Intermediate Econometrics

IMQF 2023/24

Aleksandra Nojković

Heteroskedastiticy and Autocorrelation

(Verbeek, Chapter 4)

Gauss-Markov conditions and OLS

Recall the *Gauss-Markov conditions* for the linear model

$$y_i = x_i'\beta + \varepsilon_{i,} \qquad (4.1)$$

which state:

(A1) Error terms have mean zero: $E\{\varepsilon_i\}=0$

(A2) *All* error terms are independent of *all* x variables:

$\{\varepsilon_i, \ldots \varepsilon_N\}$ is independent of $\{x_1, \ldots x_N\}$

(A3) All error terms have the same variance (**homoskedasticity**): $V\{\varepsilon_i\} = \sigma^2$.

(A4) The error terms are mutually uncorrelated (**no autocorrelation**): $\text{cov}\{\varepsilon_{i,}, \varepsilon_j\} = 0, \; i \neq j$.

Estimator properties

Under assumptions (A1) and (A2):

1. The **OLS estimator is unbiased.** That is, $E\{b\} = \beta$.

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

2. **The variance of the OLS estimator** is given by

$$V\{b\} = \sigma^2 \left(\Sigma_i \, x_i \, x_i' \right)^{-1} \qquad (2.33)$$

3. And s^2 (see (2.35)) is unbiased for σ^2.
4. The OLS estimator is BLUE: best linear unbiased estimator for β.

Gauss-Markov conditions

- Denoting the *N*-dimensional vector of all error terms by ε, and the entire matrix of explanatory variables by X, the two essential implications of the Gauss-Markov conditions are:

  $$E\{ \varepsilon \mid X\} = 0 \qquad (4.3)$$

  and

  $$V\{\varepsilon \mid X\} = \sigma^2 I, \qquad (4.4)$$

where *I* is the *N*x*N* identity matrix.

- This says: the distribution of error terms given X has **means of zero** and **constant variances** and **zero covariances (**spherical correlation matrix).

Conditional Homoscedasticity and Nonautocorrelation

Disturbances provide no information about each other.

- $\text{Var}[\varepsilon_i \mid \mathbf{X}] = \sigma^2$
- $\text{Cov}[\varepsilon_i, \varepsilon_j \mid \mathbf{X}] = 0$

$$\begin{bmatrix} Var(\varepsilon_1) & Cov(\varepsilon_1,\varepsilon_2) & Cov(\varepsilon_1,\varepsilon_3) & \ldots & Cov(\varepsilon_1,\varepsilon_N) \\ Cov(\varepsilon_2,\varepsilon_1) & Var(\varepsilon_2) & Cov(\varepsilon_2,\varepsilon_3) & \ldots & Cov(\varepsilon_2,\varepsilon_N) \\ Cov(\varepsilon_3,\varepsilon_1) & Cov(\varepsilon_3,\varepsilon_2) & Var(\varepsilon_3) & \ldots & Cov(\varepsilon_3,\varepsilon_N) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ Cov(\varepsilon_N,\varepsilon_1) & Cov(\varepsilon_N,\varepsilon_2) & Cov(\varepsilon_N,\varepsilon_3) & \ldots & Var(\varepsilon_N) \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & 0 & \ldots & 0 \\ 0 & \sigma^2 & 0 & \ldots & 0 \\ 0 & 0 & \sigma^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \sigma^2 \end{bmatrix} = \sigma^2 \mathbf{I}$$

Violation 1: heteroskedasticity

Heteroskedasticity arises if different error terms do not have the same variance. When do we expect this?

- Variances depend **upon one or more explanatory variables** (e.g., firm size);
- Variances evolve over time (time-varying volatility);

Example 1: explaining ***household food expenditures*** *from* ***household income*** *(or total expenditures). For higher income, we expect higher savings but also more uncertainty surrounding savings.*

Example 2: explaining or forecasting daily stock returns.

Violation 1 (exp1): heteroskedasticity

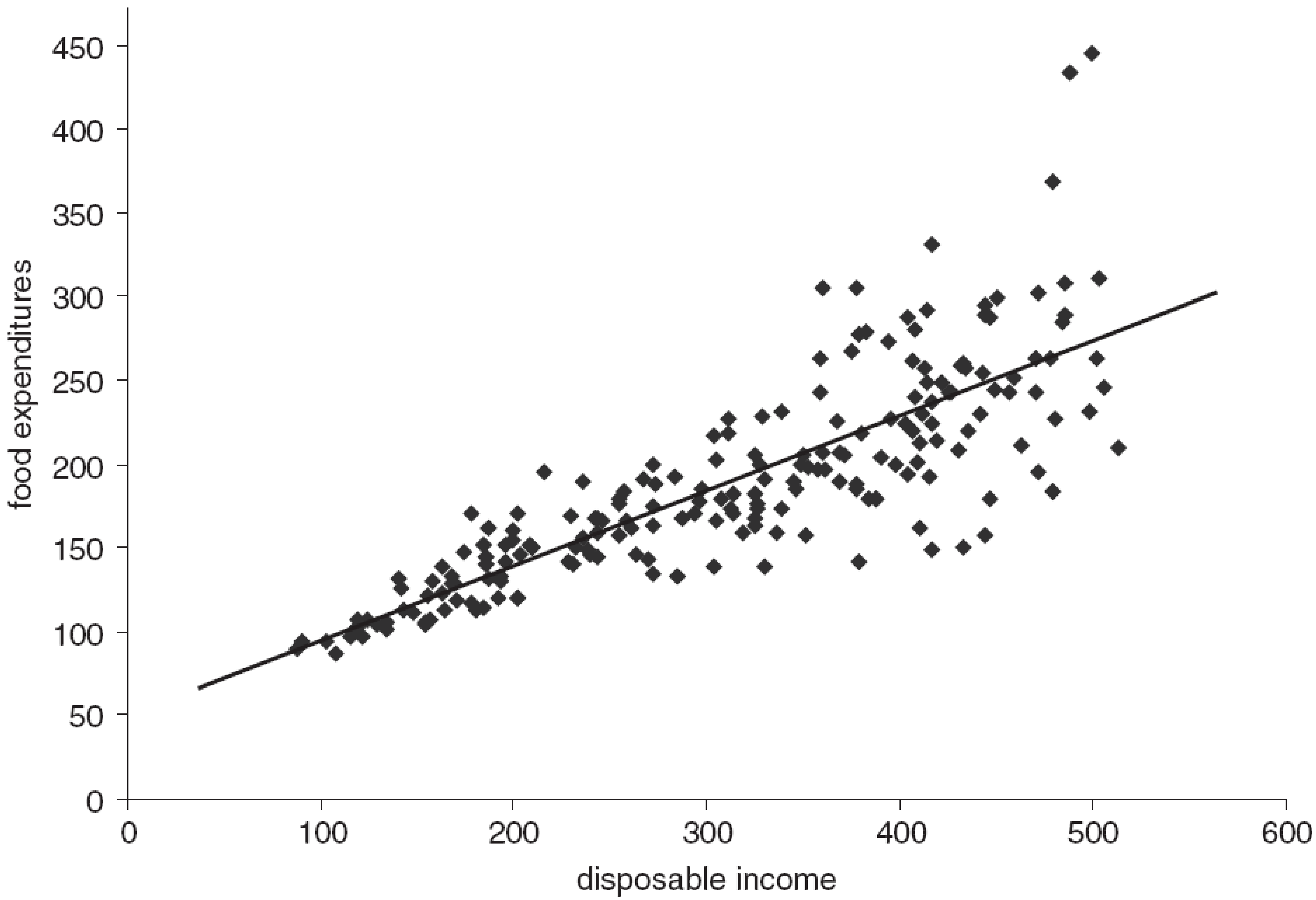
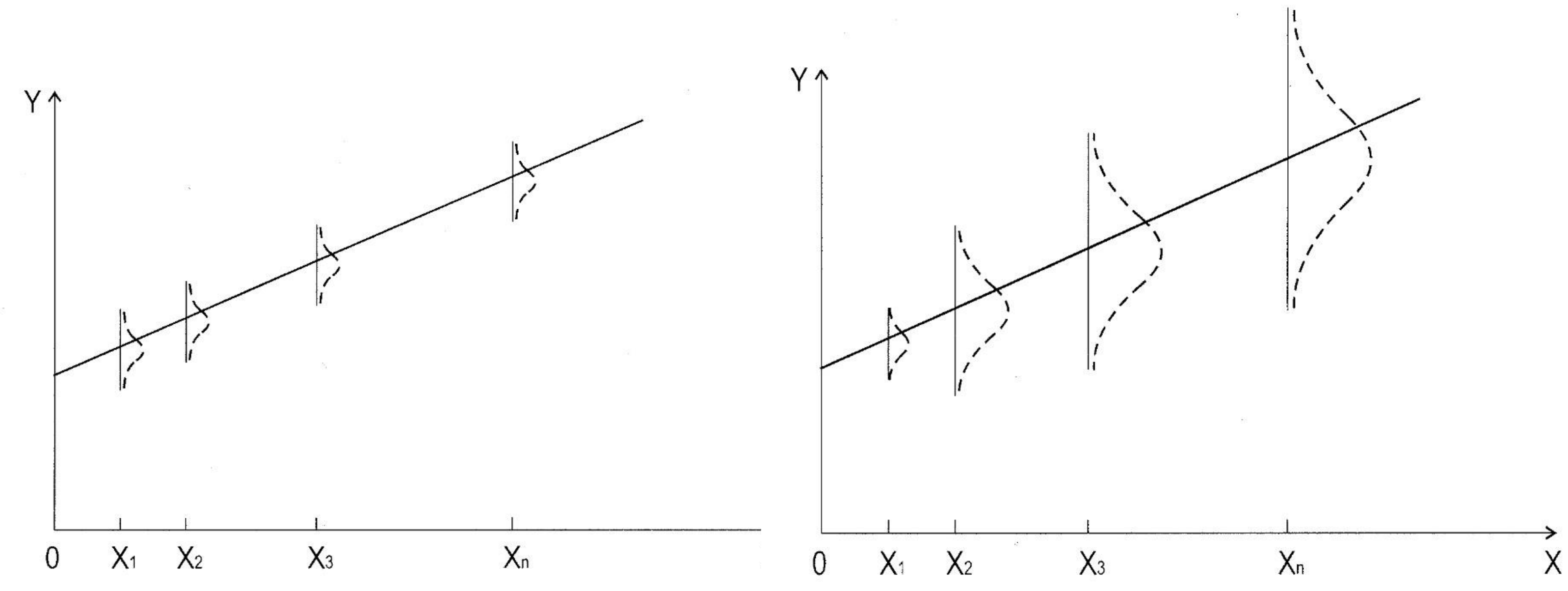


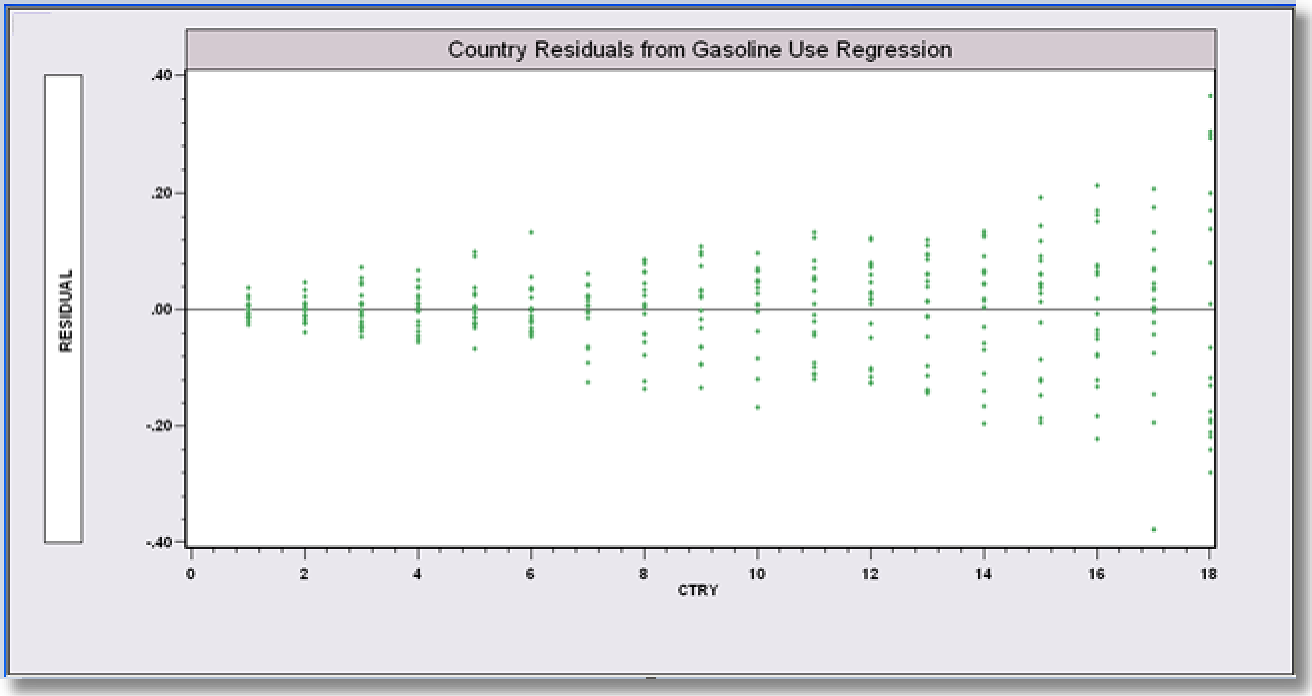
Figure 4.1 An Engel curve with heteroskedasticity

Violation 1 (exp 1): homoscedasticity and heteroscedasticity



Violation1 (exp 1): heteroscedasticity

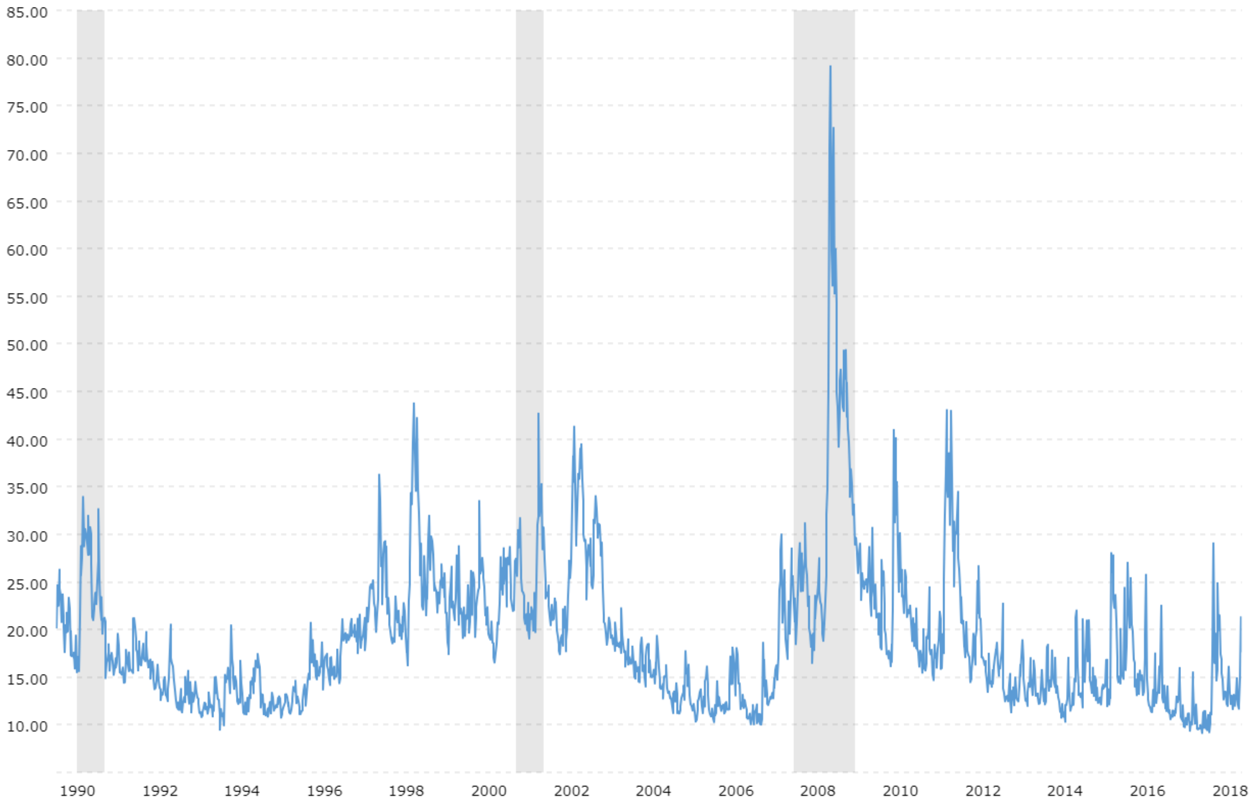
Countries are ordered by the standard deviation of their 19 residuals.



Regression of log of **per capita gasoline use** on log of **per capita income**, gasoline **price** and **number of cars per capita** for 18/30 OECD countries for 19 years. The standard deviation varies by country (source: Greene, 2018)

Violation 1 (exp 2): heteroscedasticity

The daily level of the CBOE VIX Volatility Index back to 1990. The VIX index measures the expectation of stock market volatility over the next 30 days implied by S&P 500 index options



Violation 2: autocorrelation

Autocorrelation (serial correlation) arises if different *error terms* are correlated. This mostly occurs with time-series data

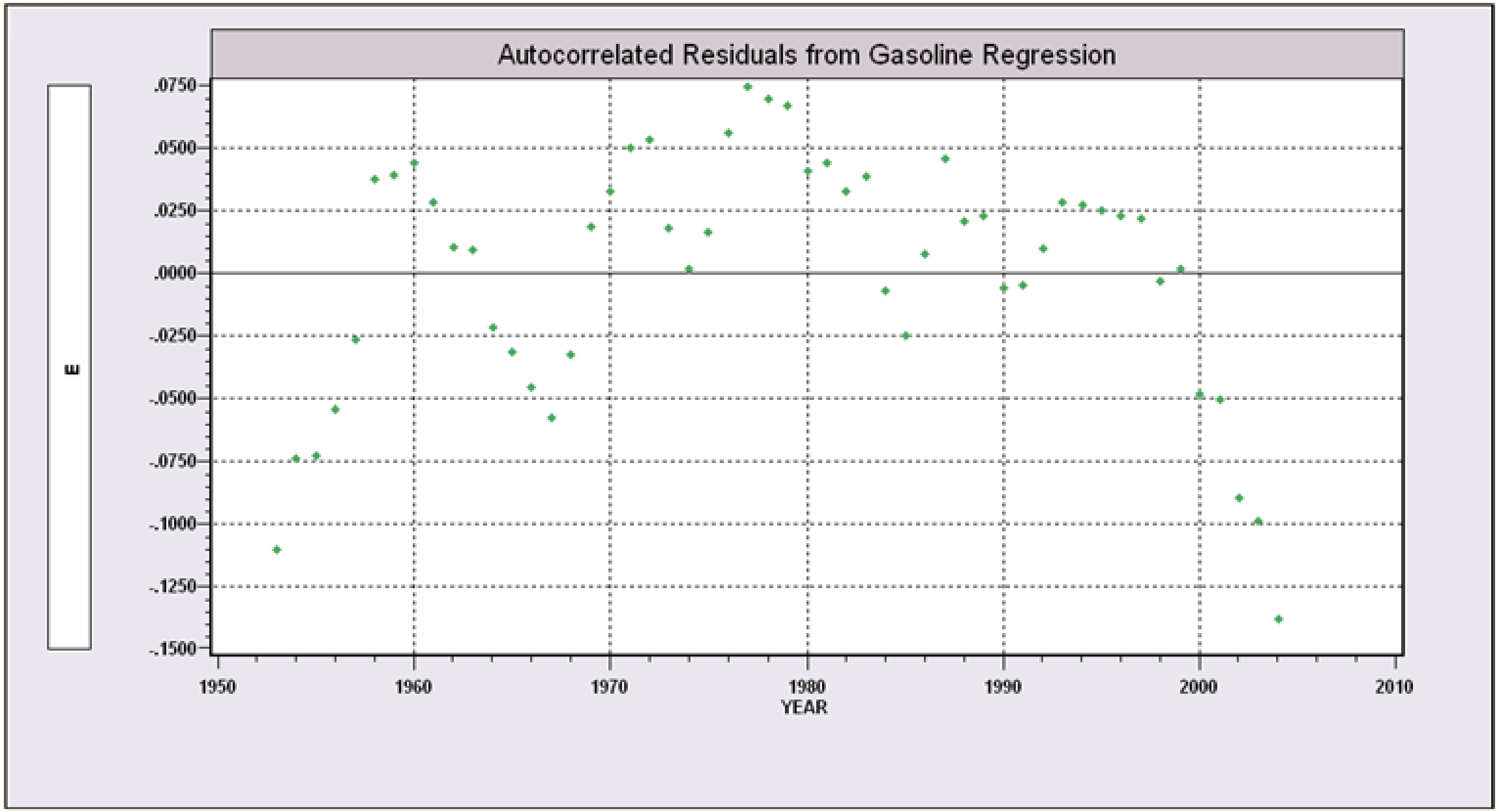
When do we expect this?

- Unobservables (model imperfections) from one period partly carry over to the next
- Model is missing seasonal patterns
- Model is based on overlapping samples (e.g., quarterly returns observed each month)
- **Model is otherwise misspecified (omitted variable, incorrect dynamics, etc.)**

Violation 2: autocorrelation

(source: Greene, 2018)

$\log(G/pop)=\beta_1 + \beta_2 \log Pg + \beta_3 \log(income/pop) + \beta_4 \log Pnc + \beta_5 \log Puc + \varepsilon$



Conditional Homoscedasticity and Nonautocorrelation

Disturbances provide no information about each other.

- $Var[\varepsilon_i \mid \mathbf{X}] = \sigma^2$
- $Cov[\varepsilon_i, \varepsilon_j \mid \mathbf{X}] = 0$

$$\begin{bmatrix} Var(\varepsilon_1) & Cov(\varepsilon_1,\varepsilon_2) & Cov(\varepsilon_1,\varepsilon_3) & ... & Cov(\varepsilon_1,\varepsilon_N) \\ Cov(\varepsilon_2,\varepsilon_1) & Var(\varepsilon_2) & Cov(\varepsilon_2,\varepsilon_3) & ... & Cov(\varepsilon_2,\varepsilon_N) \\ Cov(\varepsilon_3,\varepsilon_1) & Cov(\varepsilon_3,\varepsilon_2) & Var(\varepsilon_3) & ... & Cov(\varepsilon_3,\varepsilon_N) \\ ... & ... & ... & ... & ... \\ Cov(\varepsilon_N,\varepsilon_1) & Cov(\varepsilon_N,\varepsilon_2) & Cov(\varepsilon_N,\varepsilon_3) & ... & Var(\varepsilon_N) \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & 0 & ... & 0 \\ 0 & \sigma^2 & 0 & ... & 0 \\ 0 & 0 & \sigma^2 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \sigma^2 \end{bmatrix} = \sigma^2\mathbf{I}$$

Consequences

The consequences of both problems are similar

As long as (4.3) holds, the OLS **estimator is still unbiased**. However, if (4.4) is violated:

- *OLS is no longer BLUE*
- *Routinely computed **standard errors are incorrect***

This also indicates **three general ways** to deal with the problem:

- *Use alternative standard errors;*
- *Use an alternative estimator (more efficient than OLS); or*
- *Reconsider the employed model.*

The second option is becoming less and less popular, and the third option is often employed with autocorrelation (for heteroscedasticity – try with logs).

Solution 1: Computing robust standard errors

- When $V\{\varepsilon \mid X\}$ is diagonal, but with different diagonal elements, we have heteroskedasticity but no autocorrelation

- Thus, assumption (A3) becomes:

$$V\{\varepsilon_i\} = \sigma^2_i = \sigma^2 h^2_i$$

- Without additional assumptions, it is not possible to estimate σ^2_i . This is because each observation has its own unknown parameter

- Fortunately, it is possible to estimate standard errors for OLS without specifying σ^2_i . This is attributed to (Eicker and White)

Solution 1: Computing robust standard errors (II)

- **The White (heteroskedasticity-consistent) covariance matrix** can be computed from the regressors and the OLS residuals. Its formula is given in (4.30):

$$\left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} e_i^2 x_i x_i' \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1}$$

- If we use this formula to compute standard errors rather than the standard one from (2.36), **we can continue as before with our (t-)tests**. This is appropriate, whether or not the errors have a constant variance. We call this "heteroskedasticity-robust inference"

About White standard errors

- In many cases, using White (heteroskedasticity-consistent) standard errors is appropriate and a good solution to the problem of heteroskedasticity

- They are easily available in most modern software

- It allows one to make appropriate inference without specifying the type of heteroskedasticity

- This is (almost) **standard in many applications** in finance, most prominently **with high-frequency data** (e.g., daily returns)

- Sometimes, we would like to have a **more efficient** estimator, by making **some assumption about the form** of heteroskedasticity

Solution 2: Deriving an alternative estimator

Trick: we know that OLS is BLUE under the Gauss-Markov conditions

1. **Transform the model such that it satisfies** the Gauss-Markov assumptions again

2. Apply OLS to the transformed model

This leads to the ***generalized least squares (GLS) estimator****, which is BLUE.*

- Transformation often depends upon unknown parameters

(characterizing heteroskedasticity or autocorrelation).

3. Estimate them first and transform as before.

This leads to a feasible GLS (FGLS, EGLS) estimator, which is "approximately" BLUE

Solution 2: Deriving an alternative estimator (II)

With heteroskedasticity we have

$V\{\varepsilon_i\} = \sigma^2_i = \sigma^2 h^2_i$.

Then

$y_i / h_i = (x_i / h_i)'\beta + \varepsilon_i / h_i$ (4.16)

has an homoskedastic error term.

OLS applied to this transformed model gives

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

which is a *weighted least squares* estimator.

Solution 2: Deriving an alternative estimator (III)

- The weighted least squares estimator is a least squares estimator where **each observation is weighted** by (a factor proportional to) the inverse of the error variance

- Observations with a higher variance get a lower weight (because they provide less accurate info on β)

- The resulting estimator **is more efficient (more accurate) than OLS**

- However, it can only be applied if we know h_i (we rarely do) or if **we can estimate it by making additional restrictive** assumptions on the form of h_i (we may not like this)

Multiplicative heteroskedasticity

- Assume

$$V\{\varepsilon_i | x_i\} = \sigma_i^2 = \sigma^2 \exp\{\alpha_1 z_{i1} + \cdots + \alpha_J z_{iJ}\} = \sigma^2 \exp\{z_i' \alpha\}$$

where z_i is a function (subset) of x_i. Note that the functional form is such that the variances are never negative.

- To estimate α we run an auxiliary regression

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

where $v_i = \log(e_i^2 / \sigma_i^2)$ is an error term.

- This provides a consistent estimator for α, which can be used to transform the model

Multiplicative heteroskedasticity (II)

- To obtain the EGLS estimator, compute

$$\hat{h}_i^2 = \exp\{z_i'\hat{\alpha}\}$$

and transform all observations to obtain

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

- The error term in this model is (approximately) homoscedastic. Applying OLS to the transformed model gives the **feasible or estimated LS** estimator for β (**FGLS or EGLS**)

- *Note: the transformed regression is for computational purposes only. All economic interpretations refer to the original model!*

The Breusch-Pagan test

- The Breusch-Pagan test tests whether the error variance is a function of z_i . In particular, the alternative hypothesis is

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

for some function h with $h(0)=1$. The null is $\alpha = 0$ (**homoscedasticity**)

- It is based on regressing **the squared OLS residuals upon z_i**. In this case **we choose z_i equal to the original regressors**

- Test statistic: N multiplied by R^2 of the auxiliary regression. Has Chi-squared distribution (DF=dimension of z_i) – **Lagrange multiplier (LM)** test

The White test

- The White test tests whether the error variance is a function of the explanatory variables, with a **more general alternative** than Breusch-Pagan

- It is based on regressing the squared OLS residuals upon all regressors, their squares and their (unique) cross-products

- Test statistic: *N* multiplied by R^2 of the auxiliary regression. Has Chi-squared distribution (DF = # variables in auxiliary regression)

- Advantage: **general**

- Disadvantage: general - **low power in small samples**

Illustration: explaining labor demand

- We estimate a simple labor demand function for a sample of 569 Belgian firms (from 1996)

- We explain labor from output, wage costs and capital stock.

labour: total employment (number of workers);
capital: total fixed assets (in million euro);
wage: total wage costs divided by number of workers (in 1000 euro);
output: value added (in million euro).

- Note that the variables are scaled (to obtain coefficients in the same order of magnitude)

A linear model

Table 4.1 OLS results linear model

Dependent variable: *labour*

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

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

Breusch-Pagan test

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
wage	228.86	302.22	0.757
output	5362.21	214.35	25.015
capital	−3543.51	162.12	−21.858

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

The Breusch-Pagan test

- Striking in this auxiliary regression are the (very) high t-ratios and the high R^2
- This indicates that the squared errors are strongly related to z_i
- Recall that the expected value of ε^2_i should be equal to σ^2 in case of homoskedasticity
- Test statistic: $N \times R^2$, gives 331.0, which provides a very strong rejection!
- This is not uncommon in models like this: assume all firms are "identical", except on a different scale, then we expect the standard deviation of ε_i to be different

A loglinear model

Table 4.3 OLS results loglinear model

Dependent variable: log(*labour*)

Variable	Estimate	Standard error	*t*-ratio
constant	6.177	0.246	25.089
log(*wage*)	−0.928	0.071	−12.993
log(*output*)	0.990	0.026	37.487
log(*capital*)	−0.004	0.019	−0.197

$s = 0.465 \quad R^2 = 0.8430 \quad \bar{R}^2 = 0.8421 \quad F = 1011.02$

The White test

Table 4.4 Auxiliary regression White test

Dependent variable: e_i^2

Variable	Estimate	Standard error	t-ratio
constant	2.545	3.003	0.847
log(*wage*)	−1.299	1.753	−0.741
log(*output*)	−0.904	0.560	−1.614
log(*capital*)	1.142	0.376	3.039
$\log^2$(*wage*)	0.193	0.259	0.744
$\log^2$(*output*)	0.138	0.036	3.877
$\log^2$(*capital*)	0.090	0.014	6.401
log(*wage*)log(*output*)	0.138	0.163	0.849
log(*wage*)log(*capital*)	−0.252	0.105	−2.399
log(*output*)log(*capital*)	−0.192	0.037	−5.197

$s = 0.851 \quad R^2 = 0.1029 \quad \bar{R}^2 = 0.0884 \quad F = 7.12$

The White test

- With an R^2 of 0.1029, this leads to a value for the White test statistic of 58.5, which is highly significant for a Chi-squared with 9 degrees of freedom

- Given the strong rejection, we next **estimate the loglinear model using White standard errors**

- These are standard errors that are **robust to heteroskedasticity**. That is, are correct even if errors are heteroskedastic

- *Note: parameters estimates, and goodness-of-fit measures do not change.* ***Standard errors, and F-test are adjusted***

A loglinear model with White s.e.'s

Table 4.5 OLS results loglinear model with White standard errors

Dependent variable: log(*labour*)

Variable	Estimate	Heteroskedasticity-consistent Standard error	t-ratio
constant	6.177	0.294	21.019
log(*wage*)	−0.928	0.087	−10.706
log(*output*)	0.990	0.047	21.159
log(*capital*)	−0.004	0.038	−0.098

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

Multiplicative heteroskedasticity

Table 4.6 Auxiliary regression multiplicative heteroskedasticity

Dependent variable: $\log e_i^2$

Variable	Estimate	Standard error	t-ratio
constant	−3.254	1.185	−2.745
log(*wage*)	−0.061	0.344	−0.178
log(*output*)	0.267	0.127	2.099
log(*capital*)	−0.331	0.090	−3.659

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

Is form of heteroskedasticity is too restrictive?

Omitted Variables Test
Equation: AUX_REG
Omitted Variables: LOG(WAGE)^2 LOG(CAPITAL)^2 LOG(OUTPUT)^2
Specification: LOG(RES2*RES2) C LOG(WAGE) LOG(CAPITAL) LOG(OUTPUT)
Null hypothesis: LOG(WAGE)^2 LOG(CAPITAL)^2 LOG(OUTPUT)^2 are jointly insignificant

	Value	df	Probability
F-statistic	1.851530	(3, 562)	0.1367
Likelihood ratio	5.596165	3	0.1330

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	27.75637	3	9.252123
Restricted SSR	2836.079	565	5.019609
Unrestricted SSR	2808.323	562	4.997015

LR test summary:

	Value
Restricted LogL	-1264.368
Unrestricted LogL	-1261.570

EGLS loglinear model

Table 4.7 EGLS results loglinear model

Dependent variable: log(*labour*)

Variable	Estimate	Standard error	t-ratio
constant	5.895	0.248	23.806
log(*wage*)	−0.856	0.072	−11.903
log(*output*)	1.035	0.027	37.890
log(*capital*)	−0.057	0.022	−2.636

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

EGLS from EViews

Dependent Variable: LOG(LABOUR)/WEIGHT
Method: Least Squares
Date: 11/22/23 Time: 21:57
Sample: 1 569
Included observations: 569

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1/WEIGHT	5.895357	0.247638	23.80639	0.0000
LOG(WAGE)/WEIGHT	-0.855579	0.071876	-11.90347	0.0000
LOG(OUTPUT)/WEIGHT	1.034611	0.027306	37.88991	0.0000
LOG(CAPITAL)/WEIGHT	-0.056864	0.021576	-2.635531	0.0086

R-squared	0.905087	Mean dependent var	24.03639
Adjusted R-squared	0.904583	S.D. dependent var	8.122786
S.E. of regression	2.509096	Akaike info criterion	4.684727
Sum squared resid	3556.994	Schwarz criterion	4.715264
Log likelihood	-1328.805	Hannan-Quinn criter.	4.696643
Durbin-Watson stat	1.961772		

Remarks

- Comparing Table 4.7 and 4.5, we see that the **efficiency gain is substantial**

- Comparison with Table 4.3 is not appropriate. This table is wrong and misleading

- The coefficient estimates are fairly close to the OLS ones. Note that the effect of capital is now statistically significant

- The R^2 in Table 4.7 is misleading, because

  - it applies to the transformed model (not the original one)

  - is uncentered because **there is no intercept**

- Recall that OLS always provides **higher R^2s** than does GLS

Autocorrelation

- Autocorrelation typically occurs with time series data (where observations have a natural ordering)
- To stress this, we shall index the observations by $t = 1,\dots,T$, rather than $i = 1,..,N$
- The error term picks up the influence of those (many) variables and factors not included in the model.
- If there is some persistence in these factors, (positive) autocorrelation may arise
- Thus, autocorrelation may be an indication of a *misspecified model* (omitted variables, incorrect functional forms, incorrect dynamics)
- Accordingly, **autocorrelation tests** are often interpreted as **misspecification tests**

Positive autocorrelation

Demand for ice cream explained from income and price index

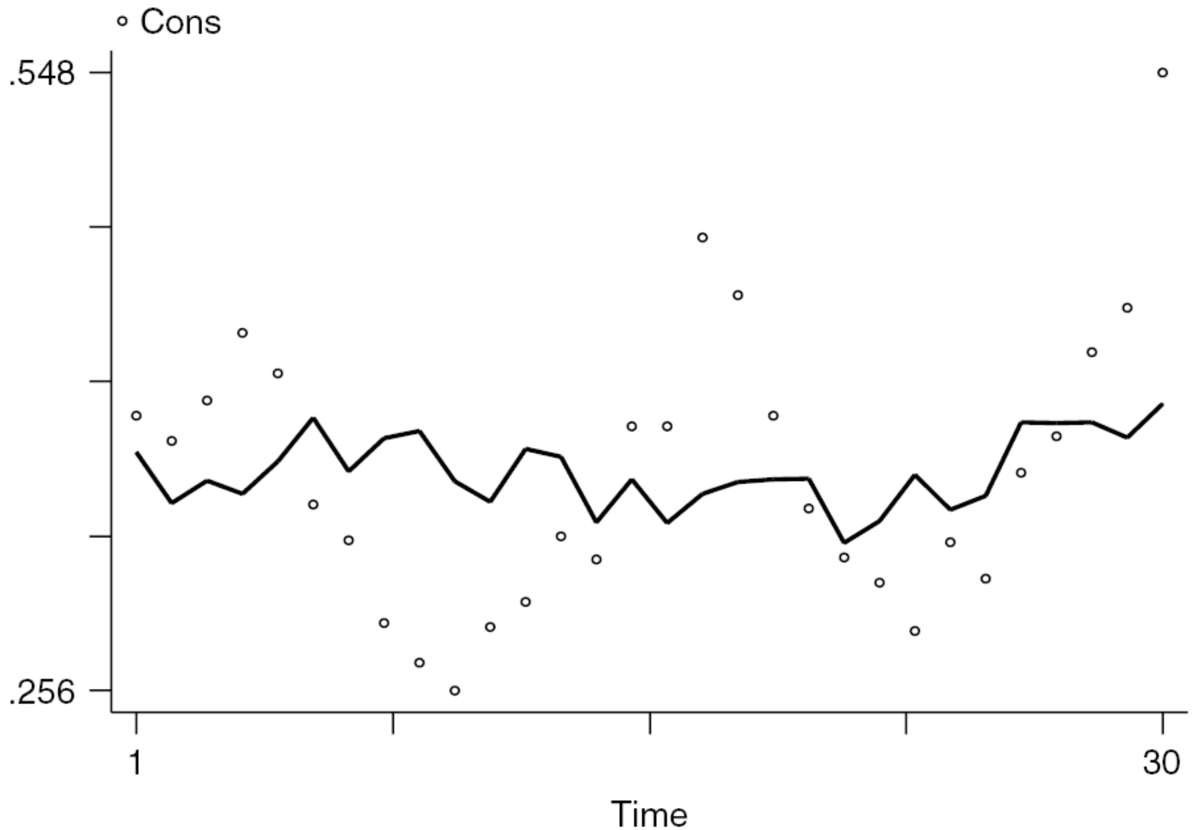
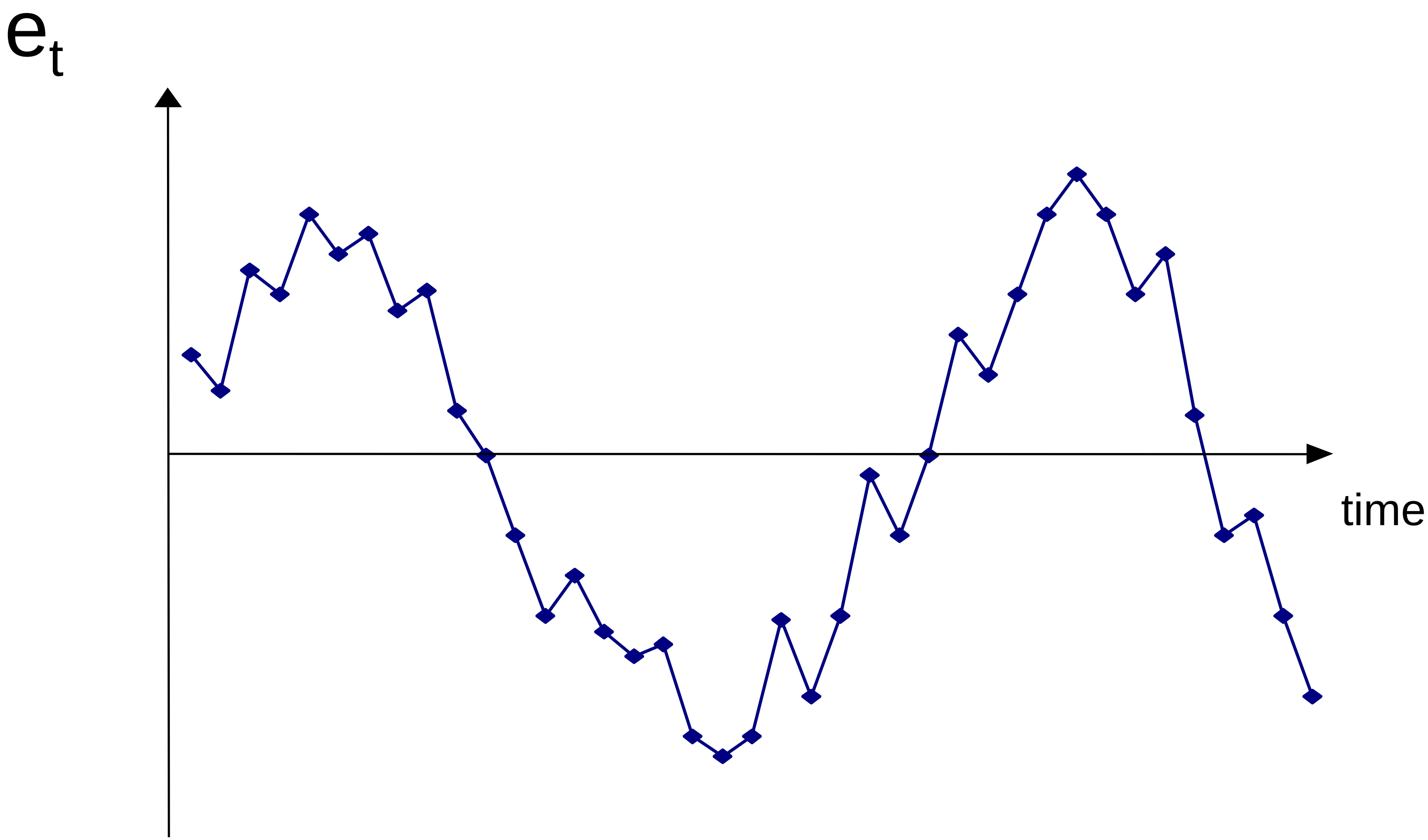
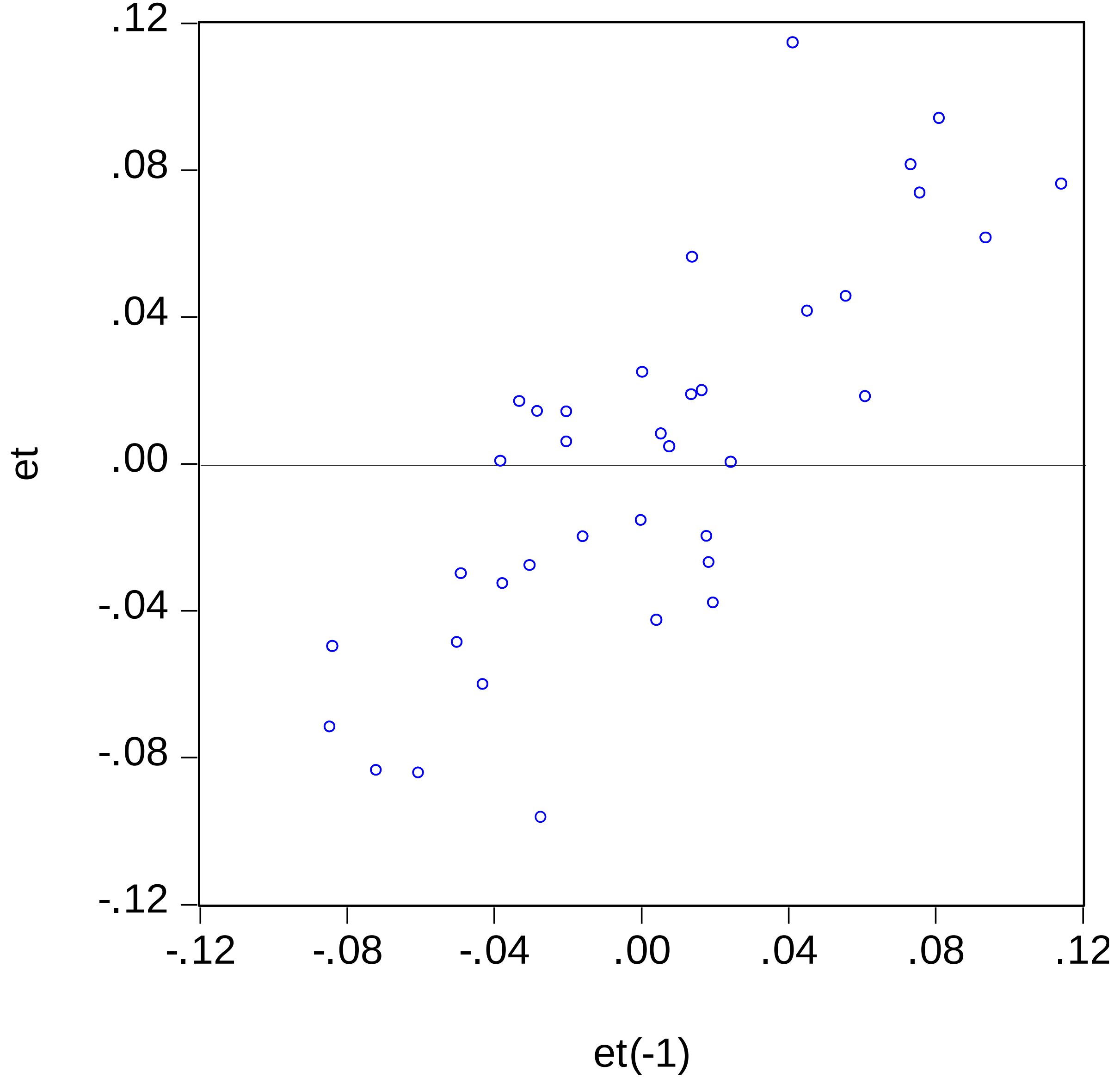


Figure 4.1 Actual and fitted consumption of ice cream, March 1951–July 1953

Violation 2: autocorrelation (positive)



Violation 2: autocorrelation (positive)



First-order autocorrelation

- Many forms of autocorrelation exist. The most popular one is first-order autocorrelation.
- Consider

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

where the error term depends upon his predecessor as

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

where v_t is an error with mean zero and constant variance.

- Assumptions are such that the Gauss-Markov conditions arise if ρ = 0.

Properties of ε_t

To determine the properties of ε_t, we assume | ρ | < 1 (stationarity – see Ch. 8 / forthcoming TSA course).

Then it holds that:

$$E\{\varepsilon_t\} = 0$$

$$V\{\varepsilon_t\} = V\{\rho\varepsilon_{t-1} + v_t\} = \rho^2 V\{\varepsilon_{t-1}\} + \sigma_v^2$$

such that

$$\sigma_\varepsilon^2 = V\{\varepsilon_t\} = \frac{\sigma_v^2}{1-\rho^2}$$

(note that this requires -1 < ρ < 1.)

Properties of ε_t

- Further

$$\text{cov}\{\varepsilon_t, \varepsilon_{t-1}\} = E\{\varepsilon_t \varepsilon_{t-1}\} =$$

$$\rho E\{\varepsilon_{t-1}^2\} + E\{\varepsilon_{t-1} v_t\} = \rho \frac{\sigma_v^2}{1-\rho^2}$$

and

$$E\{\varepsilon_t \varepsilon_{t-2}\} = \rho E\{\varepsilon_{t-1}\varepsilon_{t-2}\} + E\{\varepsilon_{t-2} v_t\} = \rho^2 \frac{\sigma_v^2}{1-\rho^2}$$

and in general

$$E\{\varepsilon_t \varepsilon_{t-s}\} = \rho^s \frac{\sigma_v^2}{1-\rho^2}$$

Solution 1: Computing HAC standard errors

- Similar to the White standard errors for heteroskedasticity, it is also possible to correct OLS standard errors for hetero *and* autocorrelation

- This is typically attributed to Newey and West (HAC- heteroscedasticity-and-autocorrelation-consistent standard errors)

- It is appropriate if the autocorrelation is restricted to a maximum number of lags (so strictly speaking only with moving average errors)

- The number of lags can be chosen by the researcher, although some programmes (e.g., EViews) have standard choices depending upon sample size.

Solution 2: Deriving an alternative estimator- first-order autocorrelation

- Thus, this form of autocorrelation implies that all error terms are correlated. Their covariance decreases if the distance in time gets large

- **To transform the model** such that it satisfies the Gauss-Markov conditions we use

$$y_t - \rho y_{t-1} = (x_t - \rho x_{t-1})'\beta + v_t, \quad t = 2, 3, \ldots, T$$

- With known ρ, this produces (almost) **the GLS estimator**. Note: first observation is lost by this transformation (see p. 115 on how to handle this)

- Of course, typically ρ is unknown

Estimating ρ

- First estimate the original model by OLS. This gives the OLS residuals

- Starting from

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

it seems natural to estimate ρ by regressing the OLS residual e_t upon its lag e_{t-1}. This gives

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

- While this estimator is typically biased, it is consistent for *ρ* under weak conditions

Testing for first-order auto-correlation. 1. Asymptotic tests

- The auxiliary regression producing $\hat{\rho}$ also provides a standard error to it. The resulting t-test statistic is approximately equal to

$$t \approx \sqrt{T}\,\hat{\rho}$$

- We reject the null (no autocorrelation) against the alternative of nonzero autocorrelation if |t| > 1.96 (95% confidence)

- Another form is based on (T-1) x R^2 of this regression, to be compared with Chi-squared distribution with 1 DF (reject if > 3.86)

Testing for first-order auto-correlation. 1. Asymptotic tests

Two remarks:

- If the model of interest contains lagged values of y_t (or other explanatory variables that may be correlated with lagged error terms), the auxiliary regression should also include all explanatory variables (just to make sure the distribution of the test is correct) – special case of Breusch-Godfrey LM test

- If we also suspect heteroskedasticity, White standard errors may be used in the auxiliary regression.

Testing for first-order auto-correlation. 2. Durbin-Watson test

- This is a very popular test, routinely computed by most regression packages (also if it is not appropriate!)

- Requirements: (a) intercept in the model, and (b) assumption (A2), so **no lagged dep. variables!**

- The test statistic is given by

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

which is approximately equal to

$$dw \approx 2 - 2\hat{\rho}$$

Testing for first-order auto-correlation. 2. Durbin-Watson test

- Distribution is “peculiar”.
- Moreover, it depends upon x_t’s.

In general, *dw* values close to 2 are fine, while *dw* values close to 0 imply positive autocorrelation.

The exact critical value is unknown, but **upper and lower bounds can be derived** (see Table 4.8).

Thus (to test for positive autocorrelation):

- *dw* is less than lower bound: reject
- *dw* is larger than upper bound: not reject
- *dw* is in between: inconclusive.

The inconclusive region becomes smaller if T gets large.

Bounds on critical values Durbin-Watson test

Table 4.8 Lower and upper bounds for 5% critical values of the Durbin–Watson test (Savin and White, 1977)

Number of observations	Number of regressors (incl. intercept)							
	$K = 3$		$K = 5$		$K = 7$		$K = 9$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
$T = 25$	1.206	1.550	1.038	1.767	0.868	2.012	0.702	2.280
$T = 50$	1.462	1.628	1.378	1.721	1.291	1.822	1.201	1.930
$T = 75$	1.571	1.680	1.515	1.739	1.458	1.801	1.399	1.867
$T = 100$	1.634	1.715	1.592	1.758	1.550	1.803	1.506	1.850
$T = 200$	1.748	1.789	1.728	1.810	1.707	1.831	1.686	1.852

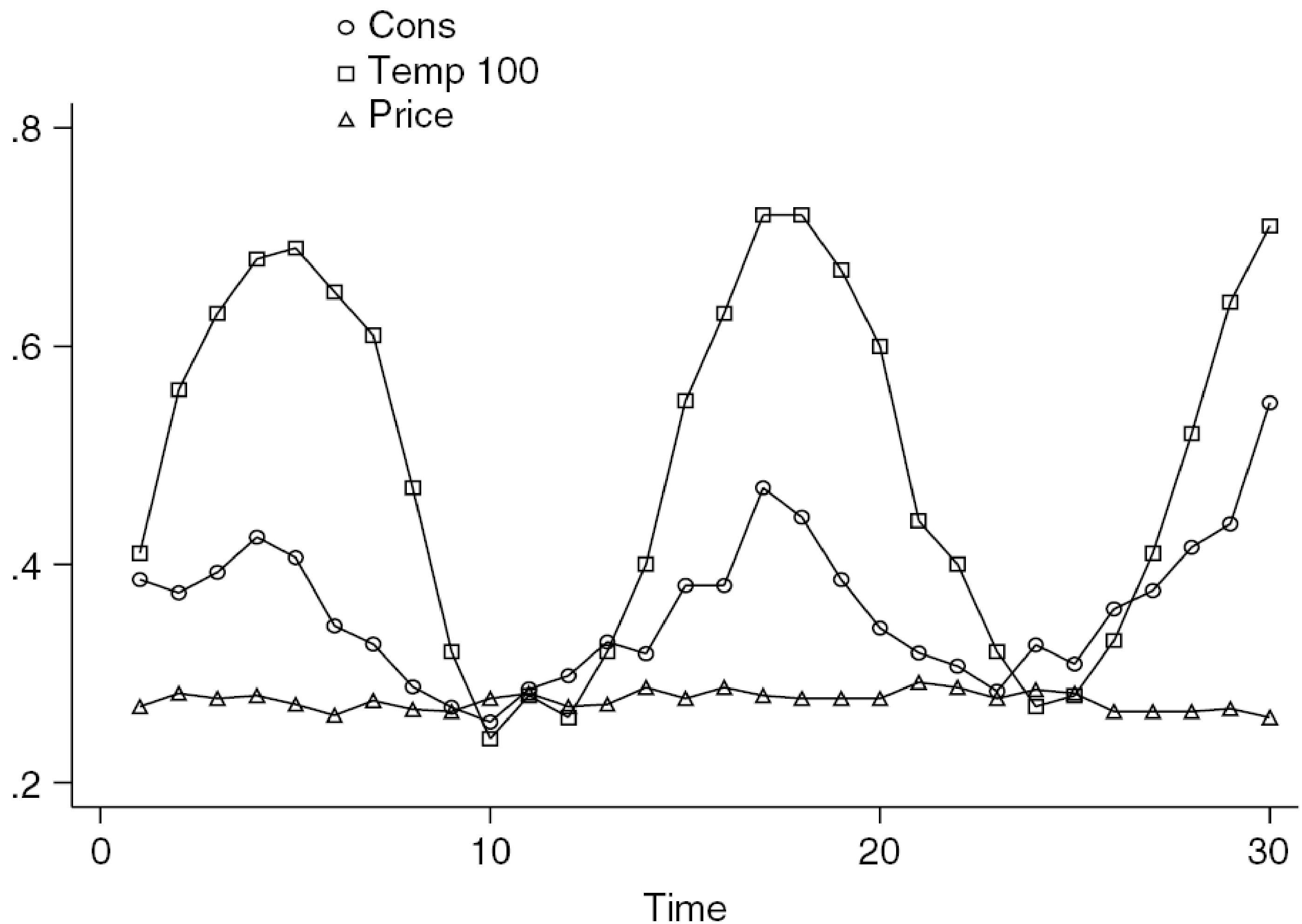
Illustration: the demand for ice cream

Based on classic article Hildreth and Lu (1960), based on a time-series of 30 (!) four-weekly observations 1951-1953.

cons: consumption of ice cream per head (in pints);
income: average family income per week (in US Dollars);
price: price of ice cream (per pint);
temp: average temperature (in Fahrenheit).

See Figure 4.2 for three of these series

Illustration: the demand for ice cream



OLS results

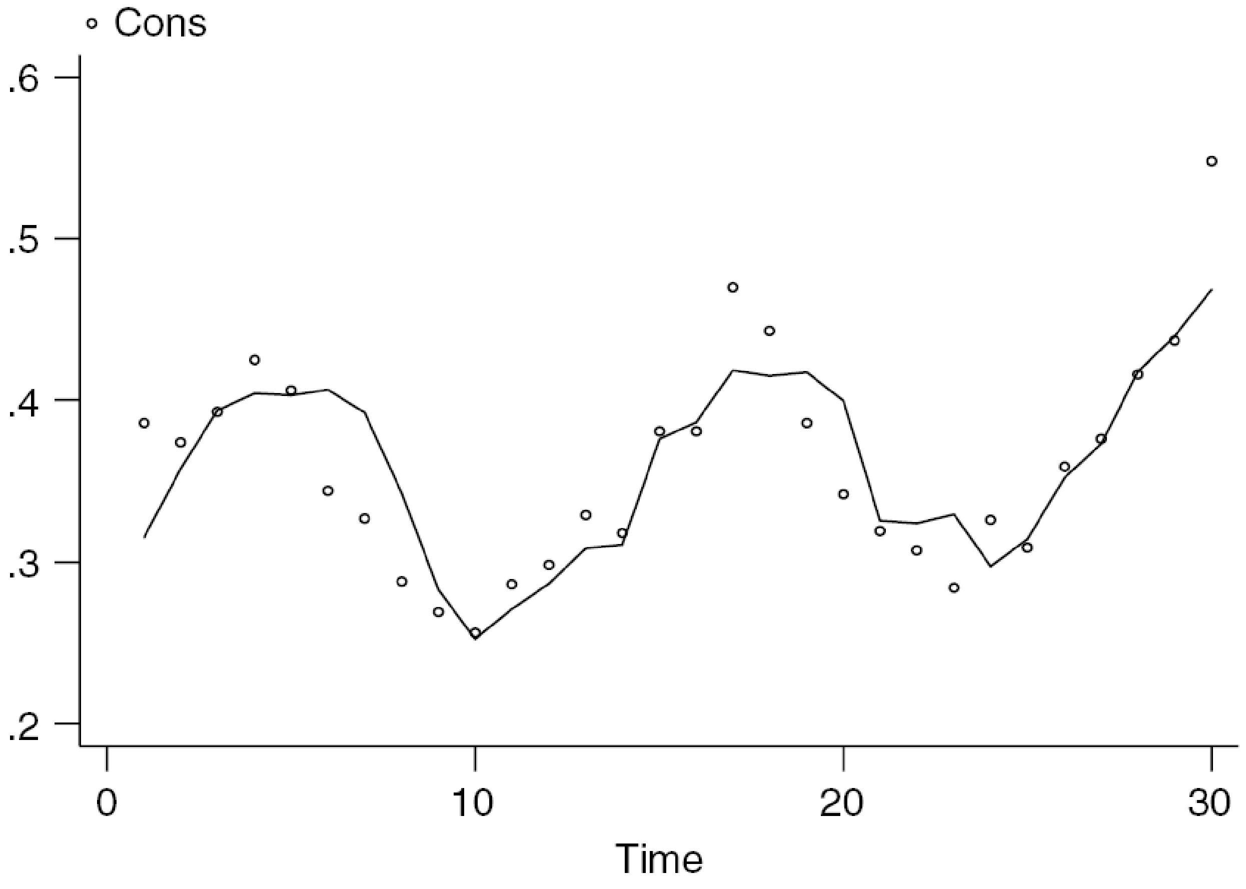
Table 4.9 OLS results

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

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

Actual and fitted values



Estimating ρ

- Regressing the OLS residuals upon their lag gives

$$\hat{\rho} = 0.401$$

$$\text{with an } R^2 \text{ of } 0.149$$

- This gives test statistics:

$$\sqrt{T}\hat{\rho} = 2.19$$

$$(T - 1)R^2 = 4.32$$

- Both reject the null of no autocorrelation. EGLS or change model specification?

EGLS

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

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

Note: starred statistics are for the transformed model.

Alternative model with lagged temperature

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

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

Testing for first-order auto-correlation. 2a. Durbin-Watson test (in models with lagged dep. var.)

- Durbin h- statistics:

$$h = \left(1 - \frac{d}{2}\right)\sqrt{\frac{T}{1 - Ts_{b2}^2}},$$

where S_{b2} is estimated standard error of b2 (estimated coefficient corresponding to the lagged dependent variable), d is DW statistics

- h - statistic has normal standardized dis. (you reject null hypothesis of no autocorrelation if $|h| > 1.96$.

Alternative autocorrelation patterns

- Consider

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

with first order (autoregressive) autocorrelation

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

- This implies that all errors are correlated with each other, with correlations becoming smaller if they are further apart.

Two alternatives:

1. higher order patterns;
2. moving average patterns.

Higher order autocorrelation

- With quarterly or monthly (macro) data, higher order patterns are possible (due to a periodic effect). For example, with quarterly data:

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

or, more generally

$$\varepsilon_t = \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-2} + \gamma_3 \varepsilon_{t-3} + \gamma_4 \varepsilon_{t-4} + v_t,$$

known as 4^{th} order (autoregressive) autocorrelation

- Correlations between different error terms are more flexible than with 1^{st} order

Breusch–Godfrey test

- Baseline regression:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

- Auxiliary regression:

$$e_t = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \rho_1 e_{t-1} + \rho_2 e_{t-2} + ... + \rho_m e_{t-m} + v_t,$$

- Null hypothesis: : $\rho_1 = \rho_2 = ... = \rho_m = 0$

- Test statistic: T multiplied by R^2 of the auxiliary regression. Has Chi-squared distribution (DF=dimension of *m*) – **Lagrange multiplier (LM)** test

- **Works in models with lagged dep. variable**

Moving average autocorrelation

- Arises if the correlation between different error terms is limited by a maximum time lag

- Simplest case (1st order):

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

- This implies that ε_t is correlated with ε_{t-1} , but not with ε_{t-2} or ε_{t-3}, etc.

- Moving average errors arise by construction when "overlapping samples" are used (see Illustration in Section 4.11)

What to do when you find autocorrelation?

In the preferred order:

1. Reconsider the model:

    1a: change functional form (e.g., use log(x) rather than x), see Figure 4.5.

    1b: extend the model by including additional explanatory variables (seasonals) or additional lags;

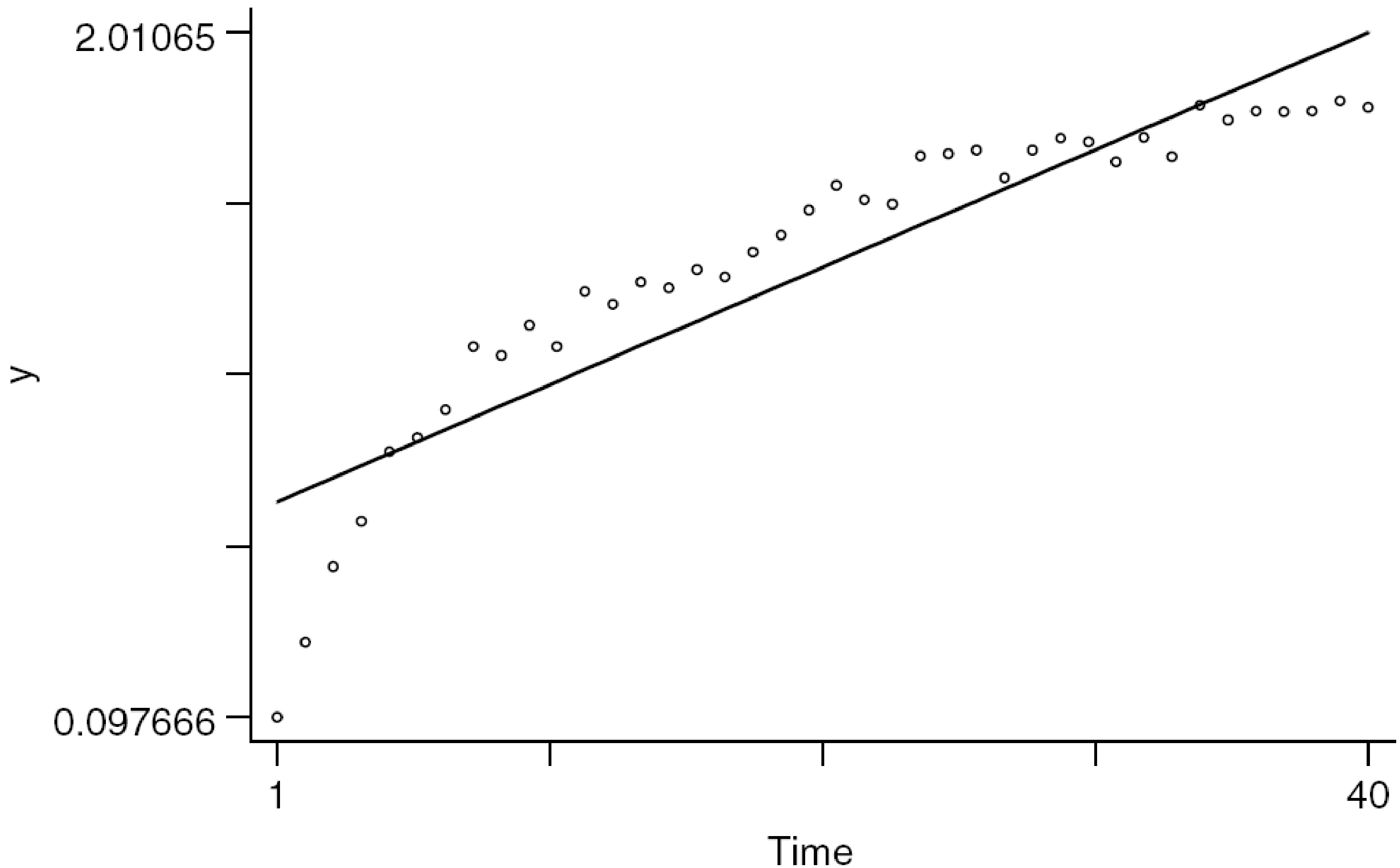
2. Compute heteroskedasticity-and-autocorrelation consistent standard errors (HAC standard errors) for the OLS estimator;

3. Reconsider options 1 and 2;

if you are sure:

4. Use EGLS with existing model.

Wrong functional form



Incomplete dynamics

- Consider the model

  $y_t = x_t'\beta + \varepsilon_t$

which describes $E\{y_t \mid x_t\} = x_t'\beta$, even if $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$

- However, it also describes

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

- Accordingly, we can also write the linear model

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

where the error term does not exhibit serial correlation

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