

ECONOMIC STATISTICS





NEW

SZÉCHENYI PLAN

ECONOMIC STATISTICS

Sponsored by a Grant TÁMOP-4.1.2-08/2/A/KMR-2009-0041

Course Material Developed by Department of Economics,

Faculty of Social Sciences, Eötvös Loránd University Budapest (ELTE)

Department of Economics, Eötvös Loránd University Budapest

Institute of Economics, Hungarian Academy of Sciences

Balassi Kiadó, Budapest



The project is supported
by the European Union.

National Development Agency
www.ujszechenyiterv.gov.hu
06 40 638 638



HUNGARY'S RENEWAL



The projects have been supported
by the European Union.

ELTE Faculty of Social Sciences, Department of Economics

ECONOMIC STATISTICS

Author: Anikó Bíró

Supervised by Anikó Bíró

June 2010

ECONOMIC STATISTICS

Week 9

Distributed lag models

Anikó Bíró

Discussion of 2nd exam

Time series analysis

- Cross sectional vs. time series data
- Difficulties in time series analysis:
 - Lagged effects – dynamic model
 - Not stationary variables – “spurious regression”
- Complicated with Excel – instead: EViews (basics: seminar)

Lagged variables

$Y_t, t = 1, \dots, T$ T observations

$X_t, t = 1, \dots, T$ T observations

$W_t = X_{t-1}, t = 2, \dots, T$ T - 1 observations

$Z_t = X_{t-j}, t = j + 1, \dots, T$ T - j observations

Regression of Y on W: T-1 observations

- OLS: like before (if stationary variables)
- Examples:
 - Interest rate cut – inflation
 - Participation on training – earnings

Example 1 – industrial safety training

safety.xls – data of a firm over 60 months

Y: loss due to accidents (GDP/month)

X: safety training (h/worker/month)

4 period distributed lag model:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + e_t$$

Lag in EViews: X(-1)

Estimation results

Included observations: 56 after adjustments

Variable	Coefficient	t-Statistic	Prob.
C	91173,32	46,759	0,0000
TRAINING	-131,99	-2,783	0,0076
TRAINING(-1)	-449,86	-9,459	0,0000
TRAINING(-2)	-422,52	-9,032	0,0000
TRAINING(-3)	-187,10	-3,927	0,0003
TRAINING(-4)	-27,77	-0,583	0,5627

Interpretation of coefficients:

- Effect of training after j months, ceteris paribus
- Sum: total effect
- U-shape

Example 2 – patents

1960-1993 U.S. yearly data

(source: Gretl-Ramanathan)

Y: number of patents (thousand)

X: R&D expenditures (billion USD)

- Are lagged regressors needed?
- Lag length (how many periods)?

Estimation results

Dependent Variable: PATENT

Method: Least Squares

Sample(adjusted): 1964 1993

Included observations: 30 after adjusting endpoints

Variable	Coefficient	Std. error	t-statistic	P-value
C	26,327	4,148	6,347	0,000
RD	-0,597	0,459	-1,298	0,207
RD(-1)	0,867	0,971	0,893	0,381
RD(-2)	0,013	1,098	0,012	0,991
RD(-3)	-0,640	0,995	-0,649	0,526
RD(-4)	1,347	0,494	2,727	0,012
R-squared	0,964			

Lag length selection

1. Choose longest lag which is still reasonable, q_{\max} .
2. Estimation, testing significance of q_{\max} . If significant: finished. If not significant: decrease lag length by one.
3. Repeat point 2 with $q_{\max} - 1$, $q_{\max} - 2$, ... lags.

Summary

Distributed lag models

- Model specification
- Interpretation of coefficients
- Lag length selection

EViews

Seminar 9

EViews and other softwares

- Statistical – econometric software
- User friendly, especially suitable for time series analysis
- Help files (User's guide)
- Stata:
 - More built-in procedures, easier to program
 - More suitable for cross-sectional and time series analysis
- Gretl:
 - Free
 - Weaknesses: panel and multivariate time series models
- Excel:
 - Not suitable for time series analysis
 - Difficult to use with large database

Loading data I.

- Example: educ.xls
 - File/new/workfile – annual
 - Objects/new object/series – edit
 - Copy – paste
 - Name

Loading data II.

- Example: safety.xls
 - File/new/workfile – undated
 - Procs/Import/Read text-lotus-excel
 - Source file should be closed!
 - Excel sheet name: safety
 - Upper-left data cell: pl. A2
 - Names for series: pl. loss train – reads both variables

Data manipulation

- Open, descriptive statistics, graphs
 - View/Descriptive statistics
 - View/Graph
 - Select several variables jointly: open as group
- Generate variables (genr)
- Sample: `smpl`
 - `smpl 1 20`
 - `smpl @all`
 - Or: `quick/sample`

Regression

- Quick/estimate equation ...
 - Include constant!
- Or:
equation name.ls ...
- Output formats, view options

Practice

Import Educ.xls file to EViews

Y: GDP annual growth rate

X: expenditure on schooling/number of children below age 16

- Descriptive statistics of GDP growth
- Estimate distributed lag model with 5 lags (lag: $X(-1)$)
- Lag length selection (assume: max. 10 years)

Lag length selection

1. Choose longest lag which is still reasonable, q_{\max} .
2. Estimation, testing significance of q_{\max} . If significant: finished. If not significant: decrease lag length by one.
3. Repeat point 2 with $q_{\max} - 1$, $q_{\max} - 2$, ... lags.

Practicing examples

8.1, 8.2 – with EViews!

Safety.xls file – 6 months max. lag length

Homework 5 (individual)

Estimate a cross sectional model (which has economic meaning) with EViews

- Describe briefly the model
- Descriptive statistics and histogram of the dependent variable
- Correlation between two selected variables
- Estimate the model, interpret the results