:



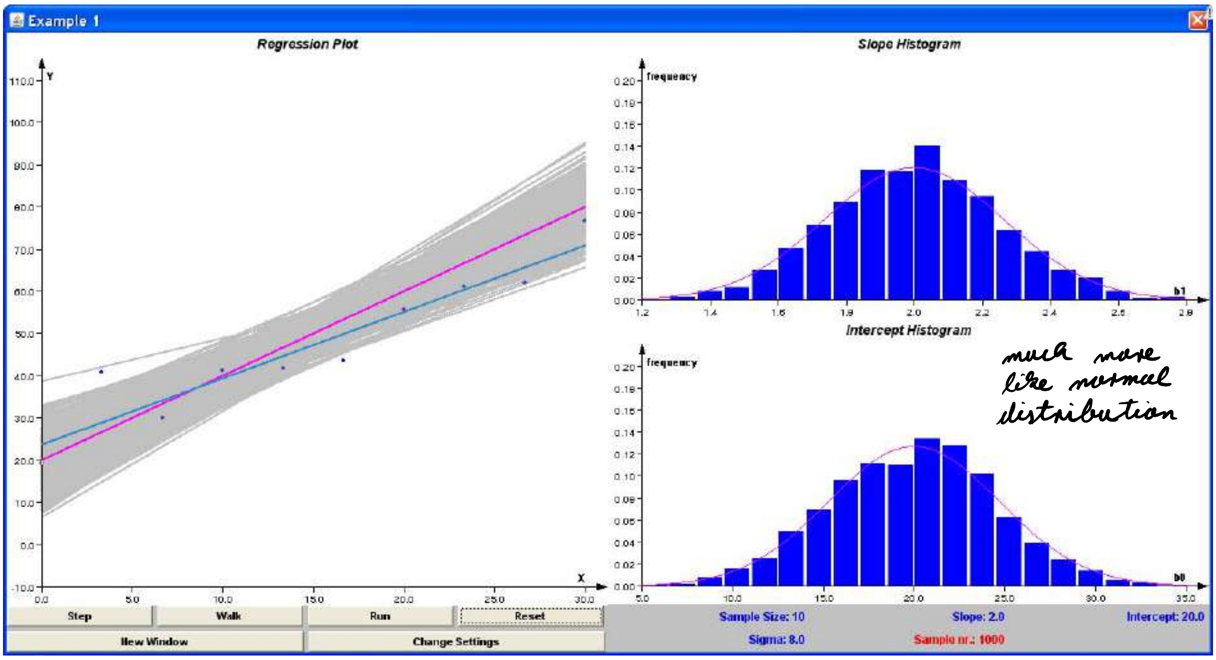
INITIAL SETTINGS:



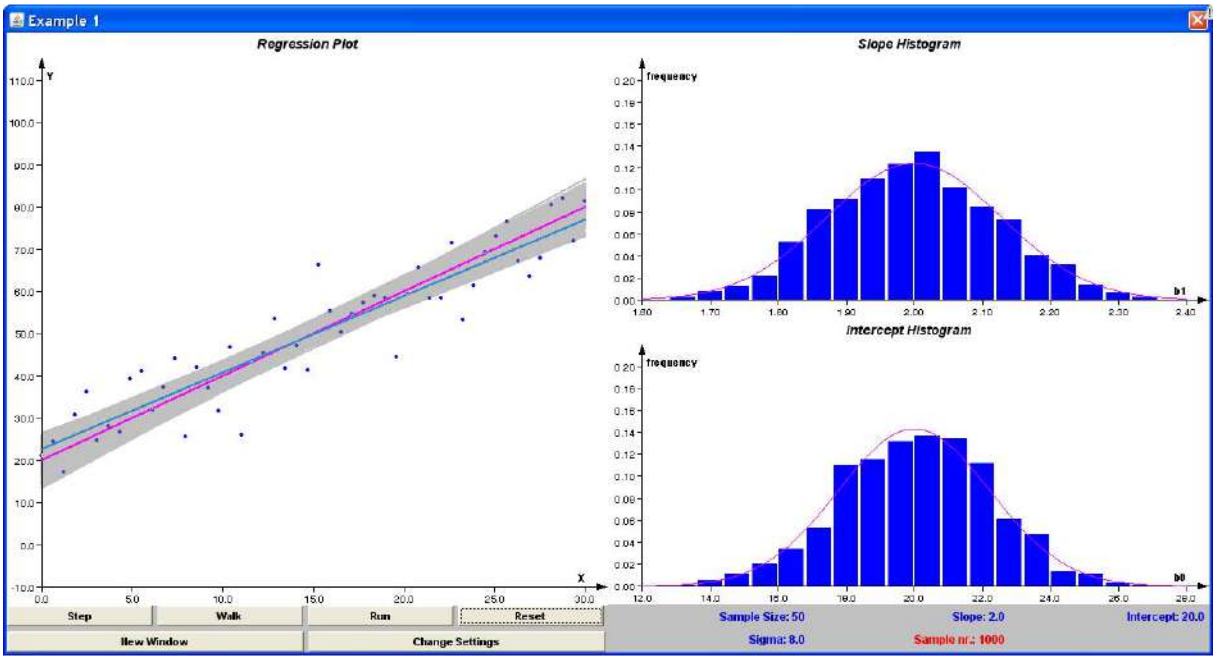
EXPERIMENT 1:



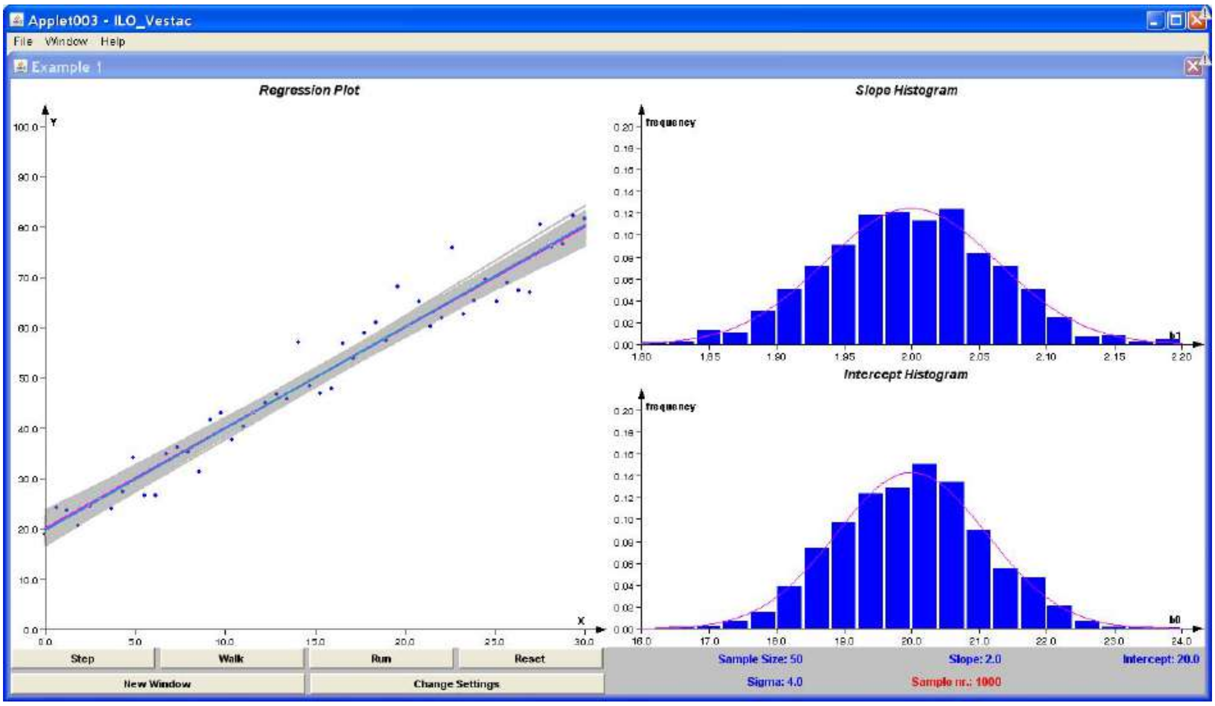
EXPERIMENT 2:



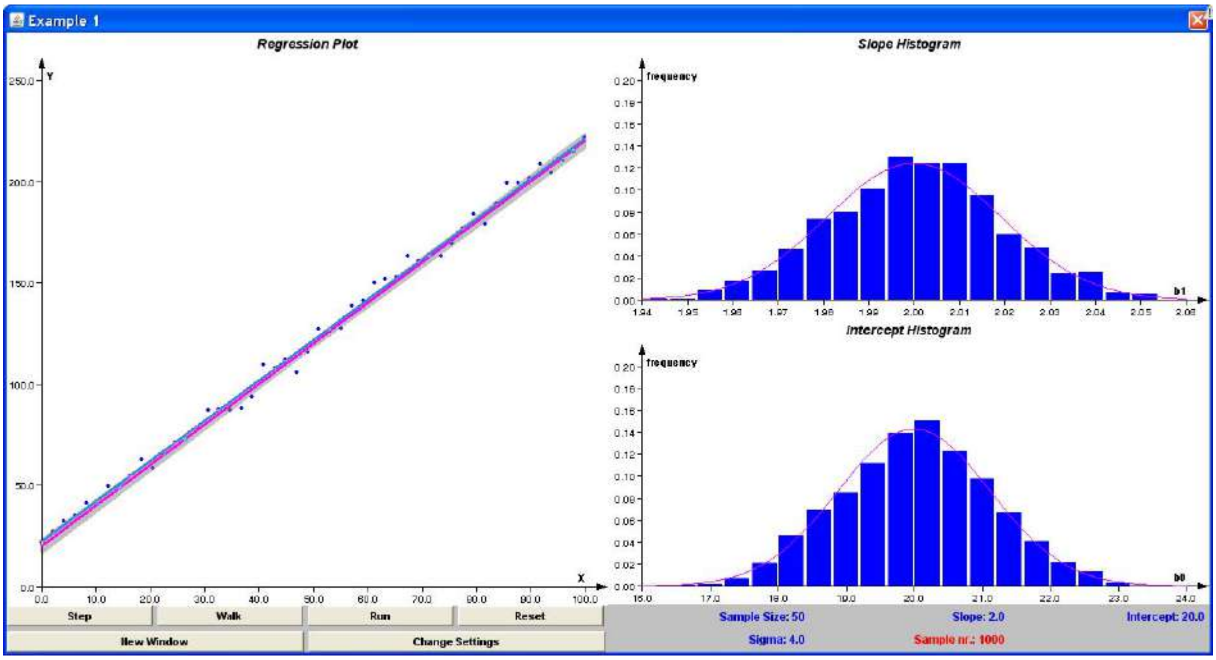
EXPERIMENT 3:



EXPERIMENT 4:



EXPERIMENT 5:



Monte Carlo experiments PDF

Experiment 1

PRM: $y_i = \beta_1 + \beta_2 x_{2i} + u_i$
SRM: $y_i = b_1 + b_2 x_{2i} + e_i$

$$\left.\begin{array}{l}\beta_1 = 20 \\ \beta_2 = 2 \\ n = 10\end{array}\right\} E(y_i \mid x_{2i}) = 20 + 2x_{2i}\,;\ 0 \leq x_{2i} \leq 30$$

Monte Carlo replications

We prepare 50 samples, $R=50$:

$y_i = 20 + 2x_{2i} + \underbrace{u_i}_{\text{drawing from the distribution}}\,;\ i = 1, \dots, 10$

where $u \sim N(0, 8^2)$, $\sigma_u = 8$ and thus
$y \sim N(20 + 2x_{2i}, 8^2)$.

We estimate the PRM on $R=50$ samples and analyze the distribution of estimates.

Experiment 2

$\uparrow R = 1000$, the rest as in E1.

increase the number of samples $\Rightarrow$ the distribution is much more closer to the normal distribution.

Slide 46

Experiment 3

več enot v vzorcu

$n=50$, the rest as in E2.

↳ we are going to increase efficiency

↳ gray set of lines is narrower ⇒ variability is lower

Experiment 4

lower standard deviation ⇒ you choose it in the simulation

$\sigma_u = 4$, the rest as in E3.

↳ effeciency is increased ⇒ variability is even lower

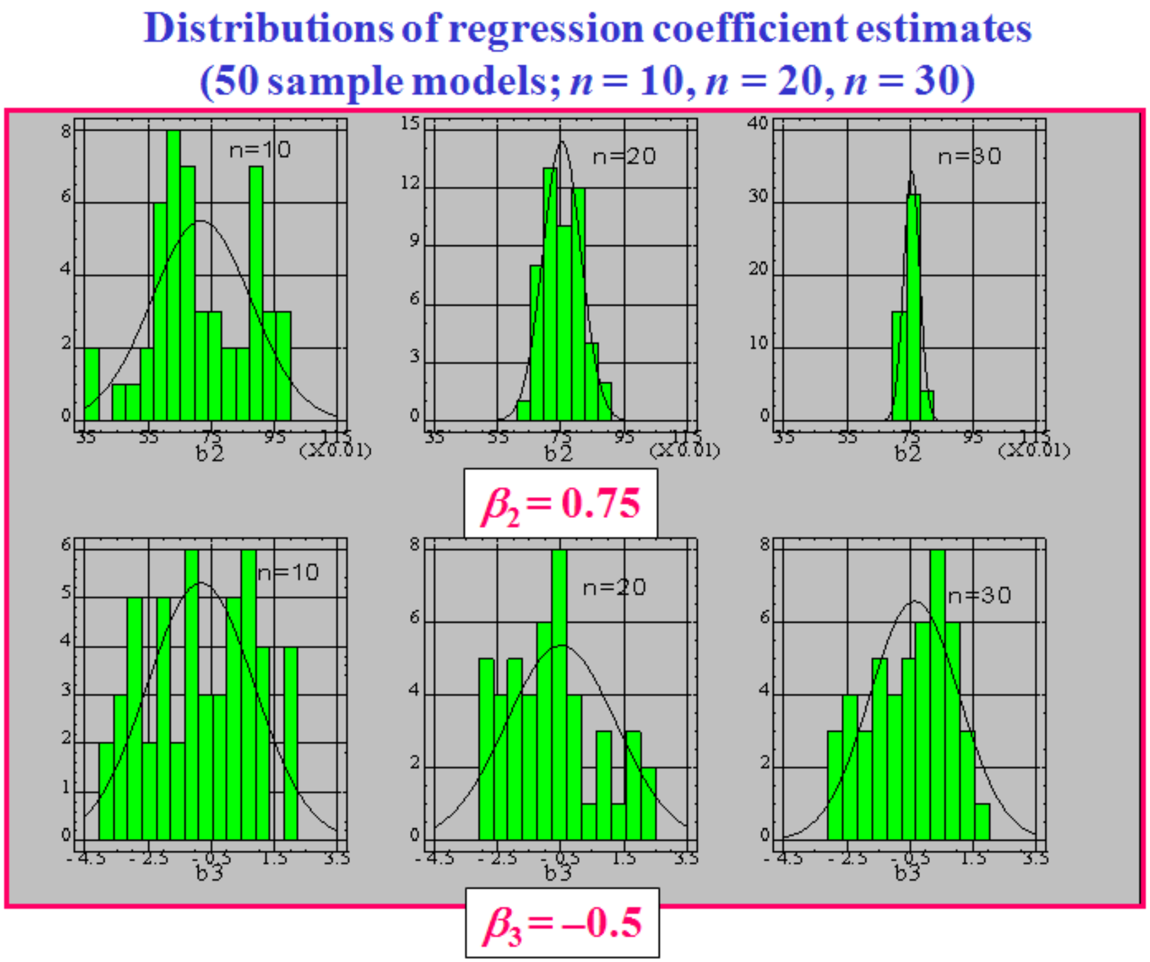
Experiment 5

doing this leads to lower variance of b_2

$Var(x_2)\uparrow$, such that $0 \leq x_{2i} \leq 1000$, the rest as in E4.

↳ efficiency is increased once again

Monte Carlo Experiments



Variance estimator of disturbances *u*

Variance estimator of disturbances u, σ_u^2:

$$s_e^2 = \frac{\mathbf{e}^{\mathrm{T}}\mathbf{e}}{n-k}$$

degrees of freedom

sample size - number of parameters

s_e^2 is an unbiased estimator of σ^2

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Estimator of the var–cov matrix of **b**

Estimator of the var–cov matrix of regression coefficient estimates, Var – Cov(**b**):

Variance covariance matrix of the regression coefficient estimates

$$\text{var}-\text{cov}(\mathbf{b}) = s_e^2(\mathbf{X}^T\mathbf{X})^{-1}$$

Estimator of the var–cov matrix of b:

$\text{Var}-\text{Cov}(b) = \sigma^2 \cdot (X^T X)^{-1}$ → dejanske vrednosti

Slide 51

$\text{var}-\text{cov}(b) = s_e^2 \cdot (X^T X)^{-1}$ → ocene vrednosti

Distribution of regression coef. estimates

$$\mathbf{b} \sim N\left[\boldsymbol{\beta}, \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\right]$$

Testing a regression coefficient:

$H_0: \beta_j = \beta_{j0}$ null hypothesis

$H_1: \beta_j \neq \beta_{j0}$ alternative hypothesis

$n-k$: degrees of freedom

Slide 52-55

$$b_j \sim N(\beta_j, Var(b_j))$$

$$z = \frac{b_j - \beta_j}{\sqrt{Var(b_j)}} \sim N(0,1)$$

based on true value of the sample ⇒ must be replaced with an estimator

$var(b_j)$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Distribution of regression coef. estimates

Calculation of the *t*–statistic based on sample data

$$t_j = \frac{b_j - \beta_j}{\sqrt{\mathrm{var}(b_j)}} = \frac{b_j - \beta_j}{se(b_j)}$$

$$se(b_j) = \sqrt{\mathrm{var}(b_j)} = \sqrt{s_e^2 \left[(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1} \right]_{jj}}$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAs ACCREDITED

Distribution of regression coef. estimates

He discovered the student t' distribution

William Sealy Gosset (1876–1937)

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Distribution of regression coef. estimates

Student's *t*–distribution:

$$t_j = \frac{b_j - \beta_j}{se(b_j)} \sim t_{(n-k)}$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs ACCREDITED

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2.5 Measures of Fit of a Regression Model

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Standard Error of Regression

Standard error of regression

$$s_e = \sqrt{s_e^2} = \sqrt{\frac{\mathbf{e}^{\mathrm{T}}\mathbf{e}}{n-k}}$$

Coefficient of variation

$$KV = \frac{s_e}{\bar{y}} \quad \text{or} \quad KV\% = \frac{s_e}{\bar{y}}100$$

mean of
the dependant
variable

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Analysis of Variance (ANOVA)

Decomposition of sum of squares (SS):

total sum of squares

residual sum of squares

$$TSS = ESS + RSS$$

explained sum of squares

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \mathbf{y}'\mathbf{y} - n\bar{y}^2$$

fitted values

$$ESS = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n} \hat{y}_i^2 - n\bar{y}^2 = \hat{\mathbf{y}}'\hat{\mathbf{y}} - n\bar{y}^2$$

$$ESS = \mathbf{b}'\mathbf{X}'\mathbf{y} - n\bar{y}^2 = \mathbf{y}'\mathbf{X}\mathbf{b} - n\bar{y}^2$$

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2 = \mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{y}}'\hat{\mathbf{y}}$$

value of the dependent variable

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
Association of MBAs AMBA ACCREDITED

Determination Coefficient

Determination coefficient of multiple regression
Multiple determination coefficient

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$R^2 = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}} - n\bar{y}^2}{\mathbf{y}'\mathbf{y} - n\bar{y}^2} = \frac{\mathbf{b}'\mathbf{X}'\mathbf{y} - n\bar{y}^2}{\mathbf{y}'\mathbf{y} - n\bar{y}^2}$$

– if you have the intercept

$$\boxed{R_*^2 = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}} = \frac{\mathbf{b}'\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}}}$$

– if you do not have the intercept

Multiple determination coefficient:

$0 \leq R^2 \leq 1$ Slide 59

$$R^2 = (r_{y\hat{y}})^2 = r_{y\hat{y}}^2$$

Determination Coefficient

Adjusted determination coefficient (H. Theil, 1971)

$$\text{TSS} = \text{ESS} + \text{RSS}$$

$$\text{TSS} = (\text{TSS} - \text{RSS}) + \text{RSS}$$

Degrees of freedom:

number of parameters in the system that can change freely
↳ it is not limited to statistics

$$(n-1) = ((n-1)-(n-k)) + (n-k)$$

$$(n-1) = (k-1) + (n-k)$$

adjusted determination coefficient

$$\bar{R}^2 = 1 - \frac{\frac{\text{RSS}}{(n-k)}}{\frac{\text{TSS}}{(n-1)}} = 1 - \frac{\text{RSS}}{\text{TSS}}\frac{(n-1)}{(n-k)}$$

$$\bar{R}^2 = 1 - (1-R^2)\frac{(n-1)}{(n-k)}$$

Example: of negative $\bar{R}^2$ Slide 60

$R^2 = 0.25$
$n = 12$
$k = 5$

$\bar{R}^2 = R^2_{adj} = 1-(1-R^2)\cdot\frac{n-1}{n-k} =$
$= 1 - 0.75 \cdot 1.571 = -0.179$

Determination Coefficient

Weaknesses of determination coefficient and adjusted determination coefficient:

- A high value of determination coefficient does not yet mean that the right explanatory variables are included in the model;
- Values of determination coefficient are not comparable among models with differently defined dependant variable;
- Values of determination coefficient are in general higher for time series models than for cross-section based models;
- A low value of determination coefficient does not mean that the model does not include relevant explanatory variables.

↳ it just means that it does not have all of the explanatory variables

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAs ACCREDITED

Testing the Model as a Whole

Testing for the statistical significance of a regression model as a whole:

R. A. Fisher established in 1922 that the ratio between explained and residual variance, taking into account the respective degrees of freedom, is distributed according to a specific distribution, named after him the ***F*–distribution**.

Testing stat. sign. of the model as a whole:

for every

$H_0: \beta_j = 0, \forall j = 2, \ldots, k$ Slide 62

$H_1: \beta_j \neq 0, \exists j = 2, \ldots, k$

for at least one

this tells you if the model actually explains anything

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Testing the Model as a Whole

Ronald Aylmer Fisher (1890–1962)

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Testing the Model as a Whole

F - test

$$F = \frac{\frac{ESS}{k-1}}{\frac{RSS}{n-k}} \sim F_{(k-1,n-k)}$$

$$\frac{\text{explained sum of squares}}{\text{residual sum of squares}}$$

$$F = \frac{\frac{ESS}{TSS(k-1)}}{\frac{RSS}{TSS(n-k)}} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \sim F_{(k-1,n-k)}$$

Testing the Model as a Whole

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

$$F > F_c$$

$$R^2 > R^2_c$$

$$R_c^2 = \frac{F_c(k-1)}{F_c(k-1)+(n-k)}$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

Likelihood Function and Info. Criteria

these models are only used for interpretation

Log–likelihood value of the model:

$$\ln L = -\frac{n}{2}\left[\ln(2\pi) + \ln\left(\frac{\text{RSS}}{n}\right) + 1\right]$$

↳ a larger value means better fit

Akaike information criterion (AIC):

$$AIC = -2\ \ln L + 2\ k$$

lower value means better fit

Schwarz criterion (SC or BIC):

Bayesian IC

$$SC = -2\ \ln L + k\ \ln(n)$$

lower value means better fit

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2.6 Interpretation of Regression Coefficients

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Representation of Estimation Results

Representation of results of an estimated regression model

$$y = \beta_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

first you need to write population regression model

$$\begin{array}{ccccccc}
\hat{y} & = & b_1 & + & b_2 x_2 & + \cdots & + b_k x_k \\
 & & (se(b_1)) & & (se(b_2)) & & (se(b_k)) \\
 & & (t_1) & & (t_2) & \cdots & (t_k) \\
 & & (p_1) & & (p_2) & \cdots & (p_k)
\end{array}$$

then you need to write a sample regression model

you can choose one of these values to write it under SRM

either $se(b_k)$, t_k or p_k

$$n = \ldots. \quad R^2 = \ldots \quad \bar{R}^2 = \ldots\ldots$$

$$s_e = \ldots.. \quad (F; DW; h)$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Model estimation

Example: Production functions PDF a,b

Theory: $Q = f(L, K)$

product (Q), labour (L), capital (K)

Data:

- Q – value added in 1.000 m.u.
- L – average number of employees
- K – sum of tangible and intangible assets in 1.000 m.u.

you would define it as number of working hours

Model 1: Linear production function

$$Q_i = \underbrace{\beta_1 + \beta_2 L_i + \beta_3 K_i}_{E(Q_i | L_i, K_i)} + u_i \qquad (PRM)$$

u_i – disturbance

tradition mathematical expectation

$$Q_i = \underbrace{b_1 + b_2 L_i + b_3 K_i}_{\hat{Q}_i} + e_i \qquad (SRM)$$

Model 2: Cobb-Douglas production function

$$Q_i = \beta_1 \cdot L_i^{\beta_2} \cdot K_i^{\beta_3} \cdot u_i \quad / \ln \qquad \text{(PRM)}$$

↳ non linear model

↓ Linearization ⇒ with logaritams ⇒ tako linearizeramo nelinearne funkcije

$$\ln Q_i = \ln [\beta_1 \cdot L_i^{\beta_2} \cdot K_i^{\beta_3} \cdot u_i]$$

$$\ln Q_i = \ln \beta_1 + \ln L_i^{\beta_2} + \ln K_i^{\beta_3} + \ln u_i$$

$$\underbrace{\ln Q_i}_{LQ_i} = \underbrace{\ln \beta_1}_{\beta_1'} + \beta_2 \underbrace{\ln L_i}_{LL_i} + \beta_3 \underbrace{\ln K_i}_{LK_i} + \underbrace{\ln u_i}_{u_i'}$$

$$LQ_i = \underbrace{b_1 + b_2 LL_i + b_3 LK_i}_{\widehat{LQ_i}} + e_i \qquad \text{(SRM)}$$

Linearized Cobb-Douglas production function.
Regression coefficients now represent elasticities.

Stata:
- data exploration;
- logarithmic transformation;
- scatter plots.

PDF, pp. 1–3.

Example: We analyse production functions for 81 manufacturing companies in the computer manufacturing industry for a given year (the data are provided in Stata Data file `production.dta`, while the programming code is given in Stata Do file `production-commands-l05.do`). We have cross-section data available for the following variables:

- value added as a proxy for the product (Q; in 1,000 monetary units);
- average number of employed workers as a proxy for labour (L);
- sum of tangible and intangible assets as a proxy for capital (K; in 1,000 monetary units).

a) Explore the data using different Stata commands. By using the scatter plots, examine the relationships of the linear and the Cobb-Douglas production function.
b) Estimate the linear and the Cobb-Douglas production function by the least squares estimator. Interpret the results of both models. Also, calculate the regression coefficients from the computer printouts manually.
c) Study direct, indirect and total effects of explanatory variables on the dependent variable for both production functions. What do the regression coefficients b_j represent?
d) Study the validity of the Frisch-Waugh-Lovell theorem on the case of explanatory variable labour in the model of linear production function.

Computer printout of the results in Stata:

a) Data exploration

```
. describe

Contains data from production.dta
  obs:            81
 vars:             4
 size:           810
-----------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-----------------------------------------------------------------------------------
n               byte   %8.0g                  Observation
q               long   %8.0g                  Value added
l               byte   %8.0g                  Average number of employed workers
k               long   %8.0g                  Tangible and intangible assets
-----------------------------------------------------------------------------------
Sorted by:

. inspect q l k

q:  Value added in 1000 monetary units          Number of Observations
--------------------------------------      ------------------------------
                                             Total    Integers   Nonintegers
|  #                         Negative            -          -            -
|  #                         Zero                -          -            -
|  #                         Positive           81         81            -
|  #                                         -----      -----        -----
|  #                         Total              81         81            -
|  #   .   .   .   .         Missing             -
+----------------------                      -----
1514             2312943                        81
  (81 unique values)
```

```
l:  Average number of employed workers        Number of Observations
--------------------------------------     ------------------------------
                                           Total    Integers   Nonintegers
|  #                         Negative          -           -             -
|  #                         Zero              -           -             -
|  #                         Positive         81          81             -
|  #                                       -----       -----         -----
|  #                         Total            81          81             -
|  #   .    .    .    .      Missing           -
+----------------------                    -----
1                    58                       81
  (22 unique values)

k:  Tangible and intangible assets in 1,000     Number of Observations
------------------------------------------- ------------------------------
                                           Total    Integers   Nonintegers
|  #                         Negative          -           -             -
|  #                         Zero              -           -             -
|  #                         Positive         81          81             -
|  #                                       -----       -----         -----
|  #                         Total            81          81             -
|  #   .    .    .    .      Missing           -
+----------------------                    -----
38               344297                       81
  (81 unique values)
```

. sum

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           n |        81          41    23.52658          1         81
           q |        81    180519.3    388072.4       1514    2312943
           l |        81    10.39506    14.81442          1         58
           k |        81    40226.93    71409.35         38     344297
```

. list q l k

```
     +-----------------------+
     |       q    l        k |
     |-----------------------|
  1. |    1514    1       43 |
  2. |    2106    1      266 |
  3. |    2758    1      639 |
  4. |    4117    1       38 |
  5. |    4243    1      656 |
     |-----------------------|
  6. |    4553    1     1597 |
  7. |    4649    1      287 |
  8. |    4963    1      560 |
  9. |    5749    1      962 |
 10. |    5807    2     2499 |
     |-----------------------|
 11. |    6267    1      418 |
 12. |    6353    2      855 |

...

 74. |  544607   48   344297 |
 75. |  580309   30    70084 |
     |-----------------------|
 76. |  798583   40   187301 |
 77. |  864224   25   124547 |
 78. |  965788   51   161713 |
 79. | 1.5e+06   51   117635 |
 80. | 1.7e+06   53   175943 |
     |-----------------------|
 81. | 2.3e+06   31   313144 |
     +-----------------------+
```

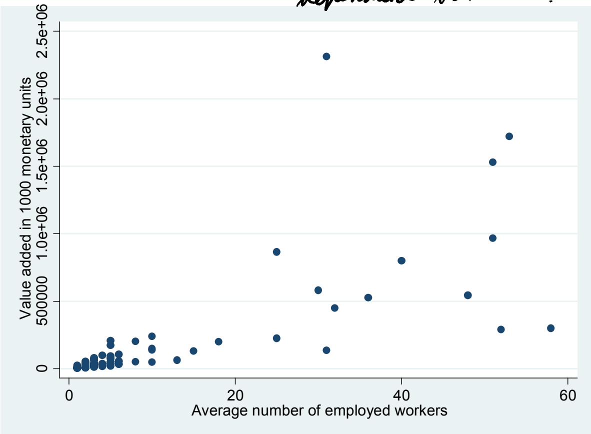
```
. sort q
```

it is the natural logarithm (ln)

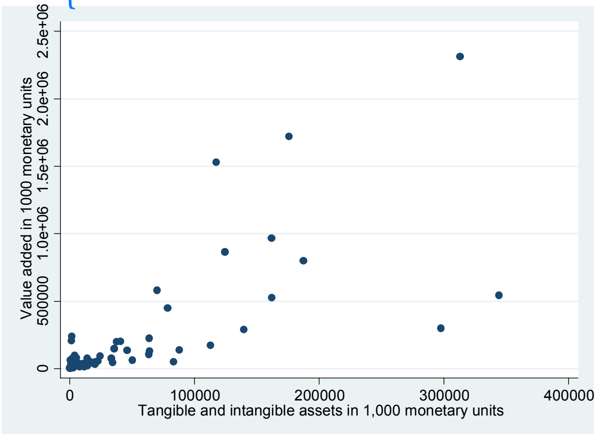
```
. gen lq=log(q)
. gen ll=log(l)
. gen lk=log(k)
```

before running the model one should look at relationship between explanatory variables and dependant variable.

```
. twoway scatter q l
```

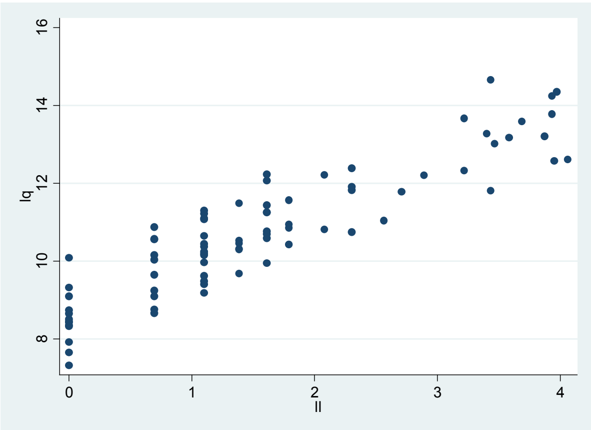



```
. twoway scatter q k
```



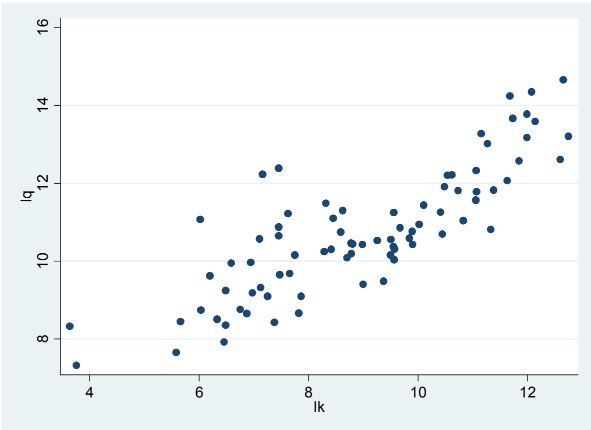

```
. twoway scatter lq ll
```

log of Q (y) – log of labour




```
. twoway scatter lq lk
```

log of capital



b) Estimation of production functions

```
. regress q l k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.6e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   9687.383   3640.852     2.66   0.009     2439.003    16935.76
           k |    2.27941   .7553228     3.02   0.003      .775678    3.783142
       _cons |  -11875.29   34865.13    -0.34   0.734    -81286.43    57535.85
------------------------------------------------------------------------------

. gen cons=1
```

```
. mkmat cons l k, matrix(xlin)
. matrix list xlin

xlin[81,3]
       cons        l        k
 r1       1        1       43
 r2       1        1      266

...

r81       1       31   313144

. matrix xxlin=(xlin)'*xlin
. matrix list xxlin

symmetric xxlin[3,3]
            cons            l            k
cons          81
   l         842        26310
   k     3258381    1.056e+08    5.390e+11

. mkmat q, matrix(ylin)
. matrix list ylin

ylin[81,1]
            q
 r1      1514
 r2      2106

...

r81   2312943

. matrix xylin=(xlin)'*ylin
. matrix list xylin

xylin[3,1]
               q
cons    14622065
   l   4.856e+08
   k   2.213e+12

. matrix xxlininv=invsym(xxlin)
. matrix list xxlininv

symmetric xxlininv[3,3]
             cons            l            k
cons    .01854393
   l   -.00067176    .00020222
   k    1.951e-08   -3.556e-08    8.703e-12

. matrix blin=xxlininv*xylin
. matrix list blin

blin[3,1]
               q
cons   -11875.29
   l   9687.3835
   k   2.2794101
```

– stata does not show you values above the diagonal because they are the same as below the line

Calculation by using matrix algebra:

$b=(X^TX)^{-1}X^Ty$ PDF, p.4.

→ model specification

```
. regress lq ll lk

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =  190.75
       Model |  178.261263     2  89.1306313           Prob > F      =  0.0000
    Residual |    36.44752    78  .467275898           R-squared     =  0.8302
-------------+------------------------------           Adj R-squared =  0.8259
       Total |  214.708783    80  2.68385978           Root MSE      =  .68358

------------------------------------------------------------------------------
          lq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ll |   .9645479   .1199229     8.04   0.000     .7257997    1.203296
          lk |   .1885438   .0673358     2.80   0.006     .0544886     .322599
       _cons |   7.546026   .4617465    16.34   0.000      6.62676    8.465293
------------------------------------------------------------------------------

. mkmat cons ll lk, matrix(xlog)
. matrix list xlog

xlog[81,3]
         cons          ll          lk
 r1         1           0   3.7612002
 r2         1           0   5.5834961

...

r81         1   3.4339871   12.654418

. matrix xxlog=(xlog)'*xlog
. matrix list xxlog

symmetric xxlog[3,3]
           cons          ll          lk
cons         81
  ll  127.86665   314.02997
  lk  725.93231   1314.3448   6861.7146

. mkmat lq, matrix(ylog)
. matrix list ylog

ylog[81,1]
            lq
 r1  7.3225102
 r2  7.6525459

...

r81  14.654032

. matrix xylog=(xlog)'*ylog
. matrix list xylog

xylog[3,1]
             lq
cons  871.43169
  ll  1515.5936
  lk  8039.3867

. matrix xxloginv=invsym(xxlog)
. matrix list xxloginv

symmetric xxloginv[3,3]
            cons           ll           lk
cons   .45628254
  ll   .08194953    .03077734
  lk  -.06396946   -.01456514    .00970328
```

Example, c) PDF, pp. 6-7.

$L_i = c_1 + c_2 K_i + e_{Li}$

$\hat{Q}_i = d_1 + d_2 e_{Li} + d_3 K_i$

compare d_2 should be equal to b_2

$\hat{Q}_i = b_1 + b_2 L_i + b_3 K_i$

$d_2 = b_2$

```
. matrix blog=xxloginv*xylog
. matrix list blog

blog[3,1]
              lq
cons   7.5460264
  ll   .96454793
  lk   .18854381
```

slide 78

c) Direct, indirect and total effects

```
. regress l k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  1,    79) =  201.49
       Model |  12612.2633     1  12612.2633           Prob > F      =  0.0000
    Residual |  4945.09471    79  62.5961355           R-squared     =  0.7183
-------------+------------------------------           Adj R-squared =  0.7148
       Total |   17557.358    80  219.466975           Root MSE      =  7.9118

------------------------------------------------------------------------------
           l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           k |   .0001758   .0000124    14.19   0.000     .0001512    .0002005
       _cons |   3.321905   1.010492     3.29   0.002     1.310571    5.333239
------------------------------------------------------------------------------

. predict resl, res
```

e

```
. regress q resl k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.6e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        resl |   9687.383   3640.852     2.66   0.009     2439.003    16935.76
           k |   3.982756   .4008578     9.94   0.000      3.18471    4.780803
       _cons |   20305.28   32700.13     0.62   0.536    -44795.69    85406.25
------------------------------------------------------------------------------

. regress q l k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.6e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   9687.383   3640.852     2.66   0.009     2439.003    16935.76
           k |    2.27941   .7553228     3.02   0.003      .775678    3.783142
       _cons |  -11875.29   34865.13    -0.34   0.734    -81286.43    57535.85
------------------------------------------------------------------------------
```

same value

```
. regress k l

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  1,    79) =  201.49
       Model |  2.9304e+11     1  2.9304e+11           Prob > F      =  0.0000
    Residual |  1.1490e+11    79  1.4544e+09           R-squared     =  0.7183
-------------+------------------------------           Adj R-squared =  0.7148
       Total |  4.0794e+11    80  5.0993e+09           Root MSE      =   38137

------------------------------------------------------------------------------
           k |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   4085.427   287.8158    14.19   0.000     3512.544     4658.31
       _cons |  -2241.338   5187.195    -0.43   0.667    -12566.19    8083.516
------------------------------------------------------------------------------

. predict resk, res

. regress q l resk

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.6e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   18999.75   1932.239     9.83   0.000     15152.95    22846.54
        resk |    2.27941   .7553228     3.02   0.003     .7756781    3.783142
       _cons |  -16984.22      34824    -0.49   0.627    -86313.49    52345.05
------------------------------------------------------------------------------
```

same value

```
. regress q l k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.6e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   9687.383   3640.852     2.66   0.009     2439.003    16935.76
           k |    2.27941   .7553228     3.02   0.003      .775678    3.783142
       _cons |  -11875.29   34865.13    -0.34   0.734    -81286.43    57535.85
------------------------------------------------------------------------------
```

slide 79

d) Frisch-Waugh-Lovell theorem

```
. regress q l k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.6e+05
```

Example, d) PDF, pp. 7-8.

$Q_i = b_1 + b_2 L_i + \underline{b_3} K_i + \underline{e_{1i}}$

$Q_i = c_1 + c_2 L_i + e_{Qi}$

$K_i = d_1 + d_2 L_i + e_{Ki}$

without the effect of labour

FWL regression: $e_{Qi} = g_1 + \underline{g_2} e_{Ki} + \underline{e_{2i}}$

FWL theorem: (1) $b_3 = g_2$

(2) $e_{1i} = e_{2i}, \forall i$

```
------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   9687.383   3640.852     2.66   0.009     2439.003    16935.76
           k |    2.27941   .7553228     3.02   0.003      .775678    3.783142
       _cons |  -11875.29   34865.13    -0.34   0.734    -81286.43    57535.85
------------------------------------------------------------------------------
```

. predict resid1, res

. regress q l

```
      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  1,    79) =   87.69
       Model |  6.3380e+12     1  6.3380e+12           Prob > F      =  0.0000
    Residual |  5.7100e+12    79  7.2278e+10           R-squared     =  0.5261
-------------+------------------------------           Adj R-squared =  0.5201
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      =  2.7e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   18999.75   2028.962     9.36   0.000      14961.2     23038.3
       _cons |  -16984.22   36567.22    -0.46   0.644    -89769.45    55801.01
------------------------------------------------------------------------------
```

. predict resqfw, res

. regress k l

```
      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  1,    79) =  201.49
       Model |  2.9304e+11     1  2.9304e+11           Prob > F      =  0.0000
    Residual |  1.1490e+11    79  1.4544e+09           R-squared     =  0.7183
-------------+------------------------------           Adj R-squared =  0.7148
       Total |  4.0794e+11    80  5.0993e+09           Root MSE      =   38137

------------------------------------------------------------------------------
           k |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   4085.427   287.8158    14.19   0.000     3512.544     4658.31
       _cons |  -2241.338   5187.195    -0.43   0.667    -12566.19    8083.516
------------------------------------------------------------------------------
```

. predict reskfw, res

. regress resqfw reskfw

```
      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  1,    79) =    9.22
       Model |  5.9698e+11     1  5.9698e+11           Prob > F      =  0.0032
    Residual |  5.1130e+12    79  6.4721e+10           R-squared     =  0.1046
-------------+------------------------------           Adj R-squared =  0.0932
       Total |  5.7100e+12    80  7.1375e+10           Root MSE      =  2.5e+05

------------------------------------------------------------------------------
      resqfw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      reskfw |    2.27941    .750527     3.04   0.003     .7855237    3.773297
       _cons |   .0005777   28267.12     0.00   1.000     -56264.3     56264.3
------------------------------------------------------------------------------
```

. predict resid2, res

```
. list resid1 resid2

     +-----------------------+
     |    resid1      resid2 |
     |-----------------------|
  1. |  3603.892    3603.892 |
  2. |  3687.584    3687.583 |
  3. |  3489.364    3489.364 |
  4. |   6218.29    6218.289 |
  5. |  4935.614    4935.614 |
     |-----------------------|
  6. |  3100.689    3100.688 |
  7. |  6182.716    6182.716 |
  8. |  5874.438    5874.437 |
  9. |  5744.114    5744.114 |
 10. | -7388.722   -7388.722 |
     |-----------------------|
 11. |  7502.114    7502.113 |
 12. | -3095.372   -3095.372 |
 13. |  -4511.34    -4511.34 |
 14. |  7905.542    7905.541 |
 15. | -9865.388   -9865.387 |
     |-----------------------|
 16. |   1406.51     1406.51 |
 17. |  10504.76    10504.76 |
 18. | -23468.89   -23468.89 |
 19. |  -30861.5    -30861.5 |
 20. | -3090.771    -3090.77 |
     |-----------------------|
 21. |  4126.923    4126.924 |
 22. | -15702.03   -15702.03 |

...

 73. | -180972.1   -180972.1 |
 74. | -693306.2   -693306.2 |
 75. |  141812.6    141812.6 |
     |-----------------------|
 76. | -3972.844    -3972.85 |
 77. |    350021      350021 |
 78. |  114996.5    114996.5 |
 79. |  778072.3    778072.3 |
 80. |  816731.7    816731.7 |
     |-----------------------|
 81. |   1310726     1310726 |
     |-----------------------|
Mean | -.0010632   -.0011936 |
     +-----------------------+
```

■

Model 1

Estimation by regress and interpretation

1) Model statistics: **PDF, pp. 3–4.**

n

F } Statistical sign. of the model as a whole

$p(F)$

R^2

$\bar{R}^2$

$s_e \rightarrow$ ROOT MSE in Stata

2) ANOVA table:

$\rightarrow$ SS in Stata

mean squares: SS/df

ESS	df: $k-1$	EMS
RSS	df: $n-k$	RMS $= s_e^2$
TSS	df: $n-1$	TMS

3) Parameter estimates:

b_j, $se(b_j)$, $\underbrace{t_j, \; p_j}$, 95% conf. interval

$H_0: \beta_j = 0$

$H_1: \beta_j \neq 0$

"e+05" means "$\cdot 10^5$", "e−05" means "$\cdot 10^{-5}$".

• $b_2 = 9687.4$

$p(b_2) = 0.009 < \alpha = 0.05$, we reject $H_0: \beta_2 = 0$ and conclude that $\beta_2 \neq 0$.

The value is non-random / stat. significant.

Interpretation:

Regression coefficient

!!

If x_j increases by 1 unit of measurement, then on average, ceteris paribus, y increases / decreases by b_j units of measurement.

If the average number of employees increases by 1 employee, then on average, ceteris paribus, value added increases (+) by 9.7 mill. m.u. (it's in 1000 m.u.)

• $b_3 = 2.279$

$p(b_3) = 0.003 < \alpha = 0.05$

If the sum of tangible and intangible assets increases by 1000 m.u., then on average, ceteris paribus, value added increases (+) by 2279 m.u.

Computer Printout of the Results

Computer printout of econometric estimation of a linear production function (Stata)

$e+12 \Rightarrow \cdot 10^{12}$

analysis of variance ANOVA

model diagnostics

```
. regress q l k

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =   52.90
       Model |  6.9350e+12     2  3.4675e+12           Prob > F      =  0.0000
    Residual |  5.1130e+12    78  6.5551e+10           R-squared     =  0.5756
-------------+------------------------------           Adj R-squared =  0.5647
       Total |  1.2048e+13    80  1.5060e+11           Root MSE      = 2.6e+05

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           l |   9687.383   3640.852     2.66   0.009     2439.003    16935.76
           k |    2.27941   .7553228     3.02   0.003      .775678    3.783142
       _cons |  -11875.29   34865.13    -0.34   0.734    -81286.43    57535.85
------------------------------------------------------------------------------
```

ESS

RSS

TSS

degrees of freedom

MS mean squares = Values in SS / Degrees of freedom

$\sqrt{6{,}5551 \cdot 10^{10}}$ 12PIT

regression coefficient estimates

b_j, $se(b_j)$, $\underbrace{t_j, \; p_j}$, 95% conf. interval

$H_0: \beta_j = 0$
$H_1: \beta_j \neq 0$

- $b_2 = 9687.4$

$p(b_2) = 0.009 < \alpha = 0.05$, we reject $H_0: \beta_2 = 0$ and conclude that $\beta_2 \neq 0$.

The value is non-random / stat. significant.

Interpretation:

Regression coefficient

> If x_j increases by 1 unit of measurement, then on average, ceteris paribus, y increases / decreases by b_j units of measurement.

12PIT

If the average number of employees increases by 1 employee, then on average, ceteris paribus, value added increases (+) by 9.7 mill. m.u.
(it's in 1000 m.u.)

Computer Printout of the Results

Computer printout of econometric estimation of a linear production function (R)

```
> mod_lin = lm(q ~ l + k, data = production)
> summary(mod_lin)

Call:
lm(formula = q ~ l + k, data = production)

Residuals:
Value added in 1000 monetary units
    Min      1Q  Median      3Q     Max
-928125  -30862   -3095    7945 1310726

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.188e+04  3.487e+04  -0.341  0.73432
l            9.687e+03  3.641e+03   2.661  0.00946 **
k            2.279e+00  7.553e-01   3.018  0.00344 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 256000 on 78 degrees of freedom
Multiple R-squared:  0.5756,   Adjusted R-squared:  0.5647
F-statistic:  52.9 on 2 and 78 DF,  p-value: 3.038e-15
```

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS AMBA ACCREDITED

Computer Printout of the Results

```
> summary.aov(mod_lin)
            Df    Sum Sq   Mean Sq F value   Pr(>F)    
l            1 6.338e+12 6.338e+12  96.688 2.64e-15 ***
k            1 5.970e+11 5.970e+11   9.107  0.00344 ** 
Residuals   78 5.113e+12 6.555e+10                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> sum(anova(mod_lin)[-3,2])
[1] 6.935018e+12

> sum(anova(mod_lin)[,2])
[1] 1.204801e+13

> nobs(mod_lin)
[1] 81

> confint(mod_lin, level=0.95)
                    2.5 %      97.5 %
(Intercept) -81286.433705 57535.852836
l             2439.003091 16935.763850
k                0.775678     3.783142
```

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Computer Printout of the Results

Computer printout of econometric estimation of a linearized Cobb-Douglas production function (Stata)

```
. regress lq ll lk

      Source |       SS       df       MS              Number of obs =      81
-------------+------------------------------           F(  2,    78) =  190.75
       Model |  178.261263     2  89.1306313           Prob > F      =  0.0000
    Residual |    36.44752    78  .467275898           R-squared     =  0.8302
-------------+------------------------------           Adj R-squared =  0.8259
       Total |  214.708783    80  2.68385978           Root MSE      =  .68358

------------------------------------------------------------------------------
          lq |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ll |   .9645479   .1199229     8.04   0.000     .7257997    1.203296
          lk |   .1885438   .0673358     2.80   0.006     .0544886     .322599
       _cons |   7.546026   .4617465    16.34   0.000      6.62676    8.465293
------------------------------------------------------------------------------
```

In this case the interpretation of the regression coefficient needs to be interpreted as elasticity

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Model 2

PDF, pp.5-6.

Estimation by regress and interpretation

- $b_2 = 0.96$

  $p(b_2) = 0.000 < \alpha = 0.05$

  Interpretation:

  Elasticity

  $\boxed{\text{If } x_j \text{ increases by 1 percent, then on average, ceteris paribus, } y \text{ increases/decreases by } b_j \text{ percent.}}$

  If the average number of employees increases by 1 percent, then on average, ceteris paribus, value added increases (+) by 0.96 percent.

- $b_3 = 0.19$

  $p(b_3) = 0.006 < \alpha = 0.05$

If the sum of tangible and intangible assets increases by 1 percent, then on average, ceteris paribus, value added increases (+) by 0.19 percent.

Computer Printout of the Results

Computer printout of econometric estimation of a linearized Cobb-Douglas production function (R)

```
> mod_log = lm(lq ~ ll + lk, data = production)
> summary(mod_log)

Call:
lm(formula = lq ~ ll + lk, data = production)

Residuals:
Value added in 1000 monetary units
     Min       1Q   Median       3Q      Max
-1.22501 -0.46545 -0.08825  0.42991  1.78679

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.54603    0.46175  16.342  < 2e-16 ***
ll           0.96455    0.11992   8.043 7.77e-12 ***
lk           0.18854    0.06734   2.800  0.00644 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6836 on 78 degrees of freedom
Multiple R-squared:  0.8302,    Adjusted R-squared:  0.8259
F-statistic: 190.7 on 2 and 78 DF,  p-value: < 2.2e-16
```

Computer Printout of the Results

```
> summary.aov(mod_log)
            Df Sum Sq Mean Sq F value   Pr(>F)    
ll           1 174.60  174.60  373.65 < 2e-16 ***
lk           1   3.66    3.66    7.84 0.00644 ** 
Residuals   78  36.45    0.47                    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> sum(anova(mod_log)[-3,2])
[1] 178.2613

> sum(anova(mod_log)[,2])
[1] 214.7088

> nobs(mod_log)
[1] 81

> confint(mod_log, level=0.95)
                2.5 %    97.5 %
(Intercept) 6.6267597 8.4652928
ll          0.7257997 1.2032961
lk          0.0544886 0.3225991
```

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Interpretation of regression coefficients for most common forms of regression models

1. Linear regr. model

error term

$$y_i = b_1 + b_2 x_{2i} + \dots + b_j x_{ji} + \dots + b_k x_{ki} + e_i$$

residual

partial regr. coefficient

on average must be included

↳ If x_j increases by 1 unit of measurement, then on average, ceteris paribus, y increases / decreases by b_j units of measurement.

Two relativisations, intercept: S 75.

2. Log-linear regression model

$$\ln y_i = b_1 + b_2 \ln x_{2i} + \dots + b_j \ln x_{ji} + \dots + b_k \ln x_{ki} + e_i$$

part. elasticity (approximation) ⇒ it relativizes the units of measurement

If x_j increases by 1 percent, then on average, ceteris paribus, y increases / decreases by b_j percent.

3. Log-lin regression model

$$\ln y_i = b_1 + b_2 x_{2i} + \dots + b_j x_{ji} + \dots + b_k x_{ki} + e_i$$

↳ the left side always has ln

part. semi elasticity (approximation)

If x_j increases by 1 unit of measurement, then on average, ceteris paribus, y increases/decreases by $100 \cdot b_j$ percent.

4. Lin-log regression model

most widely used

$$y_i = b_1 + b_2 \ln x_{2i} + \dots + b_j \ln x_{ji} + \dots + b_k \ln x_{ki} + e_i$$

part. semi elasticity (approximation)

If x_j increases by 1 percent, then on average, ceteris paribus, y increases/decreases by $b_j/100$ units of measurement.

model that doesn't fit any of the 4 groups

5. "Mixed" regression model

EXAMPLE: $\ln y_i = b_1 + b_2 x_{2i} + b_3 \ln x_{3i} + e_i$

log-lin r.m.

log-linear r.m.

Each regression coefficient is interpreted based on the functional form of the relationship between "y" and "x_j".

Indirect effects, S77-78:

Indirect effect of x_j on y (through x_l):
total effect of x_j on x_l $\cdot$ direct effect of x_l on y.

Example, c)

PDF, pp. 6-7.

$$L_i = c_1 + c_2 K_i + e_{Li}$$

$$\hat{Q}_i = d_1 + \underline{d_2} e_{Li} + d_3 K_i$$

compare

$$\hat{Q}_i = b_1 + \underline{b_2} L_i + b_3 K_i$$

$$d_2 = b_2$$

Example, d)

PDF, pp. 7-8.

$$Q_i = b_1 + b_2 L_i + \underline{b_3} K_i + \underline{e_{1i}}$$

$$Q_i = c_1 + c_2 L_i + e_{Qi}$$
$$K_i = d_1 + d_2 L_i + e_{Ki}$$

FWL regression: $e_{Qi} = g_1 + \underline{g_2} e_{Ki} + \underline{e_{2i}}$

FWL theorem:
1. $b_3 = g_2$
2. $e_{1i} = e_{2i}, \forall i$

Definition of a Regression Coefficient

$$E(y|x_2,\ldots,x_k) = \beta_1 + \beta_2 x_2 \ldots + \beta_k x_k$$

$$\beta_j = \frac{\partial E(y|x_2,\ldots,x_k)}{\partial x_j} \quad ; \quad j = 2,3,\ldots,k$$

Multiple regression coefficients

Partial regression coefficients or partial slope coefficients

value of intercept

$$\beta_1 = E(y|x_2 = 0,\ldots,x_k = 0)$$

Definition of a Regression Coefficient

12PIT

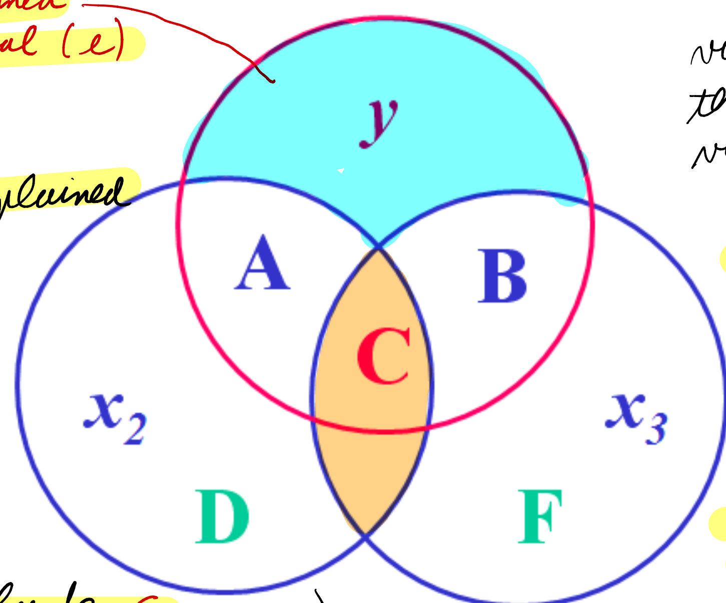
Regression coefficients represent

pure or direct or net effects ⇒ the same expression

$$\hat{y}_i = b_1 + b_2 x_{2i} + b_3 x_{3i}$$

This is unexplained
It is the residual (e)

variability of the dependent variable



A can be explained by X_2

$$\hat{x}_{2i} = c_1 + c_2 x_{3i}$$

$$e_{2i} = x_{2i} - \hat{x}_{2i}$$

B can be explained by x_3

$$\hat{x}_{3i} = f_1 + f_2 x_{2i}$$

$$e_{3i} = x_{3i} - \hat{x}_{3i}$$

C is something we try to avoid ⇒ we want to have no colinearity between the explanatory variables ⇒ in real cases there is always some colinearity present

e_2 doesn't include C

$$\hat{y}_i = d_1 + d_2 e_{2i} + d_3 x_{3i}$$

$$\hat{y}_i = g_1 + g_2 x_{2i} + g_3 e_{3i}$$

d_2 is numerically equivalent to b_2.

variability of explanatory variables

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

2.

The residual of x_3 is B plus F.

y_3 is the numerical equivalent of b_3.

x_2 affects $\hat{y}$ directly and indirectly through x_3 => that is because of collinearity

Indirect effects, S77-78:

Indirect effect of x_j on y (through x_l):
total effect of x_j on x_l $\odot$ direct effect of x_l on y.

Decomposition of the Total Effect

Effects of explanatory variable x_2

Direct effect	:	$b_2 = d_2$
Indirect effect (effect of x_2 on x_3 and through the latter on y)	:	$f_2 \cdot b_3$
Total effect	:	$g_2 = d_2 + f_2 \cdot b_3$

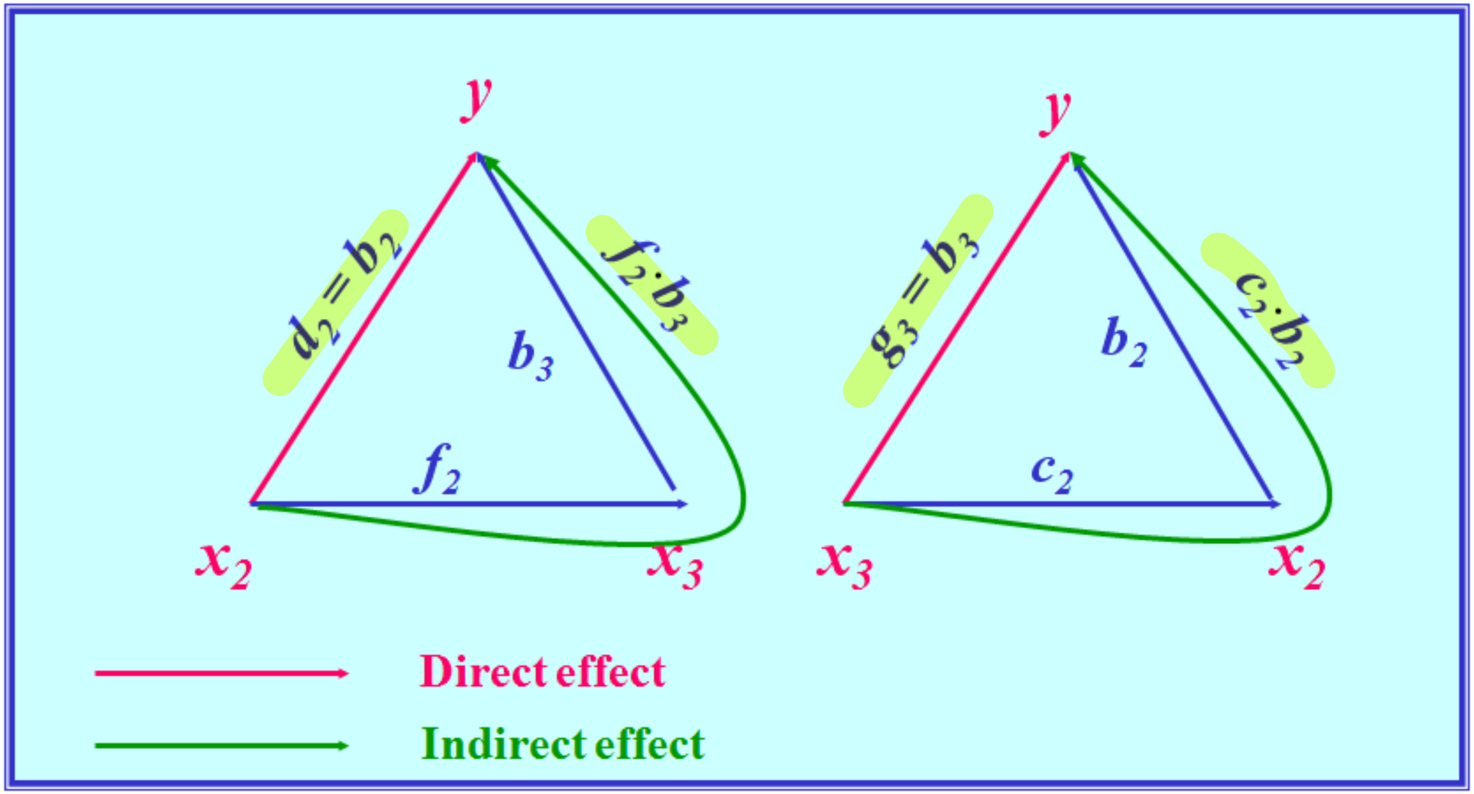
Effects of explanatory variable x_3

Direct effect	:	$b_3 = g_3$
Indirect effect (efect of x_3 on x_2 and through the latter on y)	:	$c_2 \cdot b_2$
Total effect	:	$d_3 = g_3 + c_2 \cdot b_2$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAs ACCREDITED

Decomposition of the Total Effect

Graphic illustration of direct and indirect effects of explanatory variables on the dependent variable



EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Frisch–Waugh (–Lovell) Theorem

Frisch–Waugh theorem (Econometrica, 1933)

you can get a certain direct effect and the residuals in 2 ways
- add all variables into the regression
- get rid of the effect of the variables

you can purge the effect of x_2 on y and x_3.

$$y_i = b_1 + b_2 x_{2i} + b_3 x_{3i} + e_i$$

$$y_i = c_1 + c_2 x_{2i} + e_{y_i} \quad \Rightarrow \quad \tilde{y}_i = e_{y_i} = y_i - \hat{y}_i$$

$$x_{3i} = d_1 + d_2 x_{2i} + e_{x_{3i}} \quad \Rightarrow \quad \tilde{x}_{3i} = e_{x_{3i}} = x_{3i} - \hat{x}_{3i}$$

$$\tilde{y}_i = g_1 + b_3 \tilde{x}_{3i} + e_i$$

FW(L) regression

this has different value than b_1

y without the effect of x_2

x_3 without the effect of x_2.

Frisch–Waugh (–Lovell) Theorem

AN EXAMPLE OF ITS USE:

Method of individual trend

First step

We purge the "redundant" component from the time series

Second step

We estimate the regression model that is cleaned off of the effects of "redundant" component

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Frisch–Waugh (–Lovell) Theorem

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \ldots + \beta_k x_{kt} + u_t$$

First step

$$y_t = c_1 + c_2 t + c_3 t^2 + \ldots + e_{y_t} \quad ; \quad e_{y_t} = y_t - \hat{y}_t$$

$$x_{2t} = c_{21} + c_{22} t + c_{23} t^2 + \ldots + e_{x_{2t}} \quad ; \quad e_{x_{2t}} = x_{2t} - \hat{x}_{2t}$$

$$\ldots$$

$$x_{kt} = c_{k1} + c_{k2} t + c_{k3} t^2 + \ldots + e_{x_{kt}} \quad ; \quad e_{x_{kt}} = x_{kt} - \hat{x}_{kt}$$

Second step

$$e_{y_t} = b_1 + b_2 e_{x_{2t}} + b_3 e_{x_{3t}} + \ldots + b_k e_{x_{kt}} + e_t$$

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION AMBA ACCREDITED

Frisch–Waugh (–Lovell) Theorem

AN ALTERNATIVE APPROACH:

easier approach

Method of partial time-series regression

$$y_t = d_1 + b_2 x_{2t} + b_3 x_{3t} + \ldots + b_k x_{kt} + \underline{d_2 t + d_3 t^2 + \ldots} + e_t$$

University of Ljubljana

SEB SCHOOL OF ECONOMICS AND BUSINESS

2. Multiple Regression Model

Prof. Dr. Miroslav Verbič

miroslav.verbic@ef.uni-lj.si

www.miroslav-verbic.si

EFMD EQUIS ACCREDITED

AACSB ACCREDITED

ASSOCIATION AMBA ACCREDITED

Ljubljana, October 2022