
Econometrics - Lecture 4

Heteroskedasticity and Autocorrelation

Contents

- Violations of $V\{\varepsilon\} = \sigma^2 I_N$
- Heteroskedasticity
- GLS Estimation
- Autocorrelation

Gauss-Markov Assumptions

Observation y_i is a linear function

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

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

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

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	$\text{Cov}\{\varepsilon_i, \varepsilon_j\} = 0$ for all i and j with $i \neq j$ (no autocorrelation)

In matrix notation: $E\{\varepsilon\} = 0$, $V\{\varepsilon\} = \sigma^2 I_N$

OLS-Estimator Properties

Under assumptions (A1) and (A2):

1. The OLS estimator b is unbiased: $E\{b\} = \beta$

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

2. The variance of the OLS estimator is given by

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

3. The sampling variance s^2 of the error terms ε_i ,

$$s^2 = (N - K)^{-1} \sum_i e_i^2$$

is unbiased for σ^2

4. The OLS estimator b is BLUE (best linear unbiased estimator)

Violations of $V\{\varepsilon\} = \sigma^2 I_N$

Implications of the Gauss-Markov assumptions for ε :

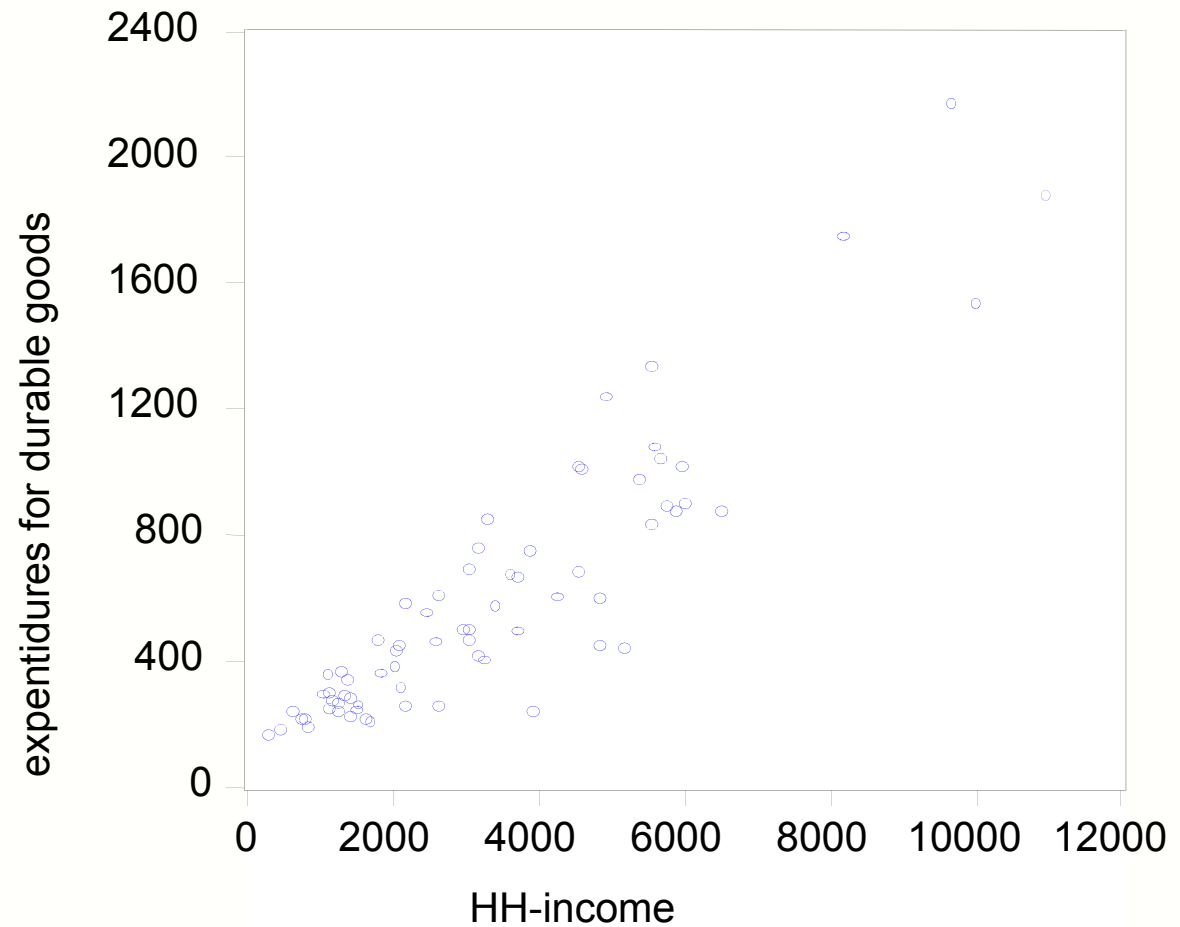
$$V\{\varepsilon\} = \sigma^2 I_N$$

Violations:

- Heteroskedasticity: $V\{\varepsilon\} = \text{diag}(\sigma_1^2, \dots, \sigma_N^2)$ or
 $V\{\varepsilon\} = \sigma^2 \Psi = \sigma^2 \text{diag}(h_1^2, \dots, h_N^2)$
- Autocorrelation: $V\{\varepsilon_i, \varepsilon_j\} \neq 0$ for at least one pair $i \neq j$ or
 $V\{\varepsilon\} = \sigma^2 \Psi$
with non-diagonal elements different from zero

Example: Household Income and Expenditures

70 households (HH):
monthly HH-
income and
expenditures for
durable goods



Household Income and Expenditures, cont'd

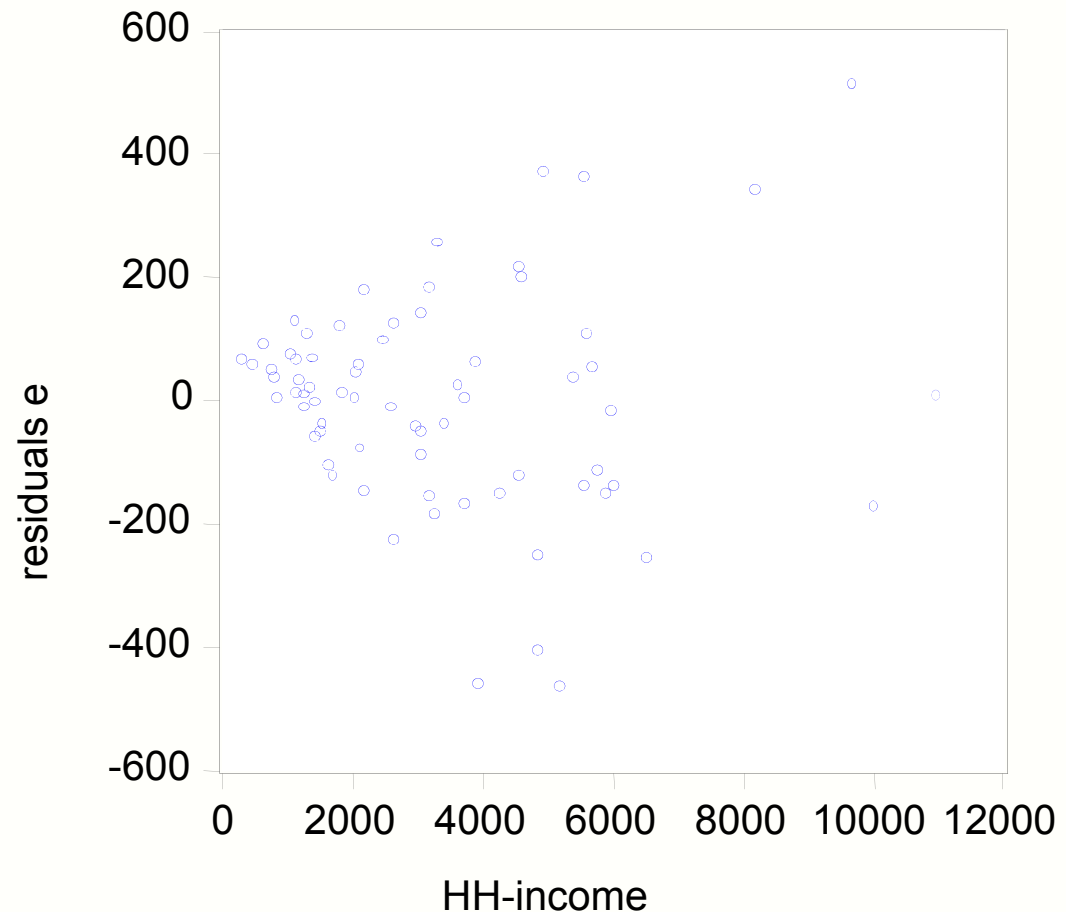
Residuals $e = y - \hat{y}$ from

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

X : monthly HH-income

Y : expenditures for durable goods

the larger the income,
the more are the
residuals scattered



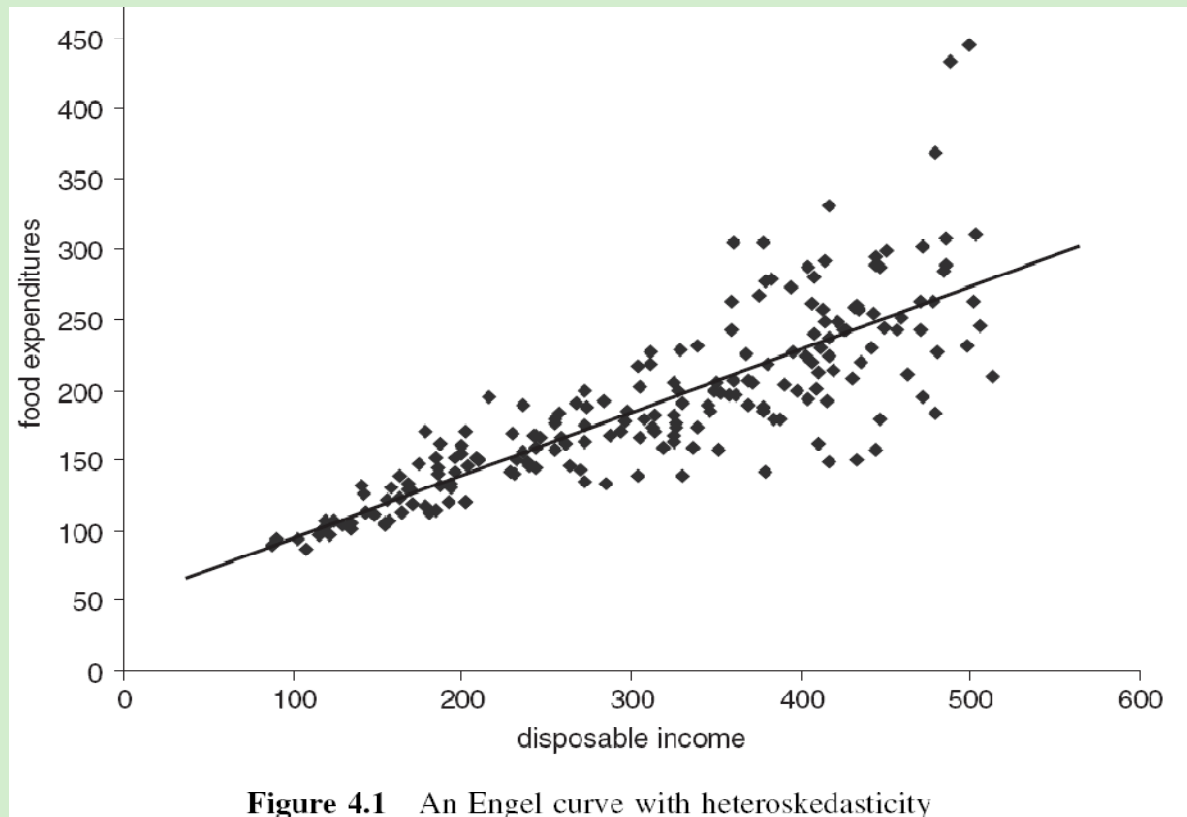
Typical Situations for Heteroskedasticity

Heteroskedasticity is typically observed

- In data from cross-sectional surveys, e.g., in households or regions
- Data with variance which depends of one or several explanatory variables, e.g., firm size
- Data from financial markets, e.g., exchange rates, stock returns

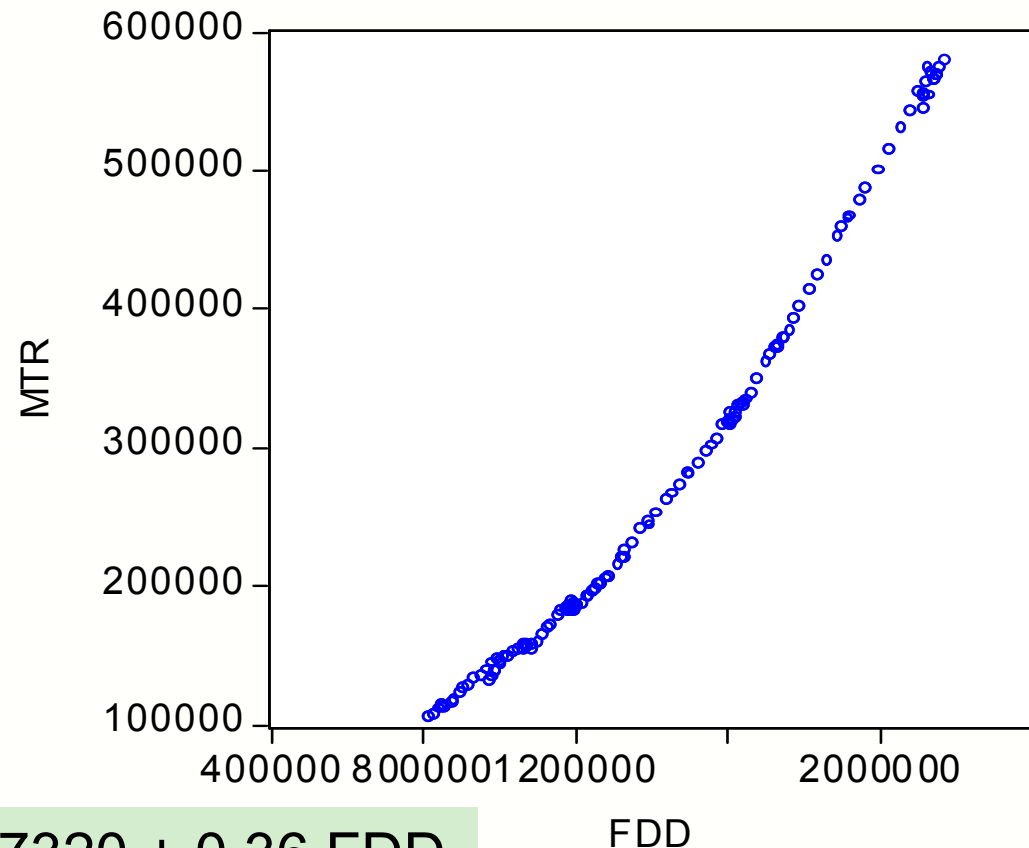
Example: Household Expenditures

With growing income increasing variation of expenditures; from Verbeek, Fig. 4.1



Example: Import Function

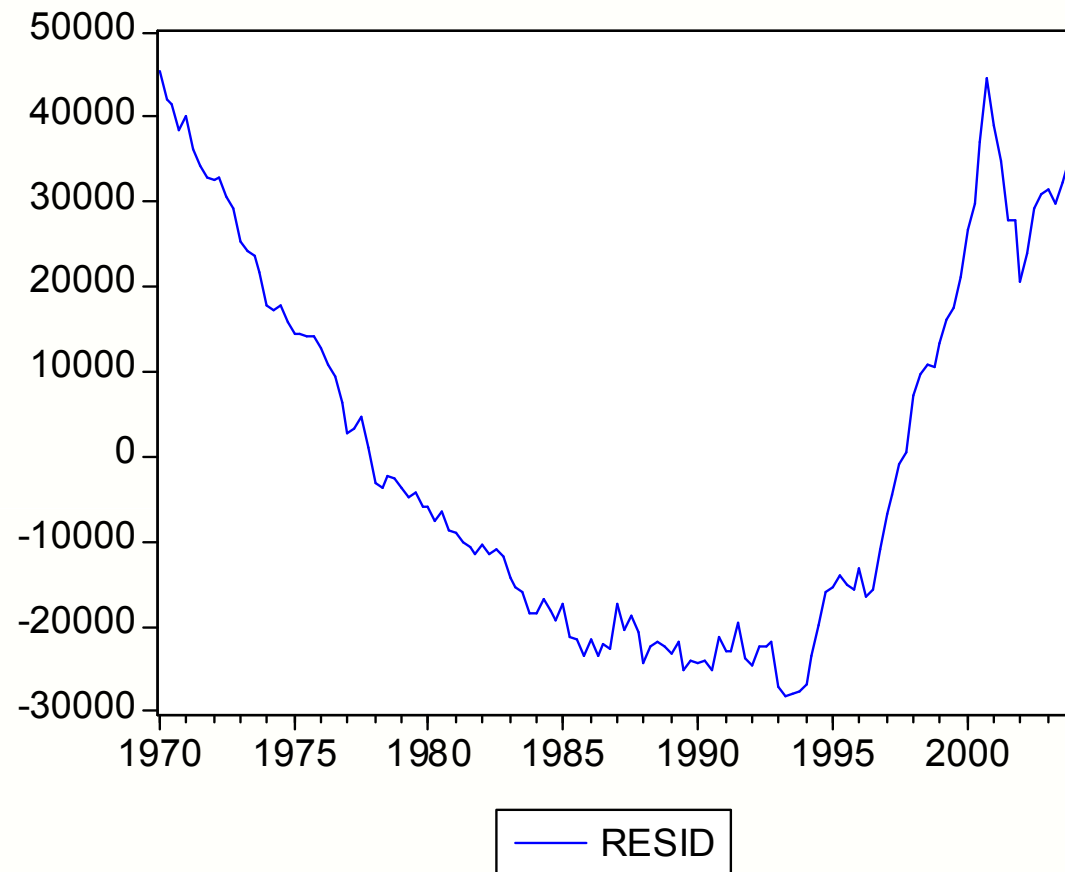
MTR: Imports
FDD: Demand
(from AWM-database)



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

Import Function, cont'd

MTR: Imports
FDD: Demand
(from AWM-database)

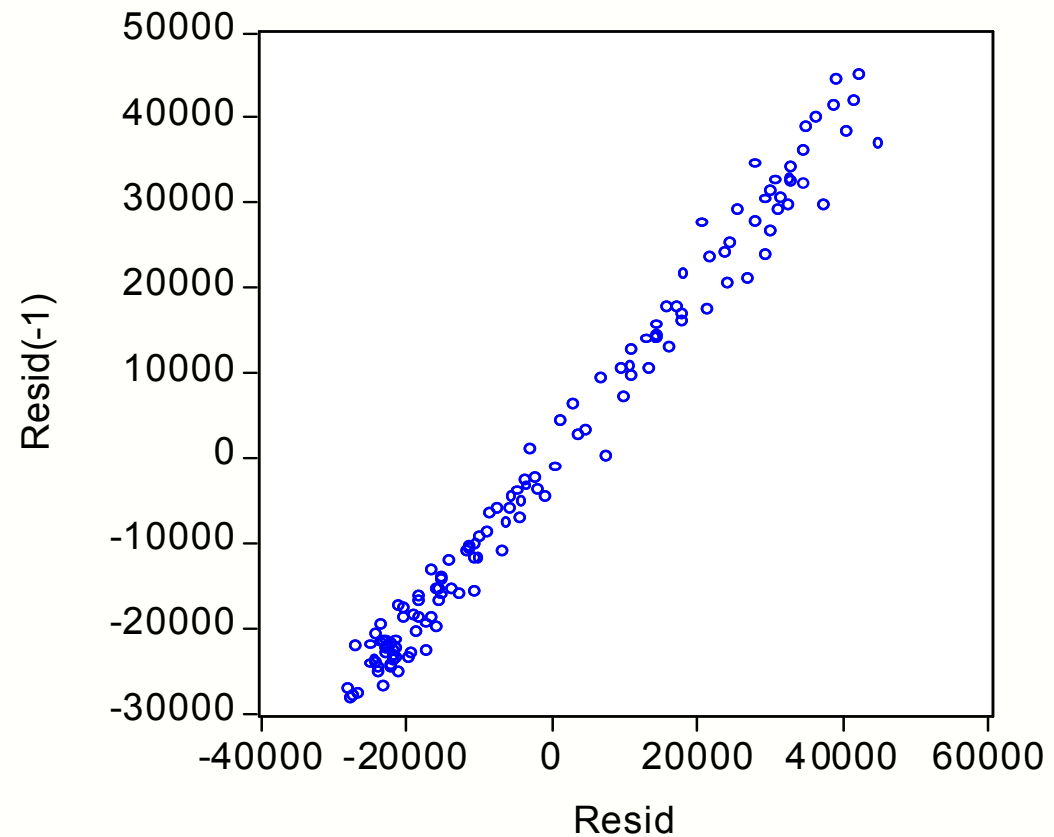


$$\text{RESID: } e_t = \text{MTR} - (-227320 + 0.36 \text{ FDD})$$

Import Function, cont'd

Scatter-diagram of by one period lagged residuals [Resid(-1)] against actual residuals [Resid]

Serial correlation!



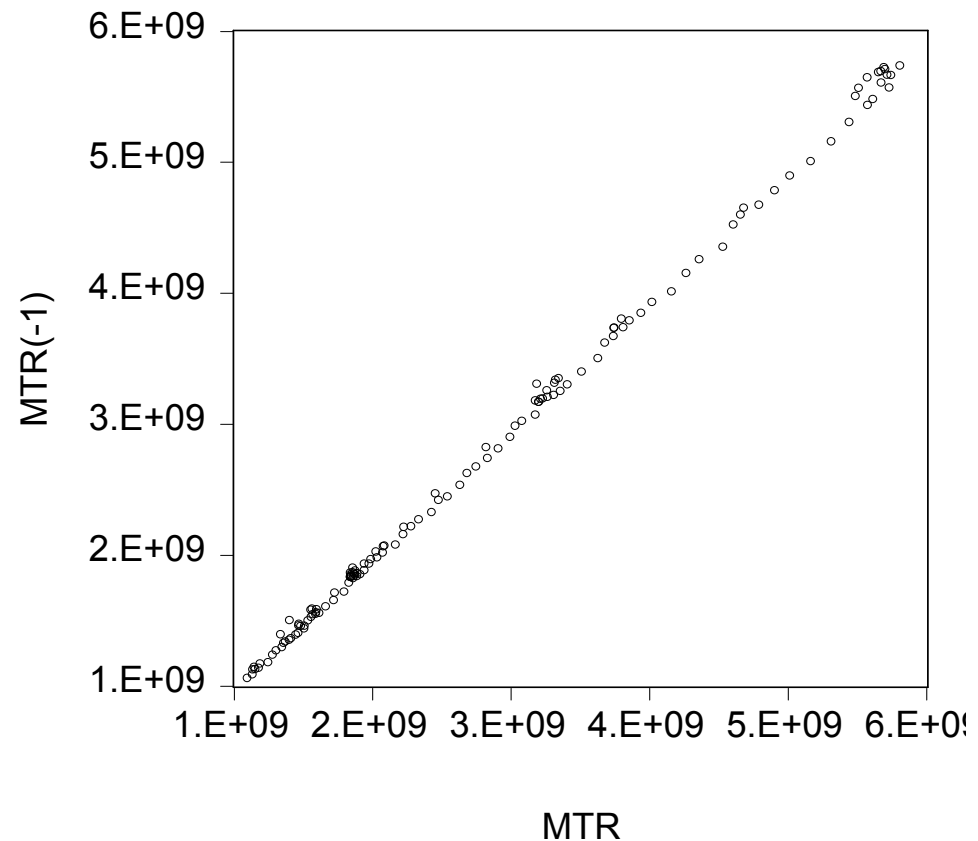
Autocorrelation of Economic Time-series

- Consumption in actual period is similar to that of the preceding period; the actual consumption „depends“ on the consumption of the preceding period
- Consumption, production, investments, etc.: it is to be expected that successive observations of economic variables correlate positively
- Seasonal adjustment: application of smoothing and filtering algorithms induces correlation

Example: Imports

Scatter-diagram of by one period lagged imports [MTR(-1)] against actual imports [MTR]

Correlation coefficient between MTR und MTR(-1): 0.9994



Typical Situations for Autocorrelation

Autocorrelation is typically observed if

- a relevant regressor with trend or seasonal pattern is not included in the model: miss-specified model
- the functional form of a regressor is incorrectly specified
- the dependent variable is correlated in a way that is not appropriately represented in the systematic part of the model

Warning! Omission of a relevant regressor with trend implies autocorrelation of the error terms; in econometric analyses autocorrelation of the error terms is always possible!

- Autocorrelation of the error terms indicates deficiencies of the model specification
- Tests for autocorrelation are the most frequently used tool for diagnostic checking the model specification

Import Functions

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

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

Autocorrelation (order 1) of residuals:

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

- Import function with trend (T)

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

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

Multicollinearity? $\text{Corr}(\text{FDD}, T) = 0.987!$

- Import function with lagged imports as regressor

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

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

Consequences of $V\{\varepsilon\} \neq \sigma^2 I_N$

OLS-estimators b for β

- are unbiased
- are consistent
- have the covariance-matrix

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

- are not efficient estimators, not BLUE
- follow – under general conditions – asymptotically the normal distribution

The estimator $s^2 = e'e/(N-K)$ for σ^2 is biased

Consequences of $V\{\varepsilon\} \neq \sigma^2 I_N$ for Applications

- OLS-estimators b for β are still unbiased
- Routinely computed standard errors are biased; the bias can be positive or negative
- t - and F -tests may be misleading

Remedies

- Alternative estimators
- Corrected standard errors
- Modification of the model

Test for identification of

- heteroskedasticity
 - autocorrelation
- are important tools

Contents

- Violations of $V\{\varepsilon\} = \sigma^2 I_N$
- Heteroskedasticity
- GLS Estimation
- Autocorrelation

Inference under Heteroskedasticity

Covariance matrix of b :

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

Use of $\sigma^2 (X'X)^{-1}$ (the standard output of econometric software) instead of $V\{b\}$ for inference on β may be misleading

Remedies

- Use of correct variances and standard errors
- Transformation of the model so that the error terms are homoskedastic

The Correct Variances

- $V\{\varepsilon_{ij}\} = \sigma_i^2 = \sigma^2 h_i^2$: each observation has its own unknown parameter h_i
- N observation for estimating N unknown parameters?

To estimate σ_i^2 – and $V\{b\}$

- Known form of the heteroskedasticity, specific correction
 - E.g., $h_i^2 = z_i' \alpha$ for some variables z_i
 - Requires estimation of α
- White's heteroskedasticity-consistent covariance matrix estimator (HCCME)

$$\tilde{V}\{b\} = \sigma^2 (X'X)^{-1} (\sum_i \hat{h}_i^2 x_i x_i') (X'X)^{-1}$$

with $\hat{h}_i^2 = e_i^2$

- Denoted as HC_0
- Inference based on HC_0 : heteroskedasticity-robust inference

White's Standard Errors

White's standard errors for b

- Square roots of diagonal elements of HCCME
- Underestimate the true standard errors
- Various refinements, e.g., $HC_1 = HC_0[N/(N-K)]$

In GRETL: HC_0 is the default HCCME, HC_1 and other refinements are optionally available

An Alternative Estimator for b

Idea of the estimator

- Transform the model so that it satisfies the Gauss-Markov assumptions
- Apply OLS to the transformed model
- Results in a BLUE

Transformation often depends upon unknown parameters that characterizing heteroskedasticity: two-step procedure

1. Estimate the parameters that characterize heteroskedasticity and transform the model
2. Estimate the transformed model

The procedure results in an approximately BLUE

An Example

Model:

$$y_i = x_i' \beta + \varepsilon_i \quad \text{with } V\{\varepsilon_i\} = \sigma_i^2 = \sigma^2 h_i^2$$

Division by h_i results in

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

with a homoskedastic error term

$$V\{\varepsilon_i/h_i\} = \sigma_i^2/h_i^2 = \sigma^2$$

OLS applied to the transformed model gives

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

It is called a generalized least squares (GLS) or weighted least squares (WLS) estimator

Weighted Least Squares Estimator

- A GLS or WLS estimator is a least squares estimator where each observation is weighted by a non-negative factor $w_i > 0$:

$$\hat{\beta}_w = \left(\sum_i w_i x_i' x_i \right)^{-1} \sum_i w_i x_i' y_i$$

- Weights proportional to the inverse of the error term variance: Observations with a higher error term variance have a lower weight; they provide less accurate information on β
- Needs knowledge of the h_i
 - seldom available
 - mostly provided by estimates of h_i based on assumptions on the form of h_i
 - e.g., $h_i^2 = z_i' \alpha$ for some variables z_i
- Analogous with general weights w_i

Example: Labor Demand

Verbeek's data set "labour2": Sample of 569 Belgian companies (data from 1996)

- Variables

- *labour*: total employment (number of employees)
- *capital*: total fixed assets
- *wage*: total wage costs per employee (in 1000 EUR)
- *output*: value added (in million EUR)

- Labour demand function

$$labour = \beta_1 + \beta_2 * wage + \beta_3 * output + \beta_4 * capital$$

Labor Demand Function

For Belgian companies, 1996; Verbeek

Table 4.1 OLS results linear model

Dependent variable: *labour*

Variable	Estimate	Standard error	<i>t</i> -ratio
constant	287.72	19.64	14.648
<i>wage</i>	-6.742	0.501	-13.446
<i>output</i>	15.40	0.356	43.304
<i>capital</i>	-4.590	0.269	-17.067

$s = 156.26$ $R^2 = 0.9352$ $\bar{R}^2 = 0.9348$ $F = 2716.02$

Labor Demand Function, cont'd

Can the error terms be assumed to be homoskedastic?

- They may vary depending of the company size, measured by, e.g., size of output or capital
- Regression of squared residuals on appropriate regressors will indicate heteroskedasticity

Labor Demand Function, cont'd

Auxiliary regression of squared residuals, Verbeek

Table 4.2 Auxiliary regression Breusch–Pagan test

Dependent variable: e_i^2

Variable	Estimate	Standard error	<i>t</i> -ratio
constant	−22719.51	11838.88	−1.919
<i>wage</i>	228.86	302.22	0.757
<i>output</i>	5362.21	214.35	25.015
<i>capital</i>	−3543.51	162.12	−21.858

$s = 94182$ $R^2 = 0.5818$ $\bar{R}^2 = 0.5796$ $F = 262.05$

Indicates dependence of error terms of *output*, *capital*, not of *wage*

Labor Demand Function, cont'd

Estimated function

$$labour = \beta_1 + \beta_2 * wage + \beta_3 * output + \beta_4 * capital$$

OLS estimates without (s.e.) and with White standard errors (White s.e.), and GLS estimates with $w_i = 1/e_i$

	β_1	β_2	β_3	β_4
Coeff OLS	287.19	-6.742	15.400	-4.590
s.e.	19.642	0.501	0.356	0.269
White s.e.	64.877	1.852	2.482	1.713
Coeff GLS	282.06	-6.609	15.235	-4.197
s.e.	1.808	0.042	0.094	0.141

The standard errors are inflated by factors between 3.7 (wage) and 6.4 (*capital*), 7.0 (*output*) wrt the White s.e.

Labor Demand Function, cont'd

With White standard errors: Output from GRETL

Dependent variable : LABOR

Heteroskedastic-robust standard errors, variant HCO,

	coefficient	std. error	t-ratio	p-value
const	287,719	64,8770	4,435	1,11e-05 ***
WAGE	-6,7419	1,8516	-3,641	0,0003 ***
CAPITAL	-4,59049	1,7133	-2,679	0,0076 ***
OUTPUT	15,4005	2,4820	6,205	1,06e-09 ***
Mean dependent var		201,024911	S.D. dependent var	611,9959
Sum squared resid		13795027	S.E. of regression	156,2561
R- squared		0,935155	Adjusted R-squared	0,934811
F(2, 129)		225,5597	P-value (F)	3,49e-96
Log-likelihood		455,9302	Akaike criterion	7367,341
Schwarz criterion		-3679,670	Hannan-Quinn	7374,121

Tests against Heteroskedasticity

Due to unbiasedness of b , residuals are expected to indicate heteroskedasticity

Graphical displays of residuals may give useful hints

Residual-based tests:

- Breusch-Pagan test
- Goldfeld-Quandt test
- White test

Breusch-Pagan Test

For testing whether the error term variance is a function of variables

$$Z_2, \dots, Z_p$$

Model for heteroskedasticity

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

with function h with $h(0)=1$, p -vectors z_i und α , an intercept and $p-1$ variables Z_2, \dots, Z_p

Null hypothesis

$$H_0: \alpha = 0$$

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

Auxiliary regression of the squared OLS residuals on z_i

Test statistic: NR^2 with R^2 of the auxiliary regression; follows the Chi-squared distribution with p d.f.

Breusch-Pagan Test, cont'd

Typical functions h for $h(z_i'\alpha)$

- Linear regression: $h(z_i'\alpha) = z_i'\alpha$
- Exponential function $h(z_i'\alpha) = \exp\{z_i'\alpha\}$
 - Auxiliary regression of the log (e_i^2) upon z_i
 - “Multiplicative heteroskedasticity”
 - Variances are non-negative

Labor Demand Function, cont'd

Auxiliary regression of squared residuals, Verbeek

Table 4.2 Auxiliary regression Breusch–Pagan test

Dependent variable: e_i^2			
Variable	Estimate	Standard error	t -ratio
constant	-22719.51	11838.88	-1.919
<i>wage</i>	228.86	302.22	0.757
<i>output</i>	5362.21	214.35	25.015
<i>capital</i>	-3543.51	162.12	-21.858

$s = 94182$ $R^2 = 0.5818$ $\bar{R}^2 = 0.5796$ $F = 262.05$

$NR^2 = 331.04$, p -value = $2.17E-70$; reject null hypothesis of homoskedasticity

Goldfeld-Quandt Test

For testing whether the error term variance has values σ_A^2 and σ_B^2 for observations from regime A and B, respectively, $\sigma_A^2 \neq \sigma_B^2$ regimes can be urban vs rural area, economic prosperity vs stagnation, etc.

Example (in matrix notation):

$$y_A = X_A \beta_A + \varepsilon_A, \quad V\{\varepsilon_A\} = \sigma_A^2 I_{N_A} \quad (\text{regime A})$$

$$y_B = X_B \beta_B + \varepsilon_B, \quad V\{\varepsilon_B\} = \sigma_B^2 I_{N_B} \quad (\text{regime B})$$

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

Test statistic:

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

with S_i : sum of squared residuals for i -th regime; follows under H_0 exactly or approximately the F -distribution with $N_A - K$ and $N_B - K$ d.f.

Goldfeld-Quandt Test, cont'd

Test procedure in three steps:

1. Sort the observations with respect to the regimes
2. Separate fittings of the model to the N_A and N_B observations; sum of squared residuals S_A and S_B
3. Calculation of test statistic F

White Test

For testing whether the error term variance is a function of the model regressors, their squares and their cross-products

Auxiliary regression of the squared OLS residuals upon x_i 's, squares of x_i 's and cross-products

Test statistic: NR^2 with R^2 of the auxiliary regression; follows the Chi-squared distribution with the number of coefficients in the auxiliary regression as d.f.

The number of coefficients in the auxiliary regression may become large, maybe conflicting with size of N , resulting in low power of the White test

Labor Demand Function, cont'd

White's test for heteroskedasticity

OLS, using observations 1-569

Dependent variable: uhat^2

	coefficient	std. error	t-ratio	p-value	

const	-260,910	18478,5	-0,01412	0,9887	
WAGE	554,352	833,028	0,6655	0,5060	
CAPITAL	2810,43	663,073	4,238	2,63e-05	***
OUTPUT	-2573,29	512,179	-5,024	6,81e-07	***
sq_WAGE	-10,0719	9,29022	-1,084	0,2788	
X2_X3	-48,2457	14,0199	-3,441	0,0006	***
X2_X4	58,5385	8,11748	7,211	1,81e-012	***
sq_CAPITAL	14,4176	2,01005	7,173	2,34e-012	***
X3_X4	-40,0294	3,74634	-10,68	2,24e-024	***
sq_OUTPUT	27,5945	1,83633	15,03	4,09e-043	***

Unadjusted R-squared = 0,818136

Test statistic: $TR^2 = 465,519295$,
with p-value = $P(\text{Chi-square}(9) > 465,519295) = 0,000000$

Contents

- Violations of $V\{\varepsilon\} = \sigma^2 I_N$
- Heteroskedasticity
- **GLS Estimation**
- Autocorrelation

Generalized Least Squares Estimator

- A GLS or WLS estimator is a least squares estimator where each observation is weighted by a non-negative factor $w_i > 0$

- Example:

$$y_i = x_i' \beta + \varepsilon_i \quad \text{with } V\{\varepsilon_i\} = \sigma_i^2 = \sigma^2 h_i^2$$

- Division by h_i results in a model with homoskedastic error terms

$$V\{\varepsilon_i / h_i\} = \sigma_i^2 / h_i^2 = \sigma^2$$

- OLS applied to the transformed model results in the weighted least squares (GLS) estimator with $w_i = h_i^{-2}$:

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

- The concept of transforming the model so that Gauss-Markov assumptions are fulfilled is used also in more general situations, e.g., for autocorrelated error terms

Properties of GLS Estimators

The GLS estimator

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

is a least squares estimator; standard properties of OLS estimator apply

- The covariance matrix of the GLS estimator is

$$V \{ \hat{\beta} \} = \sigma^2 \left(\sum_i h_i^{-2} x_i x_i' \right)^{-1}$$

- Unbiased estimator of the error term variance

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

- Under the assumption of normality of errors, t - and F -tests can be used; for large N , these properties apply approximately without normality assumption

Feasible GLS Estimator

Is a GLS estimator with estimated weights w_i

- Substitution of the weights $w_i = h_i^{-2}$ by estimates \hat{h}_i^{-2}

$$\hat{\beta}^* = \left(\sum_i \hat{h}_i^{-2} x_i x_i' \right)^{-1} \sum_i \hat{h}_i^{-2} x_i y_i$$

- Feasible (or estimated) GLS or FGLS or EGLS estimator
- For consistent estimates \hat{h}_i , the FGLS and GLS estimators are asymptotically equivalent
- For small values of N , FGLS estimators are in general not BLUE
- For consistently estimated \hat{h}_i , the FGLS estimator is consistent and asymptotically efficient with covariance matrix (estimate for σ^2 : based on FGLS residuals)

$$V \{ \hat{\beta}^* \} = \hat{\sigma}^2 \left(\sum_i \hat{h}_i^{-2} x_i x_i' \right)^{-1}$$

- Warning: the transformed model is uncentered

Multiplicative Heteroskedasticity

Assume $V\{\varepsilon_i\} = \sigma_i^2 = \sigma^2 h_i^2 = \sigma^2 \exp\{z_i' \alpha\}$

- The auxiliary regression

$$\log e_i^2 = \log \sigma^2 + z_i' \alpha + v_i \quad \text{with } v_i = \log(e_i^2/\sigma_i^2)$$

provides a consistent estimator $\hat{\alpha}$ for α

- Transform the model $y_i = x_i' \beta + \varepsilon_i$ with $V\{\varepsilon_i\} = \sigma_i^2 = \sigma^2 h_i^2$ by dividing through \hat{h}_i from $\hat{h}_i^2 = \exp\{z_i' \hat{\alpha}\}$
- Error term in this model is (approximately) homoskedastic
- Applying OLS to the transformed model gives the FGLS estimator for β

FGLS Estimation

In the following steps:

1. Calculate the OLS-estimates b for β
2. Compute the OLS residuals $e_i = y_i - x_i' b$
3. Regress $\log(e_i^2)$ on z_i and a constant, obtaining estimates \hat{a} for α

$$\log e_i^2 = \log \sigma^2 + z_i' \alpha + v_i$$

4. Compute $\hat{h}_i^2 = \exp\{z_i' \hat{a}\}$, transform all variables and estimate the transformed model to obtain the FGLS estimators:

$$y_i / \hat{h}_i = (x_i / \hat{h}_i)' \beta + \varepsilon_i / \hat{h}_i$$

5. The consistent estimate s^2 for σ^2 , based on the FGLS-residuals, and the consistently estimated covariance matrix

$$\hat{V} \{ \hat{\beta}^* \} = s^2 \left(\sum_i \hat{h}_i^{-2} x_i x_i' \right)^{-1}$$

are part of the standard output when regressing the transformed model

Labor Demand Function

For Belgian companies, 1996; Verbeek

Table 4.5 OLS results loglinear model with White standard errors

Dependent variable: $\log(\text{labour})$

Variable	Estimate	Heteroskedasticity-consistent Standard error	<i>t</i> -ratio
constant	6.177	0.294	21.019
$\log(\text{wage})$	-0.928	0.087	-10.706
$\log(\text{output})$	0.990	0.047	21.159
$\log(\text{capital})$	-0.004	0.038	-0.098

$s = 0.465$ $R^2 = 0.8430$ $\bar{R}^2 = 0.8421$ $F = 544.73$

Log-transformation is expected to reduce heteroskedasticity

Labor Demand Function, cont'd

For Belgian companies, 1996; Verbeek

Table 4.6 Auxiliary regression multiplicative heteroskedasticity

Dependent variable: $\log e_i^2$

Variable	Estimate	Standard error	<i>t</i> -ratio
constant	-3.254	1.185	-2.745
$\log(\text{wage})$	-0.061	0.344	-0.178
$\log(\text{output})$	0.267	0.127	2.099
$\log(\text{capital})$	-0.331	0.090	-3.659

$s = 2.241$ $R^2 = 0.0245$ $\bar{R}^2 = 0.0193$ $F = 4.73$

Breusch-Pagan test: $NR^2 = 66.23$, p -value: 1,42E-13

Labor Demand Function, cont'd

For Belgian companies, 1996; Verbeek

Weights estimated assuming multiplicative heteroskedasticity

Table 4.7 EGLS results loglinear model

Dependent variable: $\log(\textit{labour})$

Variable	Estimate	Standard error	<i>t</i> -ratio
constant	5.895	0.248	23.806
$\log(\textit{wage})$	-0.856	0.072	-11.903
$\log(\textit{output})$	1.035	0.027	37.890
$\log(\textit{capital})$	-0.057	0.022	-2.636

$s = 2.509$ $R^2 = 0.9903$ $\bar{R}^2 = 0.9902$ $F = 14401.3$

Labor Demand Function, cont'd

Estimated function

$$\log(\textit{labour}) = \beta_1 + \beta_2 * \log(\textit{wage}) + \beta_3 * \log(\textit{output}) + \beta_4 * \log(\textit{capital})$$

The table shows: OLS estimates without (s.e.) and with White standard errors (White s.e.) as well as FGLS estimates and standard errors

	β_1	β_2	β_3	β_4
OLS coeff	6.177	-0.928	0.990	-0.0037
s.e.	0.246	0.071	0.026	0.0188
White s.e.	0.293	0.086	0.047	0.0377
FGLS coeff	5.895	-0.856	1.035	-0.0569
s.e.	0.248	0.072	0.027	0.0216

Labor Demand Function, cont'd

Some comments:

- Reduction of standard errors in FGLS estimation as compared with heteroskedasticity-robust estimation, efficiency gains
- Comparison with OLS estimation not appropriate
- FGLS estimates differ slightly from OLS estimates; effect of capital is indicated to be relevant (p -value: 0.0086)
- R^2 of FGLS estimation is misleading
 - Model is uncentered, no intercept
 - Comparison with that of OLS estimation not appropriate, explained variable differ

Contents

- Violations of $V\{\varepsilon\} = \sigma^2 I_N$
- Heteroskedasticity
- GLS Estimation
- Autocorrelation

Autocorrelation

- Typical for time series data such as consumption, production, investments, etc., and models for time series data
- Autocorrelation of error terms is typically observed if
 - a relevant regressor with trend or seasonal pattern is not included in the model: miss-specified model
 - the functional form of a regressor is incorrectly specified
 - the dependent variable is correlated in a way that is not appropriately represented in the systematic part of the model
- Autocorrelation of the error terms indicates deficiencies of the model specification such as omitted regressors, incorrect functional form, incorrect dynamic
- Tests for autocorrelation are the most frequently used tool for diagnostic checking the model specification

Example: Ice Cream Demand

Demand for ice cream explained from income and price index

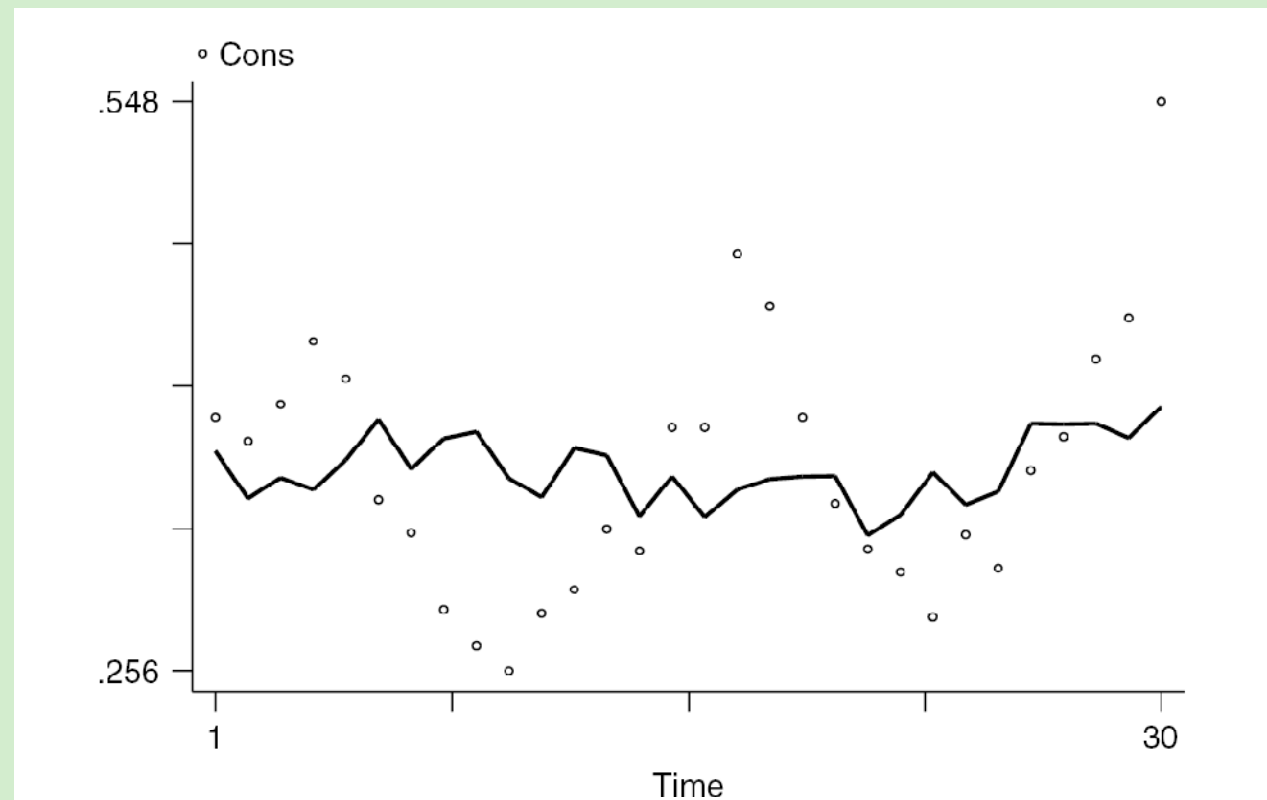


Figure 4.1 Actual and fitted consumption of ice cream, March 1951–July 1953

A Model with AR(1) Errors

Linear regression

$$y_t = x_t' \beta + \varepsilon_t \text{ } ^{1)}$$

with

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t \text{ with } -1 < \rho < 1 \text{ or } |\rho| < 1$$

where v_t are uncorrelated random variables with mean zero and constant variance σ_v^2

- For $\rho \neq 0$, the error terms ε_t are correlated; the Gauss-Markov assumption $V\{\varepsilon\} = \sigma_\varepsilon^2 I_N$ is violated
- The other Gauss-Markov assumptions are assumed to be fulfilled

The sequence ε_t , $t = 0, 1, 2, \dots$ which follows $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$ is called an autoregressive process of order 1 or AR(1) process

1) In the context of time series models, variables are indexed by „t“

Properties of AR(1) Processes

Repeated substitution of ε_{t-1} , ε_{t-2} , etc. results in

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t = v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \dots$$

With v_t being uncorrelated and having mean zero and variance σ_v^2 :

- $E\{\varepsilon_t\} = 0$
- $V\{\varepsilon_t\} = \sigma_\varepsilon^2 = \sigma_v^2(1-\rho^2)^{-1}$

This results from $V\{\varepsilon_t\} = \sigma_v^2 + \rho^2\sigma_v^2 + \rho^4\sigma_v^2 + \dots = \sigma_v^2(1-\rho^2)^{-1}$ for $|\rho| < 1$ as the geometric series $1 + \rho^2 + \rho^4 + \dots$ has the sum $(1-\rho^2)^{-1}$ given that $|\rho| < 1$

- for $|\rho| > 1$, $V\{\varepsilon_t\}$ is undefined
- $\text{Cov}\{\varepsilon_t, \varepsilon_{t-s}\} = \rho^s \sigma_v^2 (1-\rho^2)^{-1}$ for $s > 0$

all error terms are correlated; covariances – and correlations

$\text{Corr}\{\varepsilon_t, \varepsilon_{t-s}\} = \rho^s (1-\rho^2)^{-1}$ – decrease with growing distance s in time

AR(1) Process, cont'd

The covariance matrix $V\{\varepsilon\}$:

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

- $V\{\varepsilon\}$ has a band structure
- Depends only of two parameters: ρ and σ_v^2

Consequences of $V\{\varepsilon\} \neq \sigma^2 I_T$

OLS-estimators b for β

- are unbiased
- are consistent
- have the covariance-matrix

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

- are not efficient estimators, not BLUE
- follow – under general conditions – asymptotically the normal distribution

The estimator $s^2 = e'e/(T-K)$ for σ^2 is biased

For an AR(1)-process ε_t with $\rho > 0$, s.e. from $\sigma^2 (X'X)^{-1}$ underestimates the true s.e.

Inference under Autocorrelation

Covariance matrix of b :

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

Use of $\sigma^2 (X'X)^{-1}$ (the standard output of econometric software) instead of $V\{b\}$ for inference on β may be misleading

Identification of autocorrelation:

- Statistical tests, e.g., Durbin-Watson test

Remedies

- Use of correct variances and standard errors
- Transformation of the model so that the error terms are uncorrelated

Estimation of ρ

Autocorrelation coefficient ρ : parameter of the AR(1) process

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

Estimation of ρ

- by regressing the OLS residual e_t on the lagged residual e_{t-1}

$$r = \frac{\sum_{t=2}^T e_t e_{t-1}}{(T-k)s^2}$$

- estimator is
 - biased
 - but consistent for ρ under weak conditions

Autocorrelation Function

Autocorrelation of order s :

$$r_s = \frac{\sum_{t=s+1}^T e_t e_{t-s}}{(T-k)s^2}$$

- Autocorrelation function assigns r_s to s
- Correlogram: graphical representation of the autocorrelation function

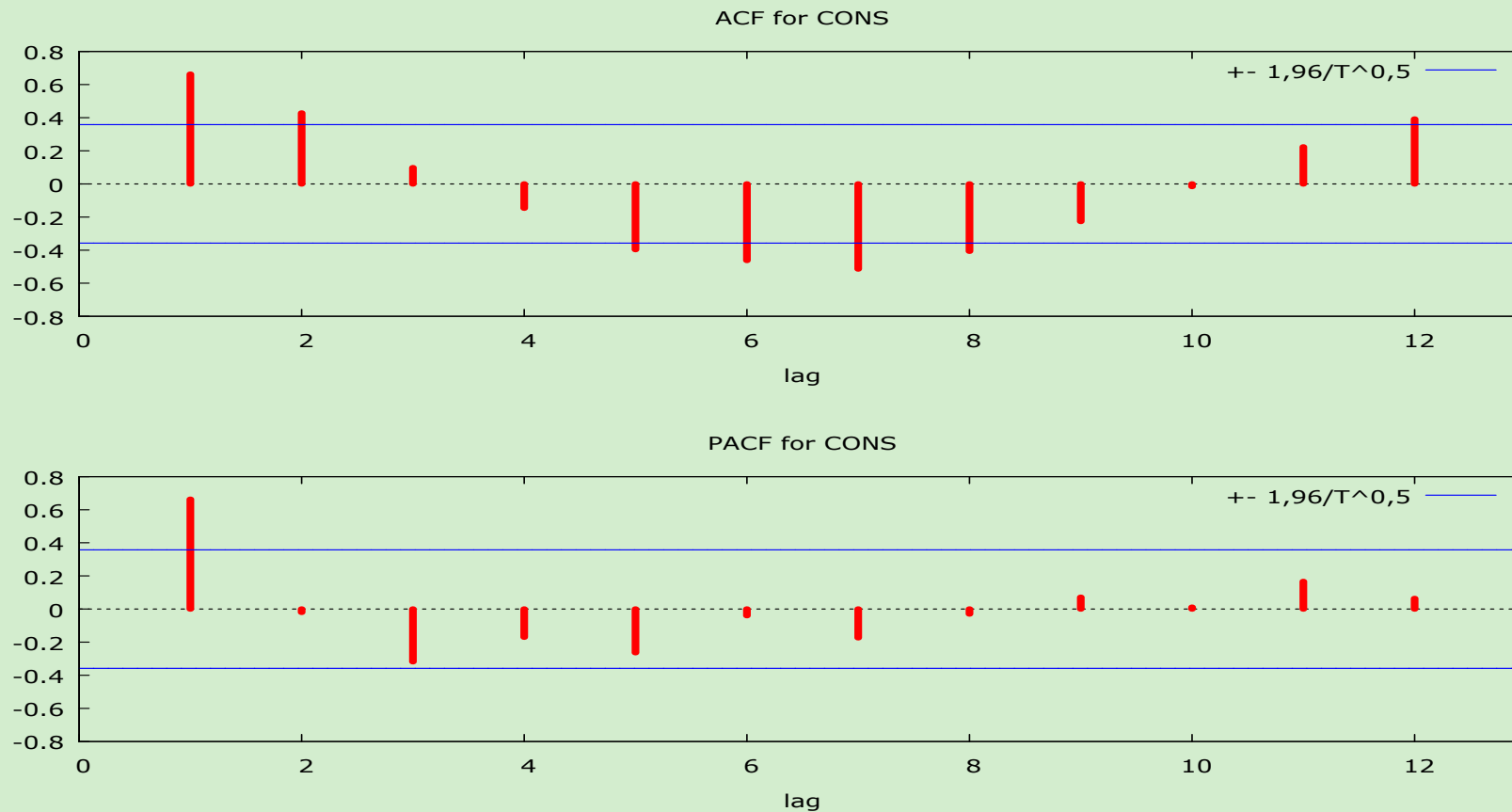
Example: Ice Cream Demand

Autocorrelation function (ACF) of *cons*

LAG	ACF		PACF	Q-stat. [p-value]
1	0,6627 ***		0,6627 ***	14,5389 [0,000]
2	0,4283 **		-0,0195	20,8275 [0,000]
3	0,0982		-0,3179 *	21,1706 [0,000]
4	-0,1470		-0,1701	21,9685 [0,000]
5	-0,3968 **		-0,2630	28,0152 [0,000]
6	-0,4623 **		-0,0398	36,5628 [0,000]
7	-0,5145 ***		-0,1735	47,6132 [0,000]
8	-0,4068 **		-0,0299	54,8362 [0,000]
9	-0,2271		0,0711	57,1929 [0,000]
10	-0,0156		0,0117	57,2047 [0,000]
11	0,2237		0,1666	59,7335 [0,000]
12	0,3912 **		0,0645	67,8959 [0,000]

Example: Ice Cream Demand

Correlogram of *cons*



Tests for Autocorrelation of Error Terms

Due to unbiasedness of b , residuals are expected to indicate autocorrelation

Graphical display, correlogram of residuals may give useful hints

Residual-based tests:

- Durbin-Watson test
- Box-Pierce test
- Breusch-Godfrey test

Asymptotic Tests

AR(1) process for error terms

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

Auxiliary regression of e_t on $x_t'\beta$ and e_{t-1} : produces

■ R_e^2

Test of $H_0: \rho = 0$

1. Breusch-Godfrey test

- R_e^2 of the auxiliary regression: close to zero if $\rho = 0$
- $(T-1) R_e^2$ follows approximately the Chi-square distribution with 1 d.f. if $\rho = 0$
- Lagrange multiplier F (LMF) statistic: F -test for explanatory power of e_{t-1} ; follows approximately the $F(1, T-K-1)$ distribution if $\rho = 0$
- General case of the Breusch-Godfrey test: Auxiliary regression based on higher order autoregressive process

Asymptotic Tests, cont'd

2. Box-Pierce test

- The corresponding t -statistic

$$t = \sqrt{(T)} r$$

follows approximately the t -distribution if $\rho = 0$

- Test based on $\sqrt{(T)} r$ is a special case of the Box-Pierce test which uses the test statistic $Q_m = T \sum_{s=1}^m r_s^2$
- Similar the Ljung-Box test, based on

$$\frac{T(T-2)}{T-1} \sum_{s=1}^m r_s^2$$

which follows the Chi-square distribution with m d.f. if $\rho = 0$

Asymptotic Tests, cont'd

Remarks

- If the model of interest contains lagged values of y the auxiliary regression should also include all explanatory variables (just to make sure the distribution of the test is correct)
- If heteroskedasticity is suspected, White standard errors may be used in the auxiliary regression

Durbin-Watson Test

Test of $H_0: \rho = 0$ against $H_1: \rho \neq 0$

Test statistic

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

- For $\rho > 0$, dw is expected to have a value in $(0,2)$
- For $\rho < 0$, dw is expected to have a value in $(2,4)$
- dw close to the value 2 indicates no autocorrelation of error terms
- Critical limits of dw
 - depend upon x_t 's
 - exact critical value is unknown, but upper and lower bounds can be derived, which depend only of the number of regression coefficients
- Test can be inconclusive

Durbin-Watson Test: Bounds for Critical Limits

Derived by Durbin and Watson

Upper (d_U) and lower (d_L) bounds for the critical limits and $\alpha = 0.05$

T	K=2		K=3		K=10	
	d_L	d_U	d_L	d_U	d_L	d_U
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

- $dw < d_L$: reject H_0
- $dw > d_U$: do not reject H_0
- $d_L < dw < d_U$: no decision (inconclusive region)

Durbin-Watson Test: Remarks

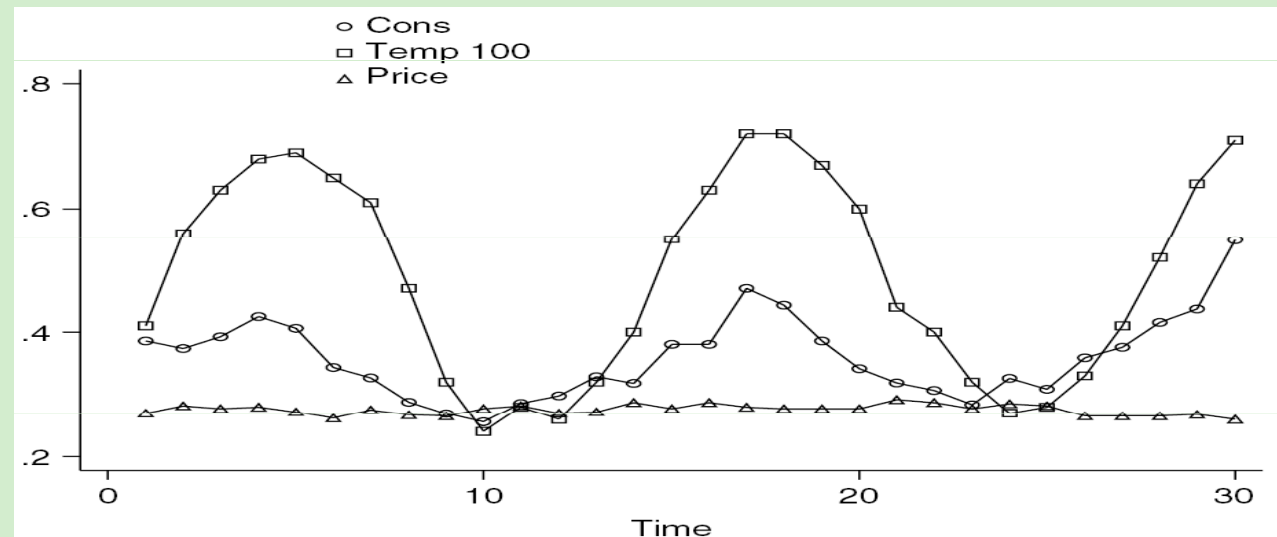
- Durbin-Watson test gives no reference to causes of rejection of the null hypothesis and how the model to modify
- Causes for rejection of the null hypothesis can be various types of misspecification
- Durbin-Watson test is a test against first-order autocorrelation; a test against autocorrelation of other orders may be more suitable, e.g., order four if the model is for quarterly data
- Use of tables unwieldy
 - Limited number of critical bounds (K , T , α) in tables
 - Inconclusive region

Example: Demand for Ice Cream

Time series of 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 (in USD per pint)
- *temp*: average temperature (in Fahrenheit)



Demand for Ice Cream, cont'd

Demand for ice cream, measured by *cons*, explained by *price*, *income*, and *temp*

Table 4.9 OLS results

Dependent variable: *cons*

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

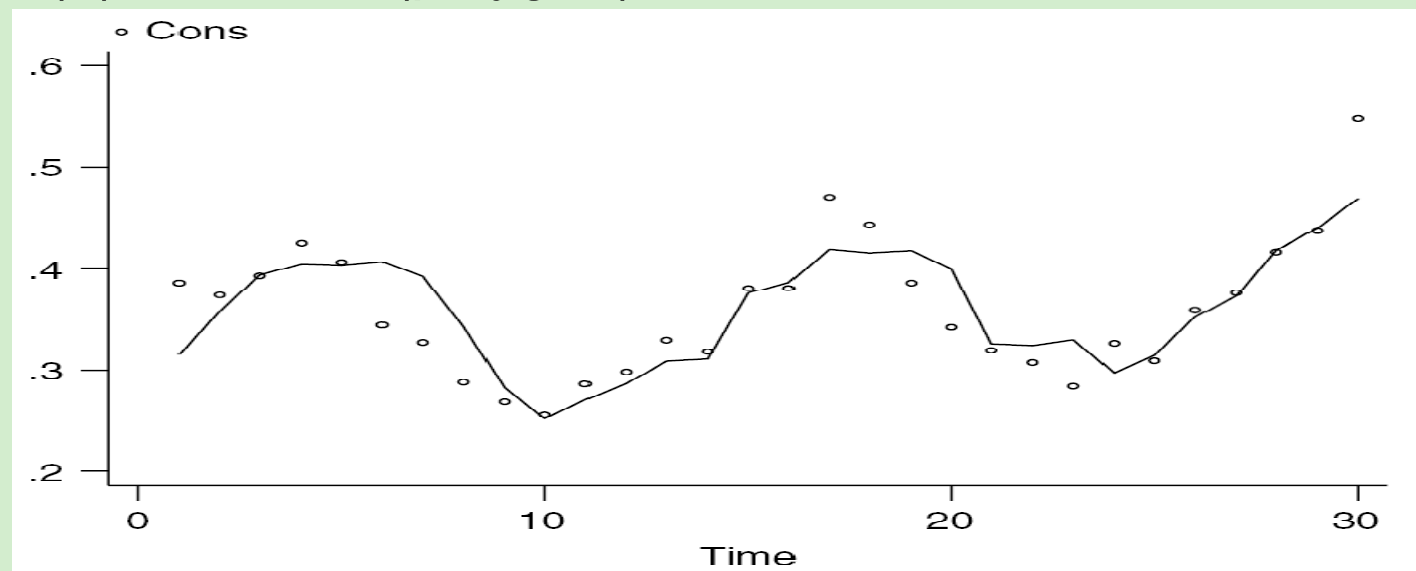
$s = 0.0368$ $R^2 = 0.7190$ $\bar{R}^2 = 0.6866$ $F = 22.175$
 $dw = 1.0212$

Demand for Ice Cream, cont'd

Test for autocorrelation of error terms

- $H_0: \rho = 0, H_1: \rho \neq 0$
- $dw = 1.02 < 1.21 = d_L$ for $T = 30, K = 4$
- Gretl also shows the autocorrelation coefficient: $r = 0.401$

Plot of actual (o) and fitted (polygon) values



Demand for Ice Cream, cont'd

Auxiliary regression $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$: OLS-estimation gives

$$e_t = 0.401 e_{t-1}$$

with $\text{s.e.}(r) = 0.177$, $R^2 = 0.154$

Test of $H_0: \rho = 0$ against $H_1: \rho > 0$

1. Box-Pierce test:

- $t \approx \sqrt{(30)} 0.401 = 2.196$, p -value: 0.018
- t -statistic: 2.258, p -value: 0.016

2. Breusch-Godfrey test

- $(T-1) R^2 = 4.47$, p -value: 0.035

Both reject the null hypothesis

Inference under Autocorrelation

Covariance matrix of b :

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

Use of $\sigma^2 (X'X)^{-1}$ (the standard output of econometric software) instead of $V\{b\}$ for inference on β may be misleading

Remedies

- Use of correct variances and standard errors
- Transformation of the model so that the error terms are uncorrelated

HAC-estimator for $V\{b\}$

Substitution of Ψ in

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

by a suitable estimator

- Newey-West: substitution of $S_x = \sigma^2(X'\Omega X)/T = (\sum_t \sum_s \sigma_{ts} x_t x_s')/T$ 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 $w_j = j/(p+1)$; p , the *truncation lag*, is to be chosen suitably

- The estimator

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

for $V\{b\}$ is called *heteroskedasticity and autocorrelation consistent* (HAC) estimator, the corresponding standard errors are the HAC s.e.

Demand for Ice Cream, cont'd

Demand for ice cream, measured by *cons*, explained by *price*, *income*, and *temp*, OLS and HAC standard errors

	coeff	s.e.	
		OLS	HAC
<i>constant</i>	0.197	0.270	0.288
<i>price</i>	-1.044	0.834	0.881
<i>income</i>	3.308	1.171	1.151
<i>temp</i>	3.458	0.446	0.449

Cochrane-Orcutt Estimator

GLS estimator

- With transformed variables $y_t^* = y_t - \rho y_{t-1}$ and $x_t^* = x_t - \rho x_{t-1}$, also called quasi-differences, the model $y_t = x_t' \beta + \varepsilon_t$ with $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$ can be written as

$$y_t - \rho y_{t-1} = y_t^* = (x_t - \rho x_{t-1})' \beta + v_t = x_t^{*'} \beta + v_t \quad (\text{A})$$

- The model in quasi-differences has error terms which fulfill the Gauss-Markov assumptions
- Given observations for $t = 1, \dots, T$, model (A) is defined for $t = 2, \dots, T$
- Estimation of ρ using, e.g., the auxiliary regression $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$ gives the estimate r ; substitution of r in (A) for ρ results in FGLS estimators for β
- The FGLS estimator is called Cochrane-Orcutt estimator

Cochrane-Orcutt Estimation

In following steps

1. OLS-estimation of b for β from $y_t = x_t'\beta + \varepsilon_t, t = 1, \dots, T$
2. Estimation of r for ρ from the auxiliary regression $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$
3. Calculation of quasi-differences $y_t^* = y_t - ry_{t-1}$ and $x_t^* = x_t - rx_{t-1}$
4. OLS-estimation of β from

$$y_t^* = x_t^*\beta + v_t, t = 2, \dots, T$$

resulting in the Cochrane-Orcutt estimators

Steps 2. till 4. can be repeated: iterated Cochrane-Orcutt estimator

Demand for Ice Cream, cont'd

Iterated Cochrane-Orcutt estimator

Table 4.10 EGLS (iterative Cochrane–Orcutt) results

Dependent variable: *cons*

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

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

Durbin-Watson test: $dw = 1.55$; $d_L = 1.21 < dw < 1.65 = d_U$

Demand for Ice Cream, cont'd

Demand for ice cream, measured by *cons*, explained by *price*, *income*, and *temp*, OLS and HAC standard errors, and Cochrane-Orcutt estimates

	coeff	s.e.		Cochrane-Orcutt	
		OLS	HAC	coeff	se
<i>constant</i>	0.197	0.270	0.288	0.157	0.300
<i>price</i>	-1.044	0.834	0.881	-0.892	0.830
<i>income</i>	3.308	1.171	1.151	3.203	1.546
<i>temp</i>	3.458	0.446	0.449	3.558	0.555

Demand for Ice Cream, cont'd

Model extended by $temp_{-1}$

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

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

Durbin-Watson test: $dw = 1.58$; $d_L = 1.21 < dw < 1.65 = d_U$

Demand for Ice Cream, cont'd

Demand for ice cream, measured by *cons*, explained by *price*, *income*, and *temp*, OLS and HAC standard errors, Cochrane-Orcutt estimates, and OLS estimates for the extended model

	OLS		Cochrane-Orcutt		OLS	
	coeff	HAC	coeff	se	coeff	se
<i>constant</i>	0.197	0.288	0.157	0.300	0.189	0.232
<i>price</i>	-1.044	0.881	-0.892	0.830	-0.838	0.688
<i>income</i>	3.308	1.151	3.203	1.546	2.867	1.053
<i>temp</i>	3.458	0.449	3.558	0.555	5.332	0.670
<i>temp</i> ₋₁					-2.204	0.731

General Autocorrelation Structures

Generalization of model

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

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

Alternative dependence structures of error terms

- Autocorrelation of higher order than 1
- Moving average pattern

Higher Order Autocorrelation

For quarterly data, error terms may develop according to

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

or - more generally - to

$$\varepsilon_t = \gamma_1\varepsilon_{t-1} + \dots + \gamma_4\varepsilon_{t-4} + V_t$$

- ε_t follows an AR(4) process, an autoregressive process of order 4
- More complex structures of correlations between variables with autocorrelation of order 4 are possible than with that of order 1

Moving Average Processes

Moving average process of order 1, MA(1) process

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

- ε_t is correlated with ε_{t-1} , but not with ε_{t-2} , ε_{t-3} , ...
- Generalizations to higher orders

Remedies against Autocorrelation

- Change functional form, e.g., use $\log(y)$ instead of y
- Extend the model by including additional explanatory variables, e.g., seasonal dummies, or additional lags
- Use HAC standard errors for the OLS estimators
- Reformulate the model in quasi-differences (FGLS) or in differences

Your Homework

1. Use the data set “labour2” of Verbeek for the following analyses:
 - a. Estimate the model where $\log labor$ is explained by $\log output$ and $\log wage$; generate displays of the residuals which indicate heteroskedasticity of the error terms.
 - b. Perform the Breusch-Pagan with $h(.) = \exp(.)$ and the White test; explain the tests and compare and interpret the results.
 - c. Compare (i) the OLS and the White standard errors with (ii) HC0 and (ii) HC1 of the estimated coefficients; interpret the results.
 - d. Estimate the model of a., using FGLS and weights obtained in the auxiliary regression of the Breusch-Pagan test in b.; compare the results with that of a.

Your Homework, cont'd

2. Use the data set “icecream” of Verbeek for the following analyses:
 - a. Estimate the model where *cons* is explained by *income* and *temp*; generate displays of the residuals which may indicate autocorrelation of the error terms.
 - b. Use the Durbin-Watson and the Breusch-Godfrey test against autocorrelation; interpret the result.
 - c. Repeat a., using (i) the iterative Cochrane-Orcutt estimation and (ii) OLS estimation of the model in differences; interpret the result.
3. Show that $dw \approx 2 - 2r$