Advanced Econometrics - Lecture 2

Heteroskedasticity and Autocorrelation

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Gauss-Markov Assumptions

Observation y_i is a linear function

$$y_i = x_i'\beta + \varepsilon_i$$

of observations x_{ik}, $k = 1, \ldots, K$, of the regressor variables and the error term ε_i

for $i = 1, \ldots, N$; $x_i' = (x_{i1}, \ldots, x_{iK})$

A1	$E\{\varepsilon_i\} = 0$ for all i
A2	all ε_i are independent of all x_i (exogeneous x_i)
A3	$V\{\varepsilon_i\} = \sigma^2$ for all i (homoskedasticity)
A4	$Cov\{\varepsilon_i, \varepsilon_j\} = 0$ for all i and j with $i \neq j$ (no autocorrelation)

OLS Estimators: Properties

Under assumptions (A1) and (A2):

1. E{b} = β, the OLS estimator is unbiased

Under assumptions (A1), (A2), (A3), and (A4):

2. The variance of the OLS estimator b is

   $$V\{b\} = \sigma^2\left(\Sigma_i x_i x_i' \right)^{-1}$$

3. $s^2 = e'e/(N-K)$ is unbiased for σ^2
4. The OLS estimator b is BLUE (best linear unbiased estimator) for β

Implications of Gauss-Markov Assumptions

The conditional distribution of error terms ε given X fulfills

- $E\{\varepsilon \mid X\} = 0$
- $V\{\varepsilon \mid X\} = \sigma^2 I$

ε: N-dimensional vector of all error terms

X: NxK matrix of explanatory variables

I: NxN identity matrix

The conditional distribution of ε given X has

- zero means
- constant variances and zero covariances

Violations of $V\{\varepsilon|X\} = \sigma^2 I$

In economic reality,

1. constancy of variances of the error terms may be violated

   $V\{\varepsilon \mid X\} = \text{diag}\{\sigma_1^2, \ldots, \sigma_N^2\} = \sigma^2 \, \text{diag}\{h_1^2, \ldots, h_N^2\}$

   the error terms are denoted as heteroskedastic

2. the error terms may be correlated

   $V\{\varepsilon \mid X\} = \sigma^2\Psi$, with a positive definite matrix Ψ with diagonal elements 1; the error terms are denoted as autocorrelated or serially correlated

The notation $V\{\varepsilon \mid X\} = \sigma^2 \Psi$ with a positive definite matrix Ψ encompasses both

- Heteroskedasticity: diagonal matrix Ψ
- Autocorrelation: Ψ with diagonal elements 1

The Questions

Aspects of both heteroskedasticity and autocorrelation of error terms

- What are the consequences of violations of $V\{\varepsilon \mid X\} = \sigma^2 I$ for the OLS estimator?
- How can violations of $V\{\varepsilon \mid X\} = \sigma^2 I$ be identified?
- Which modifications of methods can be used in case of violations of $V\{\varepsilon \mid X\} = \sigma^2 I$?

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Heteroskedasticity: Typical Situations

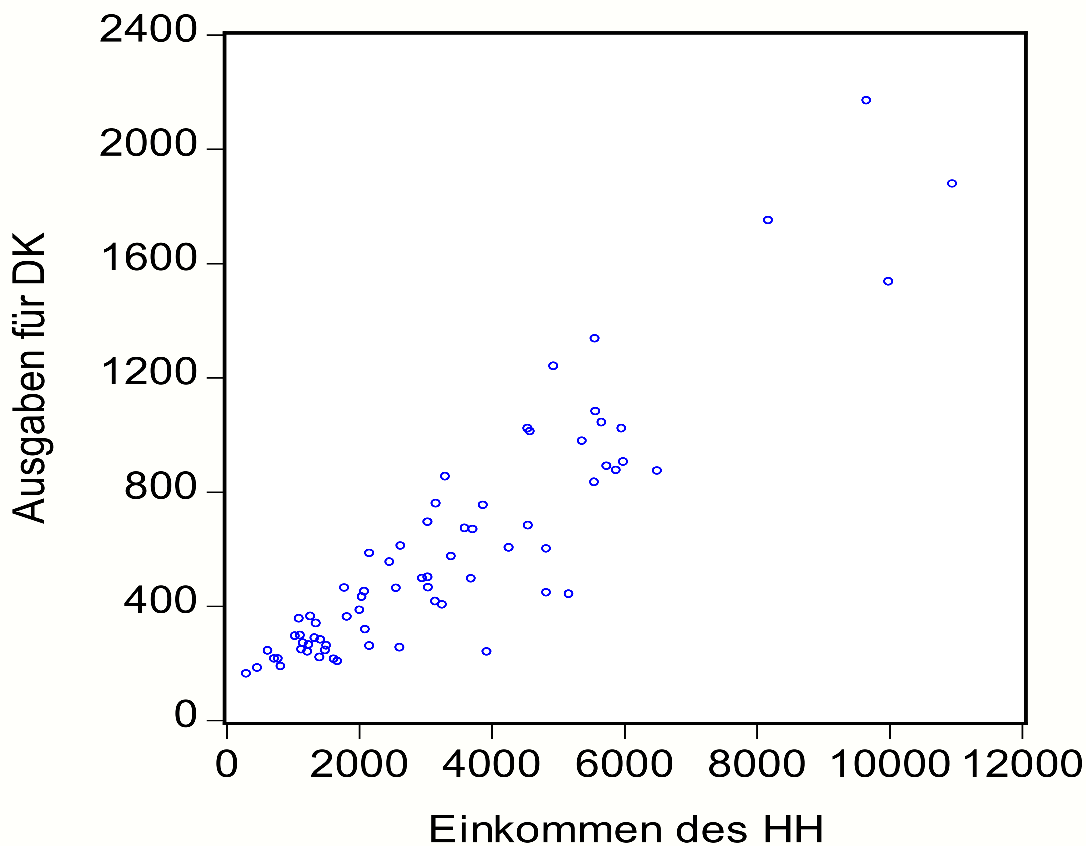
Heteroskedasticity is typically observed

- In cross sectional surveys, e.g., in household surveys:
  - Data, e.g., income from single person households vs. households with several individuals;
  - data from males and females
  - data from several regions
- For models with stochastic coefficients
- For data from financial markets, e.g., exchange rates, yields from securities

Heteroskedasticity: An Example

70 households:

monthly income of households and monthly expenditures for durable goods



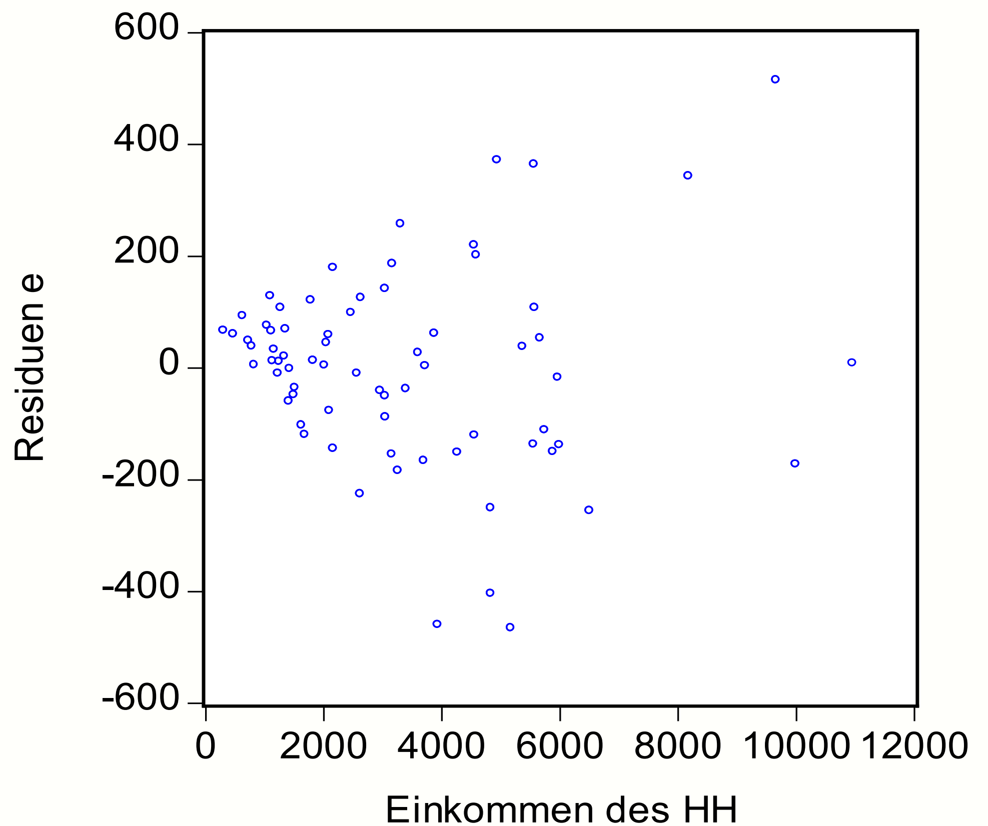
Heteroskedasticity: Example, cont`d

Residuals $e = y - \hat{y}$ from

$$\hat{Y} = 44.18 + 0.17\, X$$

X: monthly income
Y: monthly expenditures for durable goods

the larger the income, the more are the residuals scattered!



Heteroskedasticity: Stochastic Regression Coefficients

$Y_i = \alpha + \beta_i X_i + \varepsilon_i$: the coefficients are random

$$\beta_i = \beta + u_i$$

u_i: identically and independently distributed variable with variance σ_u^2 for all i

The model can be written as

$$Y_i = \alpha + \beta X_i + v_i$$

with error terms $v_i = \varepsilon_i + X_i u_i$

The variance of v_i

$$\text{Var}\{v_i\} = \sigma_e^2 + X_i^2 \sigma_u^2$$

is a function of X and not constant

Autocorrelation and economic time series

- Consumption in the current period does not differ too much from that in the previous period; the current consumption is a function of the consumption in the previous period
- Production, consumption, investment, etc.: typically, successive observations of economic variables are positively correlated
- The shorter the observation interval, the higher the correlation between economic variables
- Seasonal adjustment: applying smoothing or filtering procedures may cause correlated data

Autocorrelation: Typical Situations

Autocorrelation is typically observed for time series

- If a relevant regressor with trend is not included in the model; a case of missspecification
- If the functional form of a regressor is missspecified
- If the explained variable is autocorrelated in a form which is not adequately represented by the systematic part of the model

Attention!

- Autocorrelation of error terms is in general an indicator for not appropriate model specification
- Tests for autocorrelation are the most commonly used diagnostic tools for checking the model specification

OLS estimator in case of V{ε|X} ≠ $\sigma^2 I$

The case V{ε | X} = σ^2 Ψ with a positive definite matrix Ψ encompasses both heteroskedasticity and autocorrelation

- The OLS estimators

  $$b = (X'X)^{-1} X'y = \beta + (X'X)^{-1} X'\varepsilon$$

  are unbiased as E{ε | X} = 0; the violation V{ε | X} = $\sigma^2 I$ has no effect on a the expectation of the OLS estimators
- The covariance matrix of the OLS estimators is

  $$V\{b\} = \sigma^2 (X'X)^{-1} X' \Psi X (X'X)^{-1}$$

  with positive definite matrix Ψ from V{ε | X} = σ^2 Ψ

Consequences of $V\{\varepsilon|X\} \neq \sigma^2 I$

The consequences of both heteroskedasticity and autocorrelation are similar

- The OLS estimators are still unbiased but no longer BLUE
- Routinely computed standard errors s.e.(b) are incorrect

Ways to deal with this situation:

- Look for an alternative estimator which is more efficient than the OLS estimator
- Substitute the routinely computed, incorrect standard errors by corrected standard errors
- Reconsider the model specification

The GLS Estimator

The model $y = X\beta + \varepsilon$ with $V\{\varepsilon \mid X\} = \sigma^2 \Psi$ is transformed into a model

$$Py = y^* = PX\beta + P\varepsilon = X^*\beta + \varepsilon^*$$

such that $V\{\varepsilon^* \mid X\} = V\{P\varepsilon \mid X\} = PV\{\varepsilon \mid X\}P' = \sigma^2 P \Psi P' = \sigma^2 I$

This implies that the transformation matrix P is square and nonsingular with $P'P = \Psi^{-1}$; then $\Psi = (P'P)^{-1} = P^{-1}(P')^{-1}$

The OLS estimator □ for the transformed model $y^* = X^*\beta + \varepsilon^*$,

$$□ = (X^{*\prime}X^*)^{-1} X^{*\prime}y^* = (X'\Psi^{-1}X)^{-1} X'\Psi^{-1}y$$

is called GLS (generalized least squares) estimator

- □ is based on the Gauss-Markov assumptions!
- □ is unbiased and BLUE, $V\{□\} = \sigma^2 (X' \Psi^{-1}X)^{-1}$

The choice or derivation of P is specific for each situation or model

The EGLS Estimator

The transformation matrix *P* is a function of the elements of Ψ

To calculate the GLS estimator □, the matrix Ψ, which in most situations is unknown, is substituted by an (unbiased and consistent) estimated matrix ψ

The GLS estimator □ is derived in a two-step procedure:

1. Derive an estimate ψ for the matrix Ψ
2. Use the estimated matrix ψ to calculate the GLS estimator

$$\square = (X'\psi^{-1}X)^{-1}\, X'\psi^{-1}y$$

This estimator is called the estimated GLS or EGLS estimator for ß; it is also called FGLS (feasible GLS) estimator

- For large *N* the EGLS estimator and the GLS estimator have similar properties
- No guarantee that the EGLS outperforms the OLS estimator

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Consequences of Heteroskedasticity

- The OLS estimators b for β
  - are unbiased and consistent
  - have the covariance matrix

    $$V\{b\} = \sigma^2 (X'X)^{-1} X' \Psi X (X'X)^{-1}$$

  - are not efficient
  - follow under generally satisfied regularity conditions asymptotically the normal distribution
- The estimator $s^2 = e'e/(N-K)$ for the variance σ^2 of the error terms is biased

- Standard errors for b from $\sigma^2(X'X)^{-1}$ are biased
- Attention! The sign of the bias can be positive ore negative!
- t- and F-Test may be misleading

The GLS Estimator □

Heteroskedasticity: The error terms ε_i of $y_i = x_i'\beta + \varepsilon_i$ have variances

$$V\{\varepsilon_i | X\} = \sigma^2_i = \sigma^2 h_i^2$$

The transformed model ($P = \text{diag}\{h_1^{-1}, \ldots, h_N^{-1}\}$)

$$y_i / h_i = (x_i / h_i)'\beta + \varepsilon_i / h_i$$

has homoskedastic error terms: $V\{\varepsilon_i / h_i\} = \sigma^2$

GLS estimator □:

$$\tilde{b} = \left(\sum_i h_i^{-2} x_i x_i'\right)^{-1} \sum_i h_i^{-2} x_i y_i$$

- The GLS estimator is also denoted weighted least squares (WLS) estimator
- Observations with higher variance get a lower weight (they provide less accurate information on β)

Properties of □ and *t*-test

The covariance matrix of □ is

$$V\{□\} = \sigma^2 \left(\Sigma_i\, h_i^{-2}\, x_i\, x'_i\right)^{-1}$$

Estimation of the error term variance σ^2

$$\hat{\sigma}^2 = \frac{1}{N-K} \sum_i h_i^{-2} (y_i - x_i'\tilde{b})^2$$

t-statistic

$$t_k = \frac{\tilde{b}_k - q}{se(\tilde{b}_k)}$$

- Follows the *t*-distribution with *N*-*K* d.f., if the error terms are normally distributed and E{$□_k$} = q; se($□_k$) is the square root of the *k*-th diagonal element of Var{□} with estimated σ^2
- Can be used for testing H_0: $\beta_k = q$
- The *t*-distribution holds approximately under generally satisfied regularity conditions

The EGLS Estimator □ *

Estimates $\hat{h}_i$ for diagonal elements h_i from matrix Ψ

- N observations for estimating N quantities h_i
- Additional assumptions needed, depending on the form of heteroskedasticity
- Consistent estimator $\hat{h}_i^2$ implies asymptotically equivalent GLS □ and EGLS □ *

Concepts for estimating h_i

- Model for h_i as a function of regressor variables, variance function
  - $h_i^2 = \exp\{z_i'\alpha\}$; "multiplicative heteroskedasticity" (Verbeek)
  - More general, $h_i^2 = h(z_i'\alpha)$ with a non-negative function $h(.)$; see Breusch-Pagan test

Robust Standard Errors

The covariance matrix of the OLS estimator *b* is

$$V\{b\} = \sigma^2 (X'X)^{-1} X' \Psi X (X'X)^{-1}$$

Inference on β can be based on standard errors from V{*b*} if Ψ is substituted by suitable estimates :

- Heteroskedasticity-consistent covariance matrix estimator (HCCME)

$$\hat{V}\{b\} = \left(\sum_i x_i x_i'\right)^{-1} \left(\sum_i e_i^2 x_i x_i'\right) \left(\sum_i x_i x_i'\right)^{-1}$$

Heteroskedasticity-consistent standard errors: square roots of the diagonal elements of HCCME;

- also called White, heteroskedasticity-robust , or simply robust standard errors

Model-based Estimated vs Robust Standard Errors

Robust standard errors:

+ Need no information on the functional form of the variance function
+ Produce asymptotically valid inference
+ Widely available in econometric software, e.g. in GRETL

Model based standard errors:

+ Preferable to robust standard errors at least asymptotically, i.e., smaller standard errors, if true variance function is used
- Incorrect variance function may cause incorrect inference

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Tests for Heteroskedasticity

In case of heteroskedasticity: Results based on OLS may be misleading due to biased standard errors of OLS estimates *b*

Important to know whether the error terms fulfill homoskedasticity or not

Tests for checking the null hypothesis of homoskedasticity

- Breusch-Pagan-Test
- White-Test
- Goldfeld-Quandt-Test

Tests based on OLS residuals from original model

The Breusch-Pagan Test

Model for heteroskedasticity

$$\sigma_i^2 = \sigma^2 h(z_i'\alpha)$$

- function $h(.)$ with $h(.) > 0$ and $h(0) = 1$
- z_i: J variables including the intercept

Null hypothesis

$$H_0: \alpha = 0$$

i.e., $\sigma_i^2 = \sigma^2$ for all i, i.e. homoskedasticity

Breusch-Pagan test:

1. Auxiliary regression of the squared OLS residuals e_i^2 on z_i;, i.e., $h(.)$ a linear function; typically, z_i is chosen to be the model regressors; R_e^2
2. Test statistic: $BP = N R_e^2$, with R_e^2 from the auxiliary regression
3. BP follows the Chi-squared distribution with J d.f.

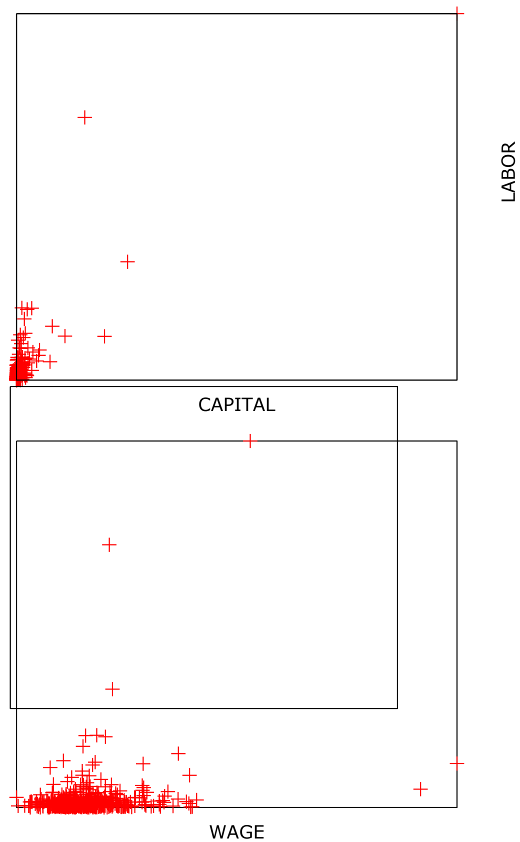
Example: Labor Demand

Labor demand function

- Explanatory variables: output, wage costs, capital stock
  - *LABOR*: total emploment (number of workers)
  - *CAPITAL*: total fixed assets (in Mio EUR)
  - *WAGE*: total wage costs per worker (in 1000 EUR)
  - *OUTPUT*: value added (in Mio EUR)
- Sample: 569 Belgian firms, data from1996
- Model specification *LABOR* = *g*(*OUTPUT, CAPITAL, WAGE*)
  - Linear model
  - Loglinear model

Labor Demand: The Data

	mean
LABOR	201.1
WAGE	38.6
OUTPUT	14.7
CAPITAL	11.5



Labor Demand Function: Linear Model

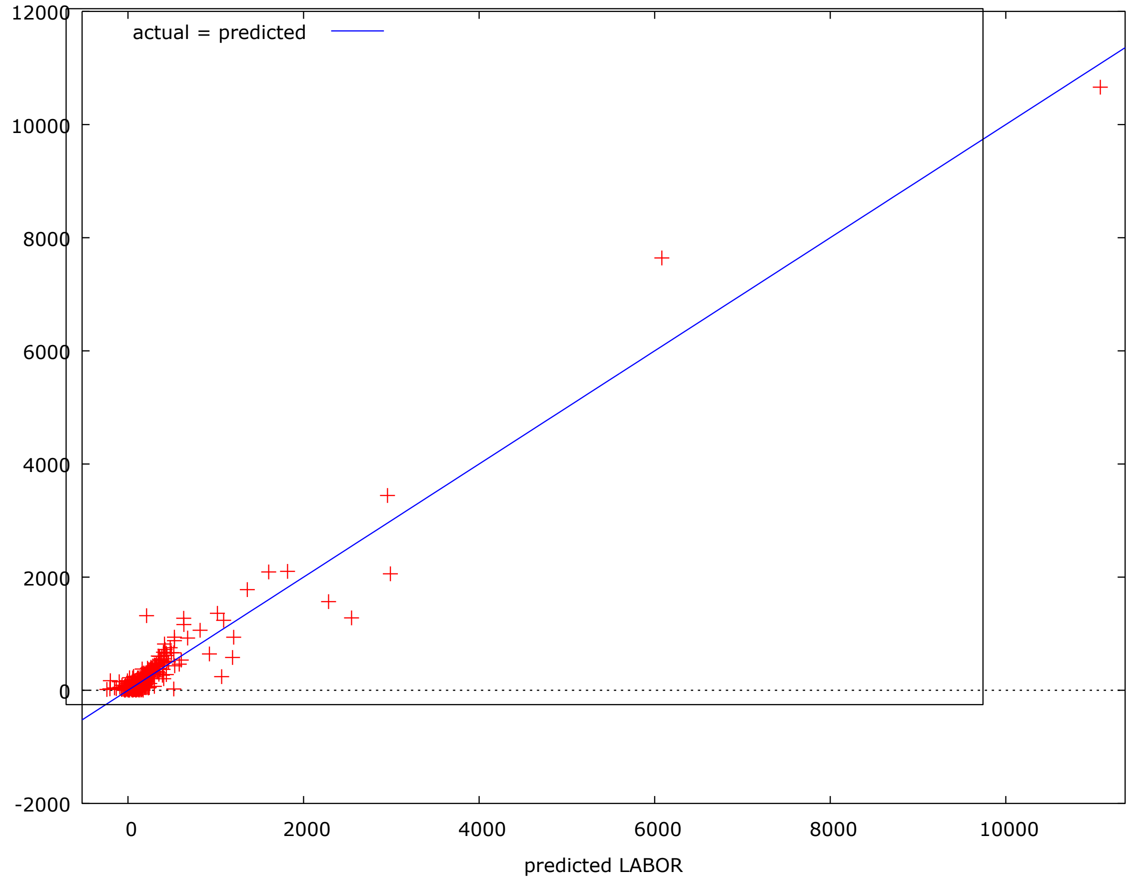
OLS estimated linear labor demand function : Output from GRETL

Modell 1:KQ, benutze die Beobachtungen 1-569
Abhängige Variable: LABOR

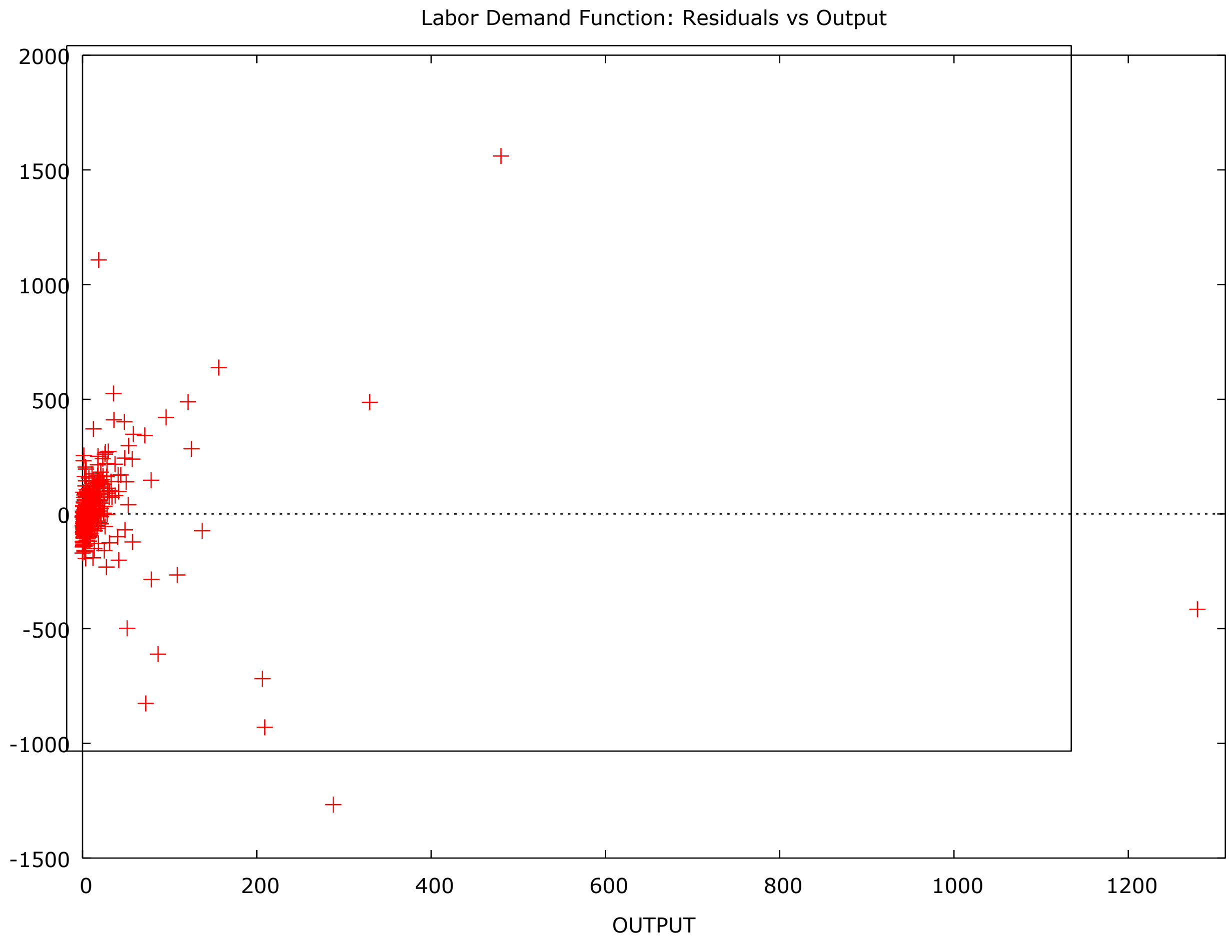
	Koeffizient	*Std. Fehler*	*t-Quotient*	*P-Wert*	
const	287,719	19,6418	14,6483	<0,00001	***
WAGE	**-6,7419**	**0,501405**	-13,4460	<0,00001	***
OUTPUT	**15,4005**	**0,355633**	43,3043	<0,00001	***
CAPITAL	**-4,59049**	**0,268969**	-17,0670	<0,00001	***

Mittel d. abh. Var.	201,0808	Stdabw. d. abh. Var.	611,9959
Summe d. quad. Res.	13795027	Stdfehler d. Regress.	156,2561
R-Quadrat	**0,935155**	Korrigiertes R-Quadrat	0,934811
F(3, 565)	**2716,024**	P-Wert(F)	0,000000
Log-Likelihood	-3679,670	Akaike-Kriterium	7367,341
Schwarz-Kriterium	7384,716	Hannan-Quinn-Kriterium	7374,121

Actual vs Predicted Labor Demand



Labor Demand: Residuals vs Output



Linear Labor Demand Function: Breusch-Pagan Test

Modell 2: KQ, benutze die Beobachtungen 1-569
Abhängige Variable: u2

	Koeffizient	Std.-fehler	t-Quotient	P-Wert	
const	-22719,5	11838,9	-1,919	0,0555	*
WAGE	228,857	302,217	0,7573	0,4492	
OUTPUT	5362,21	214,354	25,02	1,57e-093	***
CAPITAL	-3543,51	162,119	-21,86	3,25e-077	***

Mittel d. abh. Var.	24244,34	Stdabw. d. abh. Var.	145259,5
Summe d. quad. Res.	5,01e+12	Stdfehler d. Regress.	94181,84
R-Quadrat	0,581837	Korrigiertes R-Quadrat	0,579617
F(3, 565)	262,0493	P-Wert(F)	1,6e-106
Log-Likelihood	-7322,116	Akaike-Kriterium	14652,23
Schwarz-Kriterium	14669,61	Hannan-Quinn-Kriterium	14659,01

BP = 569x0.5818 = 331.1, p-value : 1.9E-71

Labor Demand Function: Loglinear Model

OLS estimated linear labor demand function : Output from GRETL

Modell 3: KQ, benutze die Beobachtungen 1-569
Abhängige Variable: l_LABOR

	Koeffizient	Std.-fehler	t-Quotient	P-Wert
const	6,17729	0,246211	25,09	6,53e-094 ***
l_WAGE	**-0,927764**	**0,0714046**	-12,99	5,85e-034 ***
l_OUTPUT	**0,990047**	**0,0264103**	37,49	2,23e-155 ***
l_CAPITAL	**-0,00369748**	**0,0187697**	-0,1970	0,8439

Mittel d. abh. Var.	4,488665	Stdabw. d. abh. Var.	1,171166
Summe d. quad. Res.	122,3388	Stdfehler d. Regress.	0,465327
R-Quadrat	0,842971	Korrigiertes R-Quadrat	0,842138
F(3, 565)	1011,023	P-Wert(F)	1,3e-226
Log-Likelihood	-370,0750	Akaike-Kriterium	748,1501
Schwarz-Kriterium	765,5256	Hannan-Quinn-Kriterium	754,9300

Loglinear Labor Demand Function, cont'd

Test for heteroskedasticity

- BP = 7.73, p-value: 0.052
- White test:
  - With regression on all regressors, squared regressors and interactions: test statistic: NxR^2 = 58.5; p-value: 2.6E-9
  - With regression on only regressors and squared regressors: test statistic: NxR^2 = 21.5; p-value: 0.0015

Loglinear Labor Demand Function, cont'd

White's Test für Heteroskedastizität
KQ, benutze die Beobachtungen 1-569
Abhängige Variable: uhat^2

	Koeffizient	Std.-fehler	t-Quotient	P-Wert	
const	2,54460	3,00278	0,8474	0,3971	
l_WAGE	-1,29900	1,75274	-0,7411	0,4589	
l_OUTPUT	-0,903725	0,559854	-1,614	0,1070	
l_CAPITAL	1,14205	0,375822	3,039	0,0025	***
sq_l_WAGE	0,192741	0,258954	0,7443	0,4570	
X2_X3	0,138038	0,162563	0,8491	0,3962	
X2_X4	-0,251779	0,104967	-2,399	0,0168	**
sq_l_OUTPUT	0,138198	0,0356469	3,877	0,0001	***
X3_X4	-0,191605	0,0368665	-5,197	2,84e-07	***
sq_l_CAPITAL	0,0895374	0,0139874	6,401	3,27e-010	***

White = 0.1029x569 = 58.5; *p*-value: 2.6E-9

The White Test

Generalizes the Breusch-Pagan test with linear function for heteroskedasticity with linear function $h(.)$

White test:

1. Auxiliary regression: the squared OLS residuals e_i^2 on all regressors, the squared regressors, and the interactions of the regressors

   $$e_i^2 = \Sigma_k \alpha_k x_{ik} + \Sigma_k \alpha_k x_{ik}^2 + \Sigma_k \Sigma_j \alpha_{kj} x_{ik}x_{ij}$$

   P: the number of coefficients in the auxiliary regression
2. Test statistic: $N R_e^2$, with R_e^2 from the auxiliary regression
3. The test statistic follows the Chi-squared distribution with P d.f.

Alternatively, the White test is based on $e_i^2 = \Sigma_k \alpha_k x_{ik} + \Sigma_k \alpha_k x_{ik}^2$

Goldfeld-Quandt-Test

Null hypothesis: homoskedasticity

Alternative: two regimes with variances of the error terms: σ_A^2 (regime A) and σ_B^2 (regime B)

Example:

$y_1 = X_1\beta_1 + u_1$, $Var\{u_1\} = \sigma_A^2 I_{N1}$ (Regime A)

$y_2 = X_2\beta_2 + u_2$, $Var\{u_2\} = \sigma_B^2 I_{N2}$ (Regime B)

Null hypothesis: $\sigma_A^2 = \sigma_B^2$

F-Test:

$$F = \frac{S_A}{S_B} \frac{N_B - K}{N_A - K}$$

S_i: sum of squared residuals for regime *i*

Goldfeld-Quandt-Test, cont'd

Test procedure:

1. Separate the N_A observations from regime A and the N_B observations from regime B

   For time series, arrange the observations in the order of increasing value of variable Z and drop $2c$ observations around the center of the ordered set of observations; $N_A = N_B = (N-c)/2$
2. Fit the model separately to the N_A and the N_B observations: OLS estimates b_i and sum of squared residuals S_i (i = A, B)
3. Determine the Goldfeld-Quandt test statistic

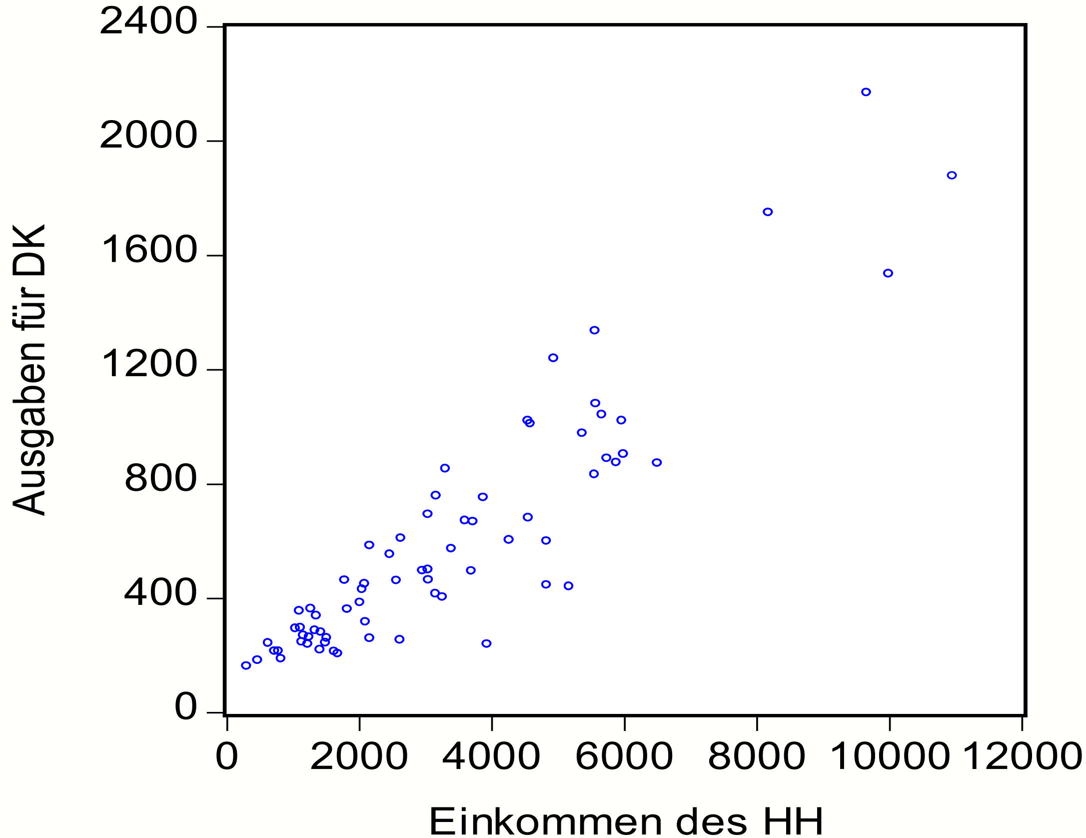
$$F = \frac{S_A}{S_B} \frac{N_B - c - K}{N_A - c - K}$$

   under the null hypothesis, F follows approximately the F-distribution with N_B-c-K and N_A-c-K d.f.

Example: Expenditures for Durable Goods

70 households:

monthly income (*X*) of households and monthly expenditures for durable goods (*Y*)



Household Expenditures

Households with (A) X<4000 and (B) X>4000: two regimes? $\sigma_A^2 \neq \sigma_B^2$?

The model $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$ is fitted

(I) to all data (N = 70): $\hat{Y} = 44.18 + 0.17\,X$, S = 2,094.511, s = 175.5

(II)(A): to data with X < 4000 (N_A = 48): $\hat{Y} = 119.71 + 0.13\,X$, S_A = 627.648, s_A = 117

(II)(B): to data with X > 4000 (N_B = 22): $\hat{Y} = -155.34 + 0.20\,X$, S_B = 1,331.777, s_B = 258

Test statistics:

$$F = \frac{1331777}{627648} \frac{48-2}{22-2} = 4.88$$

p-value: 0.000004; null hypothesis is to be rejected

Attention: Rejection can be caused by $\sigma_A^2 \neq \sigma_B^2$; but also because coefficients β_1 and β_2 change from regime A to regime B

Household Expenditures, cont'd

Breusch-Pagan test: Null hypothesis $\sigma_A^2 = \sigma_B^2$;

The alternative is: $\sigma_i^2 = \alpha_1 + \alpha_2 x_i$, i =1, …, N

1. Consumption function: $\hat{Y} = 44.18 + 0.17\,X$
2. Fitting the squared residuals e_t^2 to $\alpha_1 + \alpha_2\,x_i$ gives

   $R_e^2 = 0.2143$

   $BP = 70\,(0.2143) = 15.0$

   p-Wert: 0.0001

Null hypothesis is to be rejected

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Inference in Case of Heteroskedasticity

Under heteroskedasticity, the covariance matrix of the OLS estimators b is:

$$V\{b\} = \sigma^2 (X'X)^{-1} X' \Psi X (X'X)^{-1}$$

The use of the covariance matrix $\sigma^2(X'X)^{-1}$ and the corresponding standard errors for inference like

- t-tests, F-test
- Confidence intervals

has the risk of biased results

It is recommended

- To use corrected, i.e., robust standard errors
- To transform the model so that the error terms are homoskedastic

Labor Demand: White's s.e.

The White standard errors or robust or heteroskedasticity-consistent standard errors

variable	estimate	OLS s.e.	White s.e.
constant	6.177	0.246	0.294
log(*WAGE*)	-0.928	0.071	0.087
log(*OUTPUT*)	0.990	0.026	0.047
log(*CAPITAL*)	-0.004	0.019	0.038

The uncorrected standard errors underestimate between 20 and 100%

As a consequence, changed *p*-values are obtained for the *t*- and *F*-test

Transformation to Homoskedasticity

The transformation requires information about the function $h(.)$ from $h_i^2 = h(z_i'\alpha)$ [the error term of $y_i = x_i'\beta + \varepsilon_i$ has variances $V\{\varepsilon_i | X\} = \sigma^2_i = \sigma^2 h_i^2$]

In case of multiplicative heteroskedasticity,

$$\hat{h}_i^2 = \exp(z_i'a)$$

With transformation of the model into

$$\frac{y_i}{\hat{h}_i} = \frac{x_i'}{\hat{h}_i}\beta + \frac{\varepsilon_i}{\hat{h}_i}$$

the OLS estimators from the transformed model gives the EGLS estimators for β

Labor Demand: EGLS with $h_i^2 = \exp(z_i'\alpha)$

Estimates *b* and standard errors se(*b*) obtained from (i) the loglinear model with OLS, (ii) the loglinear model with OLS and White s.e., and (iii) the transformed model

Variable	OLS	EGLS	OLS s.e.	White s.e.	EGLS s.e.
constant	6.177	5.895	0.246	0.294	0.248
log(*WAGE*)	-0.928	-0.856	0.071	0.087	0.072
log(*OUTPUT*)	0.990	1.035	0.026	0.047	0.027
log(*CAPITAL*)	-0.004	-0.057	0.019	0.038	0.022

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Autocorrelation

Assumption (A4)

$\text{Cov}\{\varepsilon_i, \varepsilon_j\} = 0$ for all *i* and *j* with $i \neq j$

is violated

(in absence of heteroskedasticity)

$V\{\varepsilon \mid X\} = \sigma^2 \Psi$

with a positive definite matrix Ψ with diagonal elements 1

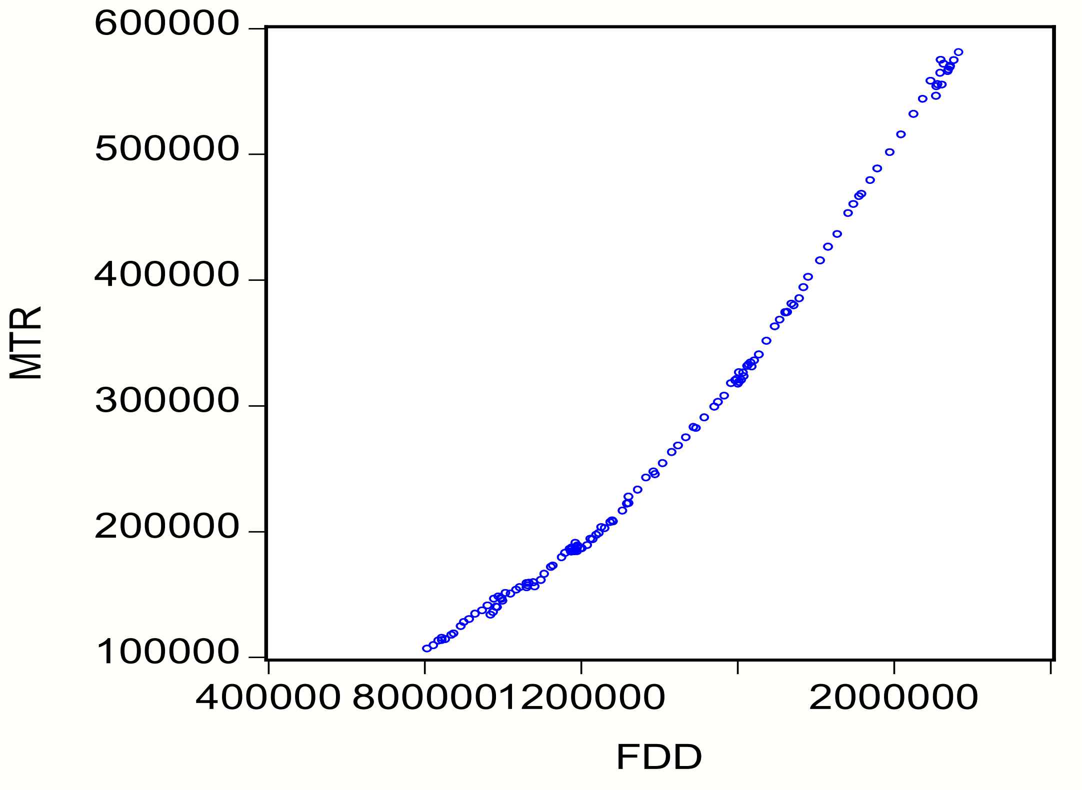
Autocorrelated or serially correlated error terms

Issues:

- What are consequences of autocorrelation?
- How to identify autocorrelation?
- What alternative methods that can be used to cope with autocorrelation?

Example: Import Function

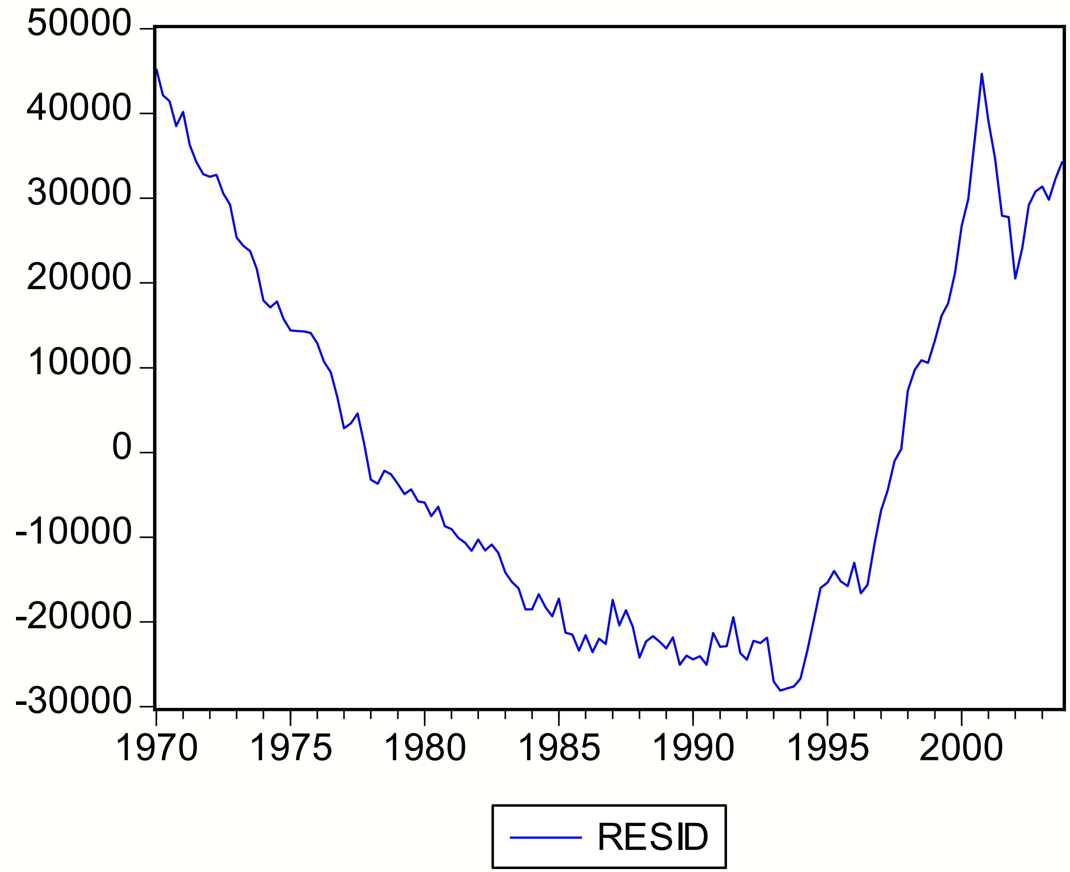
MTR: Imports
FDD: Demand
(from AWM database)



Import function: MTR = -227320 + 0.36 FDD
$R^2 = 0.977$, $t_{FFD} = 74.8$

Example: Import Function, cont`d

MTR: Imports
FDD: Demand
(from AWM database)



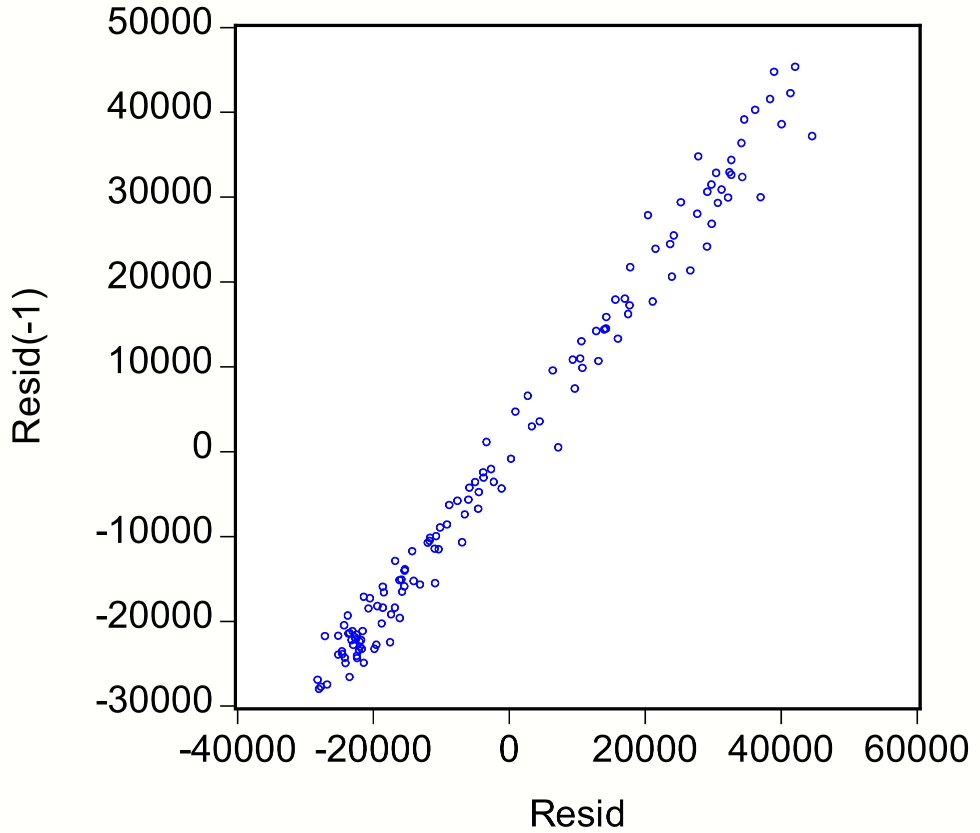
RESID: e_t = MTR - (-227320 + 0.36 FDD)

Example: Import Function, cont`d

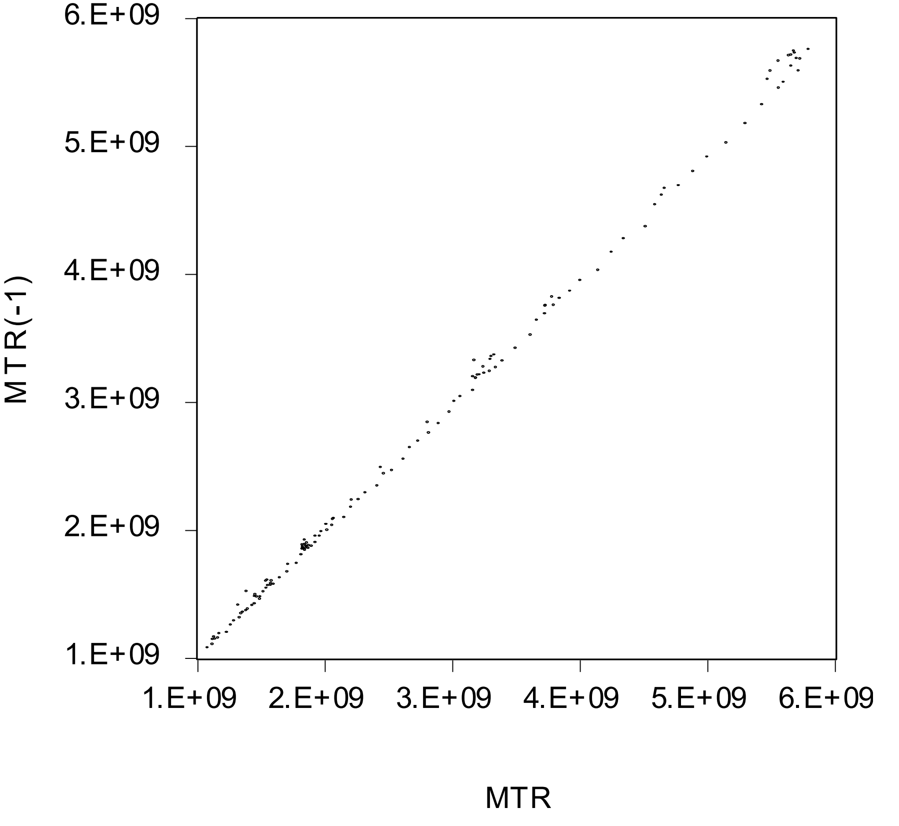
Lagged residuals [Resid(-1)] vs. actual residuals [Resid]

Attention! Serial correlation

r = 0.993



Example: Imports



Lagged imports [MTR(-1)] vs actual imports [MTR]

corr{MTR, MTR(-1)} = 0.9994

Typical Situations for Autocorrelation

Autocorrelation of the error terms typically occurs when

- a relevant regressor is not taken into account in the model, misspecification of the model
- the functional form of a regressor is erroneously specified
- the dependent variable has an autocorrelated pattern that is not adequately represented by the systematic part of the model

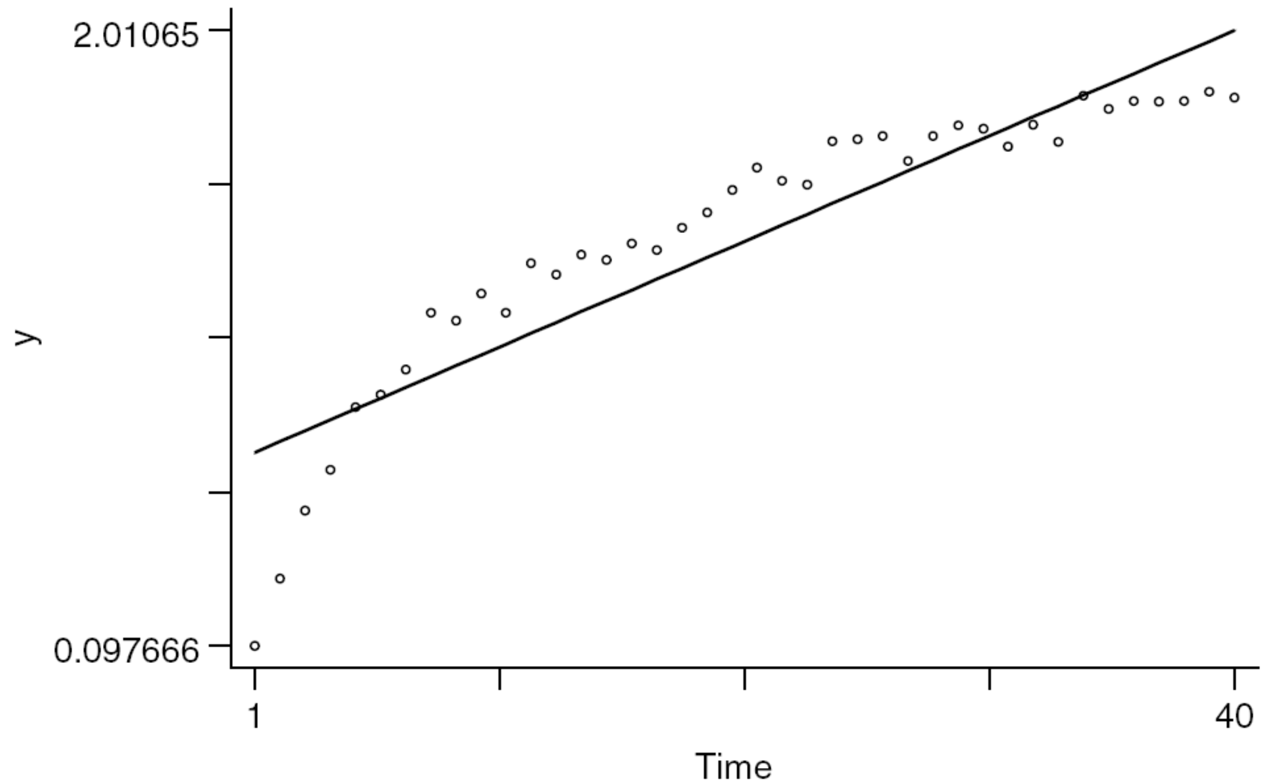
Autocorrelation of the error terms may indicate a misspecified model

- omitted variables
- incorrect functional forms
- incorrect dynamics

Autocorrelation tests are a tool for testing for misspecification

Wrong Functional Form

Simulated data (+) from 0.5 log(*Time*) and fitted linear model (Verbeek, p.117)



Positive Autocorrelation

Demand for ice cream explained from income and price index (Verbeek, p.106)

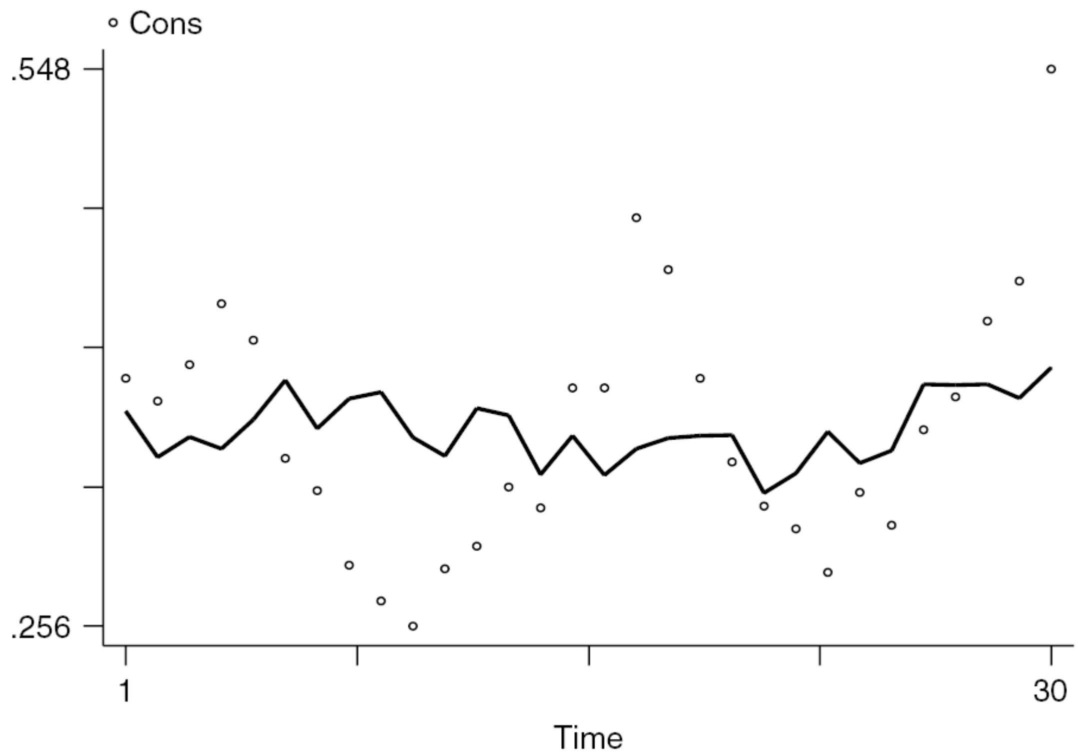


Figure 4.1 Actual and fitted consumption of ice cream, March 1951–July 1953

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

First-order Autocorrelation

A model with first-order autocorrelated error terms:

$$y_t = x_t'\beta + \varepsilon_t$$

with

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

where v_t is an error with mean zero and constant variance σ_v^2; v_t is called “white noise”

Assumptions: for $\rho = 0$, the Gauss-Markov conditions are met

The AR(1)-Process

For all *t*,

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

with white noise v_t

ε_t is generated by an autoregression or by an AR(1) process, an autoregressive process of order 1

Properties of ε_t are derived for $| \rho | < 1$

- $E\{\varepsilon_t\} = 0$ for all *t*
- $V\{\varepsilon_t\} = \sigma_v^2(1 - \rho^2)^{-1}$

  which follows from $V\{\varepsilon_t\} = V\{\rho\varepsilon_{t-1} + v_t\} = \rho^2 V\{\varepsilon_{t-1}\} + \sigma_v^2$
- $Cov\{\varepsilon_t, \varepsilon_{t-1}\} = E\{\varepsilon_t\, \varepsilon_{t-1}\} = \sigma_v^2\, \rho(1- \rho^2)^{-1}$
- $Cov\{\varepsilon_t, \varepsilon_{t-s}\} = \sigma_v^2\, \rho^s(1- \rho^2)^{-1}$ for all *s*

All error terms are correlated; the covariances decrease with growing distance *s* in time between the error terms

Autocorrelation function: $Cov\{\varepsilon_t, \varepsilon_{t-s}\}$ vs lag *s*

Imports: Autocorrelation function

Date: 05/15/05 Time: 16:57

Sample: 1970:1 2003:4
Included observations: 136

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.|*******|	.|*******|	1	0.968	0.968	130.30	0.000
.|*******|	.|. |	2	0.936	-0.017	253.06	0.000
.|*******|	.|. |	3	0.903	-0.041	368.07	0.000
.|*******|	.|. |	4	0.869	-0.021	475.47	0.000
.|****** |	*|. |	5	0.832	-0.069	574.71	0.000
.|****** |	.|. |	6	0.799	0.044	666.95	0.000
.|****** |	.|. |	7	0.768	0.001	752.66	0.000
.|****** |	.|. |	8	0.739	0.029	832.69	0.000
.|***** |	*|. |	9	0.706	-0.080	906.37	0.000
.|***** |	*|. |	10	0.668	-0.107	972.93	0.000
.|***** |	*|. |	11	0.626	-0.092	1031.8	0.000
.|**** |	*|. |	12	0.581	-0.081	1082.9	0.000

The AR(1)-Process, cont'd

Covariance matrix V{ε} of the errors ε

$$V\{\varepsilon\} = \sigma_v^2 \Psi = \frac{\sigma_v^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & \cdots & \rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \cdots & 1 \end{pmatrix}$$

with $|\rho| < 1$

V{ε}

- Has a band structure
- Depends only – besides σ_v^2 – of the parameter ρ
- Elements are decreasing with growing lag s

A Transformed Model, GLS Estimators

Model $y_t = x_t'\beta + \varepsilon_t$ with

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

where v_t is white noise

The transformed model

$$y_t - \rho y_{t-1} = (x_t - \rho x_{t-1})'\beta + v_t\,,\; t = 2, \ldots, T$$

satisfies the Gauss-Markov conditions

The differences $y_t - \rho y_{t-1}$ are called Cochrane-Orcutt transformations of the y_t; analogously for x_t

Given that ρ is known, estimation of coefficients of this model results (almost) in the GLS estimator

Note: Information of the first observation is lost by the transformation

Typically, ρ is unknown

Estimation of ρ

Model $y_t = x_t'\beta + \varepsilon_t$ with

$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$

where v_t is white noise

1. OLS estimation of β; residuals e_t
2. Auxiliary regression of residuals e_t on its lagged values e_{t-1} gives the OLS estimator $\hat{r}$ for ρ

$$\hat{r} = \left(\sum_{t=2}^{T} e_{t-1}^2\right)^{-1} \sum_{t=2}^{T} e_t e_{t-1}$$

The estimator $\hat{r}$

- is typically biased
- is consistent for ρ under weak regularity conditions

Cochrane-Orcutt Estimator

Model $y_t = x_t'\beta + \varepsilon_t$ with

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

where v_t is white noise

Two steps:

1. OLS estimation of β, estimation of $\hat{r}$ for ρ from auxiliary regression, Cochrane-Orcutt transformation $y_t^* = y_t - \hat{r}\, y_{t-1}$, $x_t^* = x_t - \hat{r}\, x_{t-1}$ for $t = 2, \ldots, T$
2. OLS estimation of β and σ_v^2 from

$$y_t^* = x_t^{*\prime}\, \beta + v_t$$

gives the Cochrane-Orcutt estimators for β (EGLS estimator)

The Cochrane-Orcutt estimator is based on only T-1 observations!

Iterative Cochrane-Orcutt estimator: repeat the estimation of ρ and step 2 until convergence

Prais-Winsten Estimator

The Prais-Winsten estimator is based on all T observations

The transformed model

$$y_t - \hat{r}\, y_{t-1} = (x_t - \hat{r}\, x_{t-1})' \beta + v_t, \quad t = 2, \ldots, T$$

is supplemented by an equation for the first observation:

$$\sqrt{1-\hat{r}^2}\, y_1 = \sqrt{1-\hat{r}^2}\, x_1'\beta + \sqrt{1-\hat{r}^2}\, \varepsilon_1$$

$$y_t - \hat{r} y_{t-1} = (x_t - \hat{r} x_{t-1})'\beta + v_t,\ t = 2,\ldots, n$$

For the first equation:

- The error term ε_1 is uncorrelated with all v_t
- $V\{\varepsilon_1\} = \sigma_v^2(1 - \rho^2)^{-1}$, so that $V\{(1 - \rho^2)^{-1/2}\, \varepsilon_1\}$ has the same variance σ_v^2 as that of all other error terms

Tests for Autocorrelation

Residuals indicate autocorrelation (*b* is an unbiased estimator)

Graphical displays of residuals give indication on autocorrelation of errors

Tests on the basis of residuals

- Durbin-Watson test
- Asymptotic tests, Breusch-Godfrey test

Advanced Econometrics - Lecture 2

- Violations of V{ε|X} = $\sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Asymptotic Tests for Autocorrelation

The auxiliary regression of residuals e_t on its lagged values e_{t-1} gives

- the OLS estimator $\hat{r}$ for ρ
- standard error *for* $\hat{r}$

The following test can be performed:

1. *t*-test: the statistic *t* for the *t*-test is approximately

   $$t \approx \hat{r}\sqrt{T}$$

   under the null hypothesis (ρ = 0) it follows approximately the *t*-distribution with T-1 df
2. Breusch-Godfrey test: $(T-1)R^2$ with R^2 from the auxiliary regression follows under the null hypothesis (ρ = 0) approximately the Chi-squared distribution with 1 df

Durbin-Watson Test

Requirements:

- the model has an intercept
- no lagged dependent variables as regressor; cf. assumption (A2)

Test statistic

$$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2} \approx 2(1-\hat{r})$$

- For ρ>0: *DW* is in the interval (0,2)
- For ρ<0: *DW* is in the interval (2,4)
- *DW* close to the value 2: no indication of autocorrelation
- *DW* close to 0 or 4: errors are highly correlated

Durbin-Watson Test, cont`d

Distribution and critical limits for *DW*:

- depends upon regressors x_t
- exact critical values are unknown, but upper bounds (d_U) and lower bounds (d_L) can be derived
  - d< d_L: H_0 rejected
  - d> d_U: H_0 not is rejected
  - d_L < d < d_U: no decision (inconclusive region)

Bounds for critical limits for $\alpha = 0.05$

T	*K*=2		*K*=3		*K*=10	
	$d_{L,0.05}$	$d_{U,0.05}$	$d_{L,0.05}$	$d_{U,0.05}$	$d_{L,0.05}$	$d_{U,0.05}$
15	1.08	1.36	0.95	1.54	0.17	3.22
20	1.20	1.41	1.10	1.54	0.42	2.70
100	1.65	1.69	1.63	1.71	1.48	1.87

Import functions, cont`d

- Regression of imports (MTR) on demand (FDD)

  $MTR = -2.27 \times 10^9 + 0.357\ FDD$, $t_{FDD} = 74.9$, $R^2 = 0.977$

  $DW = 0.014 < 1.69 = d_{L,0.05}$ for $T = 136$, $K = 2$

- Import function with trend (T)

  $MTR = -4.45 \times 10^9 + 0.653\ FDD - 0.030 \times 10^9\ T$

  $t_{FDD} = 45.8$, $t_T = -21.0$, $R^2 = 0.995$

  $DW = 0.093 < 1.68 = d_{L,0.05}$ for $T = 136$, $K = 3$

- Import function with lagged imports as regressor

  $MTR = -0.124 \times 10^9 + 0.020\ FDD + 0.956\ MTR_{-1}$

  $t_{FDD} = 2.89$, $t_{MTR(-1)} = 50.1$, $R^2 = 0.999$

  ($DW = 1.079 < 1.68 = d_{L,0.05}$ for $T = 135$, $K = 3$)

Durbin-Watson Test, cont`d

- DW test does not indicate
  - reasons for rejecting the null hypothesis
  - how the model can be improved
- Reason for rejecting the null hypothesis can be various kinds of misspecification
- Test for autocorrelation of first order; analyzing quarterly data suggests test for fourth order autocorrelation
- The inconclusive region and the limited number of critical barriers (K, T, α) make the test unwieldy

Advanced Econometrics - Lecture 2

- Violations of $V\{\varepsilon|X\} = \sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

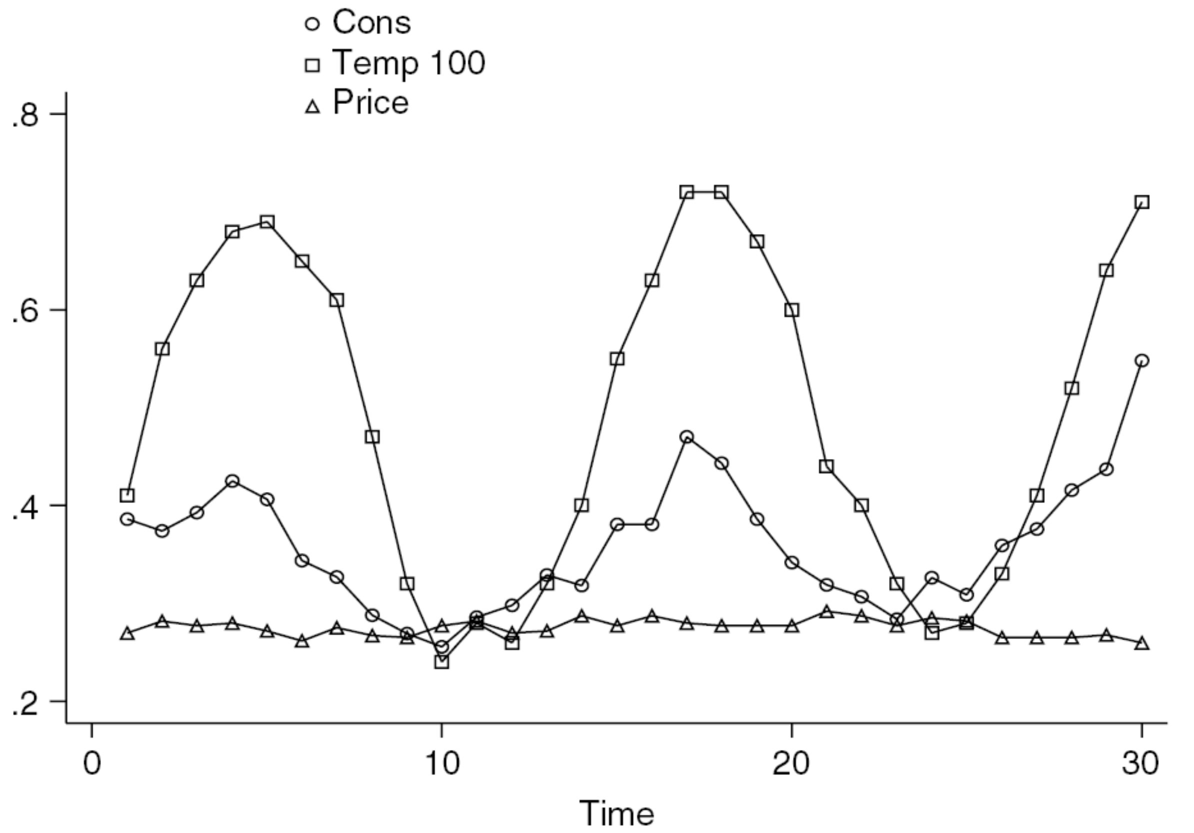
Example: Demand for Ice Cream

Time-series from Hildreth and Lu (1960):

- 30 four-weekly observations, 1951-1953
- Variables:
  - *cons*: consumption of ice cream per head (in pints)
  - *income*: average family income per week (in USD)
  - *price*: price of ice cream (per pint)
  - *temp*: average temperature (in F)

Demand for Ice Cream: Data

Consumption and price of ice cream, temperature (Verbeek, p.112)



Demand for Ice Cream: OLS Results

Consumption of ice cream, explained by three regressors (Verbeek, p.112)

D&W test (T=30, K=4):

$DW = 1.02 < d_{L;0.05} = 1.21$

Table 4.9 OLS results

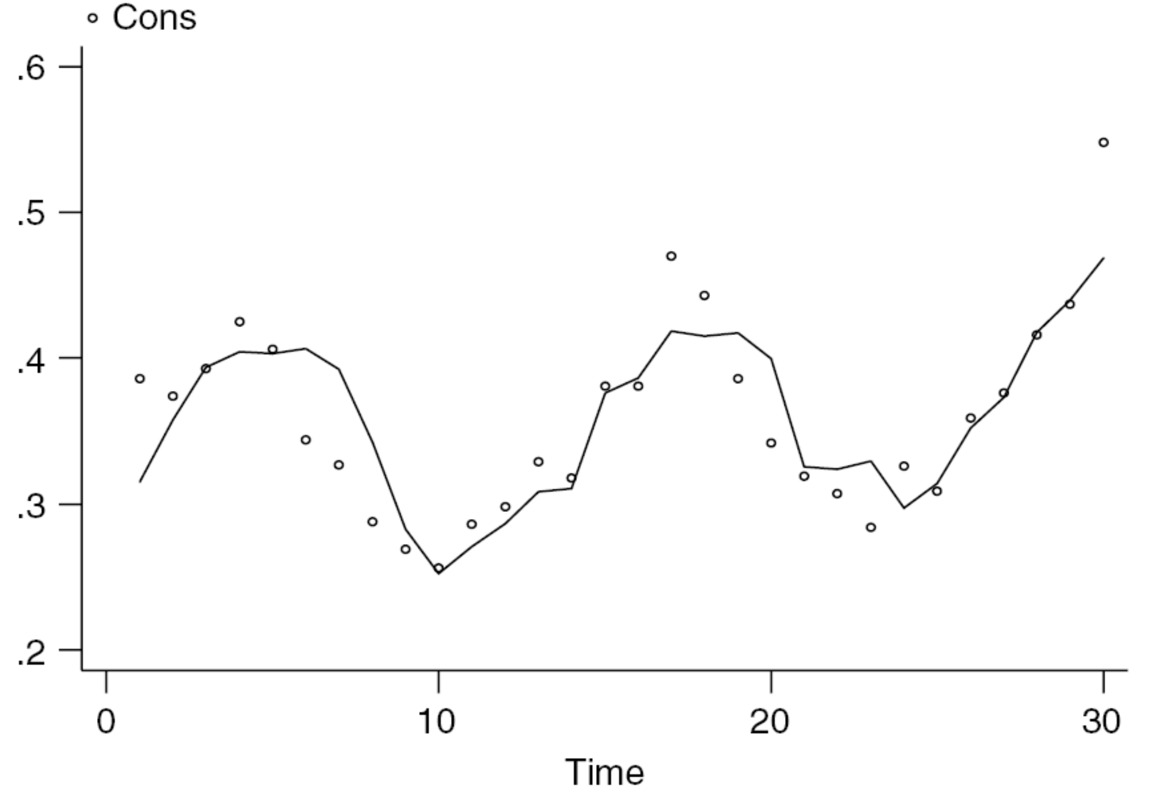
Dependent variable: *cons*

Variable	Estimate	Standard error	t-ratio
constant	0.197	0.270	0.730
price	−1.044	0.834	−1.252
income	0.00331	0.00117	2.824
temp	0.00345	0.00045	7.762

$s = 0.0368$ $R^2 = 0.7190$ $\bar{R}^2 = 0.6866$ $F = 22.175$
$dw = 1.0212$

Demand for Ice Cream: Actual and Fitted Values

Actual (+) and fitted (connected points) values of consumption of ice cream (Verbeek, p.113)



Demand for Ice Cream: Autocorrelation

Regression of the OLS residuals e_t on e_{t-1} gives

- $\hat{r} = 0.401$
- $R^2 = 0.149$

Tests for autocorrelation

- $\hat{r}\sqrt{T} = 2.19$, p-value: 0.029
- $(T-1)\, R^2 = 4.32$, p-value: 0.038

Both tests reject the null hypothesis of no autocorrelation

Demand for Ice Cream: Cochrane-Orcutt

EGLS estimates based on the iterative Cochrane-Orcutt procedure (Verbeek, p.114)

Table 4.10 EGLS (iterative Cochrane–Orcutt) results

Dependent variable: *cons*

Variable	Estimate	Standard error	t-ratio
constant	0.157	0.300	0.524
price	−0.892	0.830	−1.076
income	0.00320	0.00159	2.005
temp	0.00356	0.00061	5.800
$\hat{\rho}$	0.401	0.2079	1.927

$s = 0.0326^*$ $R^2 = 0.7961^*$ $\bar{R}^2 = 0.7621^*$ $F = 23.419$
$dw = 1.5486^*$

Demand for Ice Cream: An Alternative Model

DW in the inconclusive region for α = 0.05 (Verbeek, p.114)

Table 4.11 OLS results extended specification

Dependent variable: *cons*

Variable	Estimate	Standard error	t-ratio
constant	0.189	0.232	0.816
price	−0.838	0.688	−1.218
income	0.00287	0.00105	2.722
temp	0.00533	0.00067	7.953
$temp_{t-1}$	−0.00220	0.00073	−3.016

$s = 0.0299$ $R^2 = 0.8285$ $\bar{R}^2 = 0.7999$ $F = 28.979$
$dw = 1.5822$

Advanced Econometrics - Lecture 2

- Violations of V{ε|X} = $\sigma^2 I$
- Heteroskedasticity and Autocorrelation
- Heteroskedasticity: Estimates
- Heteroskedasticity: Tests
- Heteroskedasticity: Alternatives
- Autocorrelation: Cases and Examples
- First order Autocorrelation
- Tests for Autocorrelation
- Demand for Ice Cream
- Autocorrelation: some Extensions

Alternative Autocorrelation Patterns

Alternative patterns for autocorrelation of the error terms

- Higher order autocorrelation
- Moving average structure

Higher order autocorrelation

- Typically, for quarterly data the AR(4) structure is appropriate

$$\varepsilon_t = \gamma\varepsilon_{t-4} + v_t$$

  with white noise v_t

- More generally, the AR(4) structure is

$$\varepsilon_t = \gamma_1\varepsilon_{t-1} + \ldots \gamma_4\varepsilon_{t-4} + v_t$$

- It is called 4th order autocorrelation

Alternative Autocorrelation Patterns, cont'd

Moving average structure

- Typically, if the correlation between different error terms is limited by a maximum time lag
- MA(1) structure

  $$\varepsilon_t = v_t + \alpha\, v_{t-1}$$

  with white noise v_t
- ε_t is correlated with ε_{t-1}, but not with ε_{t-2}, ε_{t-3}, …

Inference in Case of Autocorrelation

The options – in the preferred order – are:

1. Reconstder the model:
   - Change functional form , e.g., use log(x) rather than x
   - include additional explanatory variables (seasonals) or additional lags
2. Compute heteroskedasticity-and-autocorrelation consistent standard errors (HAC standard errors) for the OLS estimator;
3. Reconsider options 1 and 2; if autocorrelation is considered certain:
4. Use EGLS with existing model.

HAC Estimator for V{*b*}

Similar to the White standard errors for heteroskedasticity

Corrects OLS standard errors for both heteroskedasticity and autocorrelation

In $V\{b\} = \sigma^2 (X'X)^{-1} X'\Psi X (X'X)^{-1}$, Ψ is substituted by an appropriate estimator

Newey-West: Substitution of $S_x = \sigma^2(X'\Psi X)/T = (\Sigma_t\Sigma_s\sigma_{ts}x_t x_s')/T$ by

$$\hat{S}_x = \frac{1}{T}\sum_t e_t^2 x_t x_t' + \frac{1}{T}\sum_{j=1}^{p}\sum_t (1-w_j)e_t e_{t-j}(x_t x_{t-j}' + x_{t-j}x_t')$$

with $w_j = j/(p+1)$; the so-called truncation lag p is to be chosen appropriately

HAC (heteroskedasticity-and-autocorrelation consistent) estimator

$$T\,(X'X)^{-1}\,\hat{S}_x\,(X'X)^{-1}$$

Import Function, cont`d

Regression of imports (MTR) on demand (FDD)

$$MTR = -2.27 \times 10^9 + 0.357\ FDD,\ t_{FDD} = 74.9,\ R^2 = 0.977$$

OLS estimator and standard errors

variable	estimate	OLS s.e.	HAC s.e.
constant	2.27 E09	0.07 E09	0.18 E09
FDD	0.357	0.0048	0.0125

The non-corrected standard errors underestimate considerably

Dynamic Models

The model

$$y_t = x_t'\beta + \varepsilon_t \quad \text{with } \varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

describes both

- $E\{y_t \mid x_t\} = x_t'\beta$
- $E\{y_t \mid x_t, x_{t-1}, y_{t-1}\} = x_t'\beta + \rho\,(y_{t-1} - x_{t-1}'\beta)$

The reformulation

$$y_t = x_t'\beta + \rho y_{t-1} - \rho x_{t-1}'\beta + v_t$$

specifies a linear model with uncorrelated error terms

In many cases, lagged values of y and/or x will eliminate the serial correlation problem

Some Import Functions

- Regression of imports (MTR) on demand (FDD)

  $$MTR = -2.27 \times 10^9 + 0.357\ FDD,\ t_{FDD} = 74.9,\ R^2 = 0.977$$

- Autocorrelation of the residuals:

  $$\mathrm{Corr}(e_t, e_{t-1}) = 0.993$$

- Import function with trend (T)

  $$MTR = -4.45 \times 10^9 + 0.653\ FDD - 0.030 \times 10^9\ T$$

  $$t_{FDD} = 45.8,\ t_T = -21.0,\ R^2 = 0.995$$

- Multicollinearity? Attention! Corr{FDD, T} = 0.987

- Import function with lagged imports as regressor

  $$MTR = -0.124 \times 10^9 + 0.020\ FDD + 0.956\ MTR_{-1}$$

  $$t_{FDD} = 2.89,\ t_{MTR(-1)} = 50.1,\ R^2 = 0.999$$

Exercise

1. Answer questions a, b, c, e, f and g of Exercise 4.1 of Verbeek.
2. Answer questions of Exercise 4.2 of Verbeek.