ECONOMETRICS

ECONOMETRICS

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

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

ECONOMETRICS

Authors: Péter Elek, Anikó Bíró

Supervised by: Péter Elek

June 2010

ECONOMETRICS

Week 12.

Time series regressions I

Péter Elek, Anikó Bíró

Plan

Regression on stationary time series

Consequences of autocorrelated error terms

Testing autocorrelation

Handling autocorrelation

Textbook: M 6.1–6.5., 6.8.

Reminder: cross sectional regression with stochastic regressors

- Fixed regressors are less sensible in case of time series
- Model with stochastic variables: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots \beta_k x_{ik} + u_i$
- Conditions for unbiasedness of OLS
  - $(y_i, x_{i1}, x_{i2}, \ldots, x_{ik})$ $(i = 1,..,n)$ random sample of the model
  - $E(u|x_1, x_2, \ldots, x_k) = 0$
  - No perfect collinearity
- If homoscedasticity is also assumed, the following statements are true
  - The usual formula of variance is valid, the OLS estimator is asymptotically normal
  - OLS is BLUE.
- If normality of error term is also assumed then small sample tests t-test, F-test) are also valid.

Regression with stationary variables

$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \ldots \beta_k x_{tk} + u_t$

If x_{ti} $(i = 1,\ldots,k)$ and y_t are stationary then sufficient conditions for consistency of OLS:

- $E(u_t | x_{t1}, x_{t2}, \ldots, x_{tk}) = 0$ and
- No perfect collinearity

Further assumptions are needed for asymptotic validity of the usual tests (validity of formulas of variance etc.):

- Homoscedasticity and
- No autocorrelation in the error terms:
- $E(u_t u_s | x_{t1}, \ldots, x_{tk}, x_{s1}, \ldots, x_{sk}) = 0 \ (t \neq s)$

Regression with stationary variables (cont.)

The same is true in case of trend stationarity, but trend has to be included as regressor

Some x_{ti} might be lagged y_t (but the exogeneity condition has to hold!)

E.g. stationary AR(1) model: $k = 1$, $x_{t1} = y_{t-1}$

- OLS estimation of the coefficient is consistent, asymptotically normal (but not unbiased!)

Autocorrelation of the error terms in the stationary regression

If the error terms are autocorrelated in a stationary regression then

- OLS is still consistent,
- But not BLUE,
- And the common formula of variance and the usual tests are not valid!

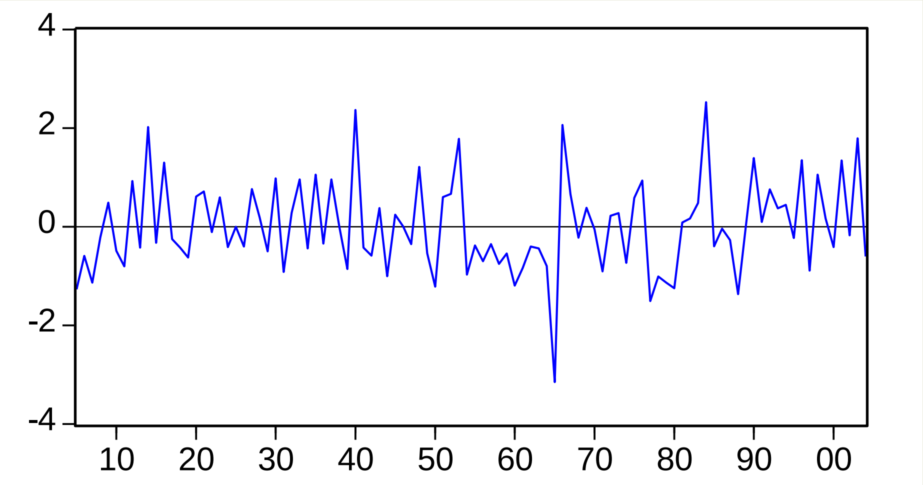
Size of bias in variance:
M page 285–286.

Testing autocorrelation of error terms

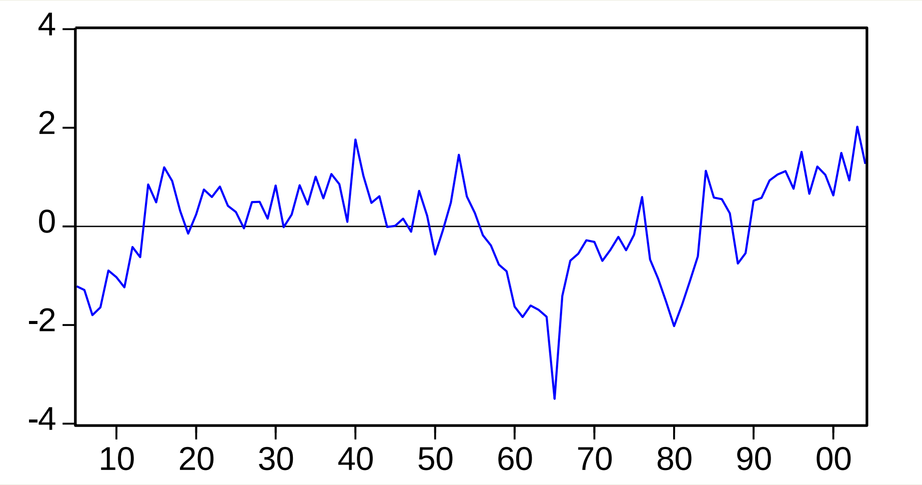
Durbin–Watson-test

Breusch–Godfrey-test

white noise error term



autocorrelated error term



Durbin–Watson-test

Analyzes the residuals of OLS regression:

$$d = \frac{\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n}\hat{u}_t^2} = \frac{\sum_{t=2}^{n}\hat{u}_t^2 - 2\sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1} + \sum_{t=2}^{n}\hat{u}_{t-1}}{\sum_{t=1}^{n}\hat{u}_t^2} \approx 2 - 2\frac{\sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=1}^{n}\hat{u}_t^2}$$

Estimator of first order autocorrelation:

$$\hat{\rho} = \frac{\sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=1}^{n}\hat{u}_t^2}$$

$$d \approx 2(1-\hat{\rho})$$

$0 \le d \le 4$ $(-1 \le \rho \le 1)$

$d = 2$ $\Leftrightarrow$ $\rho = 0$ (white noise)

$0 < d < 2$ $\Leftrightarrow$ $\rho > 0$ (positive autocorrelation)

$2 < d < 4$ $\Leftrightarrow$ $\rho < 0$ (negative autocorrelation)

DW-test, cont.

$H_0: \rho = 0$, $H_1: \rho > 0$ (one sided test!)

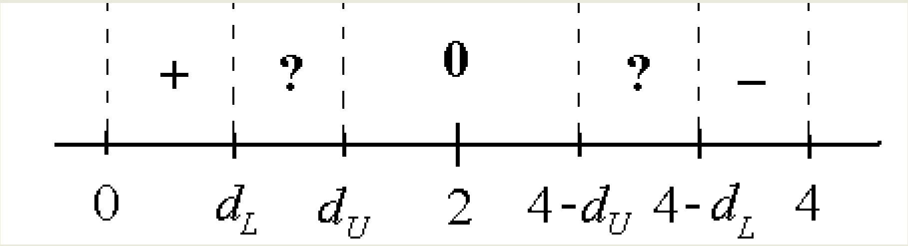
The test has two critical values because the distribution of the test statistic depends on the regressors: d_L (lower value), d_U (upper value)

Decision rule:

- Accept H_0 if $d > d_U$
- Reject H_0 if $d < d_L$
- Cannot decide if $d_L < d < d_U$ (neutral, grey zone)

Testing negative autocorrelation:

- Use $4 - d$ instead of d, otherwise everything is the same
- Accept H_0 if $4 - d > d_U$
- Reject H_0 if $4 - d < d_L$
- Cannot decide if $d_L < 4 - d < d_U$ (neutral, grey zone)



Limitations of DW-test

Can be used only for AR(1) residuals

In some cases ($d_L<d<d_U$) the test is not conclusive

Cannot be used for distributed lag models (in some cases, see later)

Breusch–Godfrey-test

AR(p) model of error terms:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots + \rho_q u_{t-p} + e_t$$

$H_0: \rho_1 = \rho_2 = \ldots = \rho_p = 0$

Regress the estimated error term on the explanatory variables and p lags of the error term

Under H_0, asymptotically $nR^2 \sim \chi_p^2$

Handling autocorrelation

- Adjusting the standard error of OLS estimation: Newey-West
  - Just as the White-procedure adjusts the standard error of OLS in case of heteroscedasticity
- Generalized Least Squares (GLS) types of estimations, e.g. Cochrane-Orcutt procedure

Cochrane–Orcutt-procedure

Model

$y_t = \beta_0 + \beta_1 x_{1t} + \beta_1 x_{2t} + \beta_1 x_{kt} + u_t$

$u_t = \rho u_{t-1} + e_t$, $e_t \sim IN$

Quasi differencing

$(y_t - \rho y_{t-1}) = (1 - \rho)\beta_0 + \beta_1(x_{1t} - \rho x_{1,t-1}) + \ldots + \beta_k(x_{kt} - \rho x_{k,t-1}) + e_t$

Regress $y_t - \rho y_{t-1}$ on $x_{it} - \rho x_{i,t-1}$ variables

Procedure

- OLS estimation, then estimation of ρ based on the residuals
- OLS of the quasi-differenced time series
- Iteration can also be used (does not improve asymptotically the efficiency)

Since ρ is estimated, if ρ is close to 0 then not necessarily better than OLS

Example: estimation of static Phillips-curve (USA) with OLS
Considerable autocorrelation

Dependent Variable: INF
Method: Least Squares
Date: 02/24/09 Time: 12:23
Sample: 1948 2003
Included observations: 56

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.053565	1.547957	0.680617	0.4990
UNEM	0.502378	0.265562	1.891752	0.0639

R-squared	0.062154	Mean dependent var	3.883929
Adjusted R-squared	0.044786	S.D. dependent var	3.040381
S.E. of regression	2.971518	Akaike info criterion	5.051084
Sum squared resid	476.8157	Schwarz criterion	5.123418
Log likelihood	-139.4304	Hannan-Quinn criter.	5.079128
F-statistic	3.578726	Durbin-Watson stat	0.801482
Prob(F-statistic)	0.063892		

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	31.52897	Prob. F(1,53)	0.0000
Obs*R-squared	20.88778	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 02/24/09 Time: 12:26
Sample: 1948 2003
Included observations: 56
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.555356	1.318280	1.938402	0.0579
UNEM	-0.457243	0.227341	-2.011267	0.0494
RESID(-1)	0.656330	0.116887	5.615066	0.0000

R-squared	0.372996	Mean dependent var	5.43E-16
Adjusted R-squared	0.349335	S.D. dependent var	2.944380
S.E. of regression	2.375048	Akaike info criterion	4.619996
Sum squared resid	298.9653	Schwarz criterion	4.728497
Log likelihood	-126.3599	Hannan-Quinn criter.	4.662061
F-statistic	15.76448	Durbin-Watson stat	1.808715
Prob(F-statistic)	0.000004		

Example, cont.: correcting autocorrelation

Strong autocorrelation, the regression is basically on the differences! These estimates are more reliable.

Dependent Variable: INF
Method: Least Squares
Date: 02/24/09 Time: 13:20
Sample (adjusted): 1949 2003
Included observations: 55 after adjustments
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.287078	2.287763	3.185242	0.0024
UNEM	-0.663959	0.330235	-2.010564	0.0496
AR(1)	0.782009	0.094683	8.259202	0.0000

R-squared	0.503681	Mean dependent var	3.807273
Adjusted R-squared	0.484592	S.D. dependent var	3.013295
S.E. of regression	2.163302	Akaike info criterion	4.434150
Sum squared resid	243.3536	Schwarz criterion	4.543641
Log likelihood	-118.9391	Hannan-Quinn criter.	4.476491
F-statistic	26.38569	Durbin-Watson stat	1.600203
Prob(F-statistic)	0.000000		

Inverted AR Roots	.78

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.734056	Prob. F(1,51)	0.3956
Obs*R-squared	0.780397	Prob. Chi-Square(1)	0.3770

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 02/24/09 Time: 13:21
Sample: 1949 2003
Included observations: 55
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.198041	2.305251	0.085908	0.9319
UNEM	-0.017397	0.331705	-0.052446	0.9584
AR(1)	-0.055103	0.114662	-0.480565	0.6329
RESID(-1)	0.151531	0.176863	0.856771	0.3956

R-squared	0.014189	Mean dependent var	-1.33E-12
Adjusted R-squared	-0.043800	S.D. dependent var	2.122863
S.E. of regression	2.168855	Akaike info criterion	4.456223
Sum squared resid	239.9006	Schwarz criterion	4.602211
Log likelihood	-118.5461	Hannan-Quinn criter.	4.512678
F-statistic	0.244685	Durbin-Watson stat	1.715550
Prob(F-statistic)	0.864729		

Distributed lag models

Model: $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_k x_{t-k} + u_t$

β_0: immediate effect of a unit shock in x on y

$\beta_0, \beta_1, \beta_2, \ldots$: lag distribution

$\beta_0 + \beta_1 + \ldots + \beta_k$: effect of a permanent unit shock in x on y (long run multiplier)

There are models with infinite lags, e.g. geometric lags (geometrically declining β-s)

Seminar

Time series regressions I

AR(2)-process:

- Stationarity, necessary conditions of stationarity
- Solution of Yule–Walker equation

Autocorrelations of ARMA(1,1) process

Simulation of ARMA and ARIMA processes

Comparison and simulation difference and trend stationary processes

Box–Jenkins modelling, forecasting from ARIMA model

- Example: modelling time series of (seasonally adjusted) industrial production