Introduction to Econometrics
Chapter 6

Ezequiel Uriel Jiménez
University of Valencia

Valencia, September 2013

6 Relaxing the assumptions in the linear classical model

6.2 Misspecification

TABLE 6.1. Summary of bias in $\tilde{\beta}_2$ when x_2 is omitted in estimating equation.

	$Corr(x_2,x_3)>0$	$Corr(x_2,x_3)<0$
$\beta_3>0$	Positive bias	Negative bias
$\beta_3<0$	Negative bias	Positive bias

6.2 Misspecification

EXAMPLE 6.1 Misspecification in a model for determination of wages (file wage06sp)

Initial model $wage = \beta_1 + \beta_2 educ + \beta_3 tenure + u$

$$\widehat{wage}_i = \underset{(1.55)}{4.679} + \underset{(0.146)}{0.681} educ_i + \underset{(0.071)}{0.293} tenure_i$$

$$R^2_{init} = 0.249 \quad n = 150$$

Augmented model $wage = \beta_1 + \beta_2 educ + \beta_3 tenure + \alpha_1 \widehat{wage}^2 + \alpha_1 \widehat{wage}^3 + u$

$$R^2_{augm} = 0.289$$

$$F = \frac{(R^2_{augm} - R^2_{init})/r}{(1 - R^2_{augm})/(n-h)} = 4.18$$

6.3 Multicollinearity

EXAMPLE 6.2 Analyzing multicollinearity in the case of labor absenteeism (file absent)

TABLE 6.2. Tolerance and *VIF*.

	Collinearity statistics	
	Tolerance	***VIF***
age	0.2346	42.634
tenure	0.2104	47.532
wage	0.7891	12.673

6.3 Multicollinearity

EXAMPLE 6.3 Analyzing the multicollinearity of factors determining time devoted to housework (file timuse03)

$$houswork = \beta_1 + \beta_2 educ + \beta_3 hhinc + \beta_4 age + \beta_5 paidwork + u$$

$$\kappa = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} = \sqrt{\frac{542.14}{7.06E-06}} = 8782$$

TABLE 6.3. Eigenvalues and variance decomposition proportions.

Eigenvalues	7.03E-06	0.000498	0.025701	1.861396	542.1400

Variance decomposition proportions

	Associated Eigenvalue				
Variable	1	2	3	4	5
C	0.999995	4.72E-06	8.36E-09	1.23E-13	1.90E-15
EDUC	0.295742	0.704216	4.22E-05	2.32E-09	3.72E-11
HHINC	0.064857	0.385022	0.209016	0.100193	0.240913
AGE	0.651909	0.084285	0.263805	5.85E-07	1.86E-08
PAIDWORK	0.015405	0.031823	0.007178	0.945516	7.80E-05

6.4 Normality test

EXAMPLE 6.4 Is the hypothesis of normality acceptable in the model to analyze the efficiency of the Madrid Stock Exchange? (file bolmadef)

n=247

TABLE 6.4. Normality test in the model on the Madrid Stock Exchange.

skewness coefficient	*kurtosis coefficient*	*Bera and Jarque statistic*
-0.0421	4.4268	21.0232

6.5 Heteroskedasticity

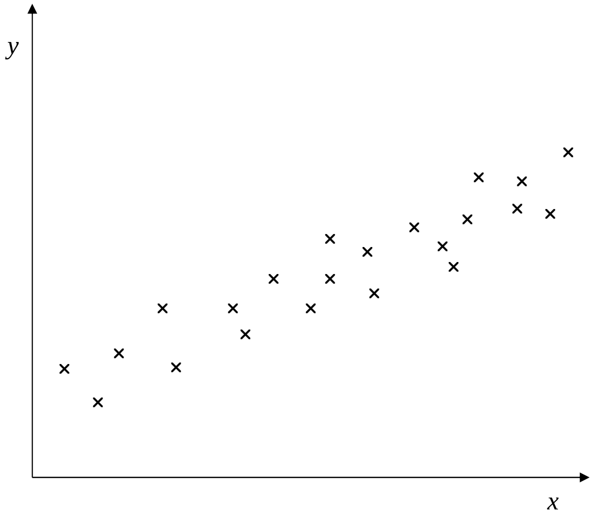


FIGURE 6.1. Scatter diagram corresponding to a model with homoskedastic disturbances.

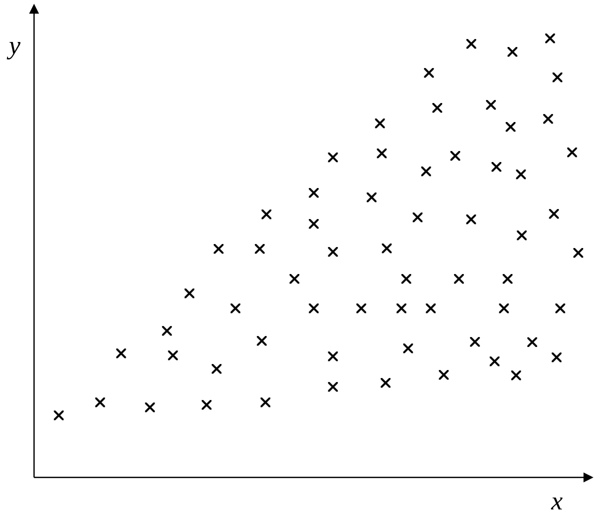


FIGURE 6.2. Scatter diagram corresponding to a model with heteroskedastic disturbances.

EXAMPLE 6.5 Application of the Breusch-Pagan-Godfrey test

TABLE 6.5. *Hostel* and *inc* data.

i	*hostel*	*inc*
1	17	500
2	24	700
3	7	250
4	17	430
5	31	810
6	3	200
7	8	300
8	42	760
9	30	650
10	9	320

Step 1. Applying *OLS* to the model, $hostel = \beta_1 + \beta_2 inc + u$

using data from table 6.5, the following estimated model is obtained:

$$\widehat{hostel_i} = \underset{(3.48)}{-7.427} + \underset{(0.0065)}{0.0533}\, inc_i$$

The residuals corresponding to this fitted model appear in table 6.6.

EXAMPLE 6.5 Application of the Breusch-Pagan-Godfrey test. (Cont.)

TABLE 6.6. Residuals of the regression of *hostel* on *inc*.

i	1	2	3	4	5	6	7	8	9	10
$\hat{u}_i$	-2.226	-5.888	1.1	1.505	-4.751	-0.234	-0.565	8.913	2.777	-0.631

Step 2. The auxiliary regression

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 inc_i + \eta_i$$

$$\hat{u}_i^2 = -23.93 + 0.0799 inc \qquad R^2 = 0.5045$$

Step 3. The *BPG* statistics is:

$$BPG = nR_{ar}^2 = 10(0.56) = 5.05$$

Step 4. Given that $\chi_1^{2(0.05)}$ =3.84, the null hypothesis of homoskedasticity is rejected for a significance level of 5%, but not for the significance level of 1%.

EXAMPLE 6.6 Application of the *White* test

Step 1. This step is the same as in the Breusch-Pagan-Godfrey test.

Step 2. The regressors of the auxiliary regression will be

$$\psi_{1i} = 1 \qquad \forall i$$

$$\psi_{2i} = 1 \times inc_i$$

$$\psi_{3i} = inc_i^2$$

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 inc_i + \alpha_3 inc_i^2 + \eta_i$$

$$\hat{u}_i^2 = 14.29 - 0.10 inc_i + 0.00018 inc_i^2 \qquad R^2 = 0.56$$

Step 3. The *W* statistic:

$$W = nR^2 = 10(0.56) = 5.60$$

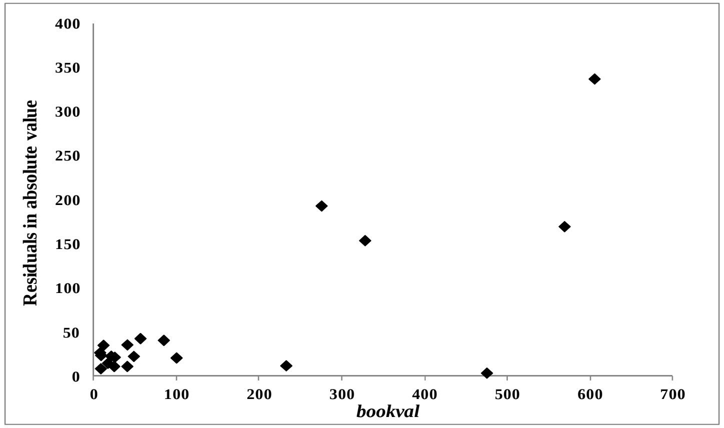
Step 4. Given that $\chi_2^{2(0.10)}$ =4.61, the null hypothesis of homoskedasticity is rejected for a 10% significance level because $W=nR^2$>4.61, but not for significance levels of 5% and 1%.

6.5 Heteroskedasticity

EXAMPLE 6.7 Heteroskedasticity tests in models explaining the market value of the Spanish banks (file bolmad95)

Heteroskedasticity in the linear model

$marktval = \beta_1 + \beta_2 bookval + u \qquad \widehat{marktval} = \underset{(30.85)}{29.42} + \underset{(0.127)}{1.219} bookval \qquad n = 20$



GRAPHIC 6.1. Scatter plot between the residuals in absolute value and the variable *bookval* in the linear model.

$$BPG = nR^2_{ar} = 20 \times 0.5220 = 10.44$$

As $\chi_1^{2(0.01)}$ =6.64<10.44, the null hypothesis of homoskedasticity is rejected for a significance level of 1%, and therefore for α=0.05 and for α =0.10.l

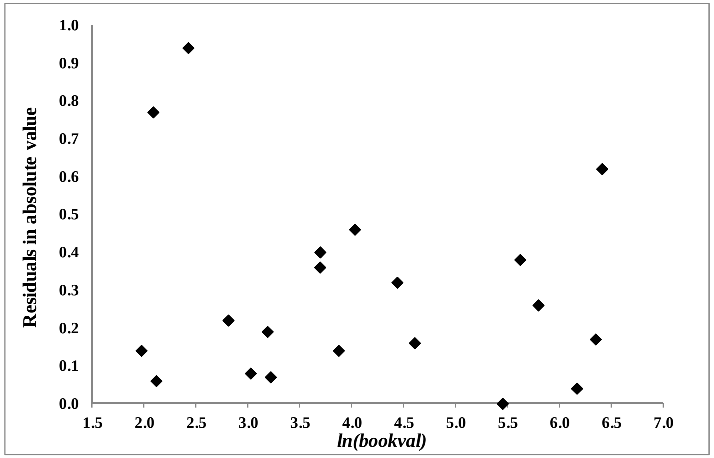
$$W = nR^2_{ar} = 20 \times 0.6017 = 12.03$$

As $\chi_2^{2(0.01)}$ =9.21<12.03, the null hypothesis of homoskedasticity is rejected for a significance level of 1%.

EXAMPLE 6.7 Heteroskedasticity tests in models explaining the market value of the Spanish banks (Cont.)

Heteroskedasticity in the log-log model

$$\widehat{\ln(marktval)} = \underset{(0.265)}{0.676} + \underset{(0.062)}{0.9384}\ln(bookval)$$



GRAPHIC 6.2. Scatter plot between the residuals in absolute value and the variable *bookval* in the log-log model.

TABLE 6.7. Tests of heteroskedasticity on the log-log model to explain the market value of Spanish banks.

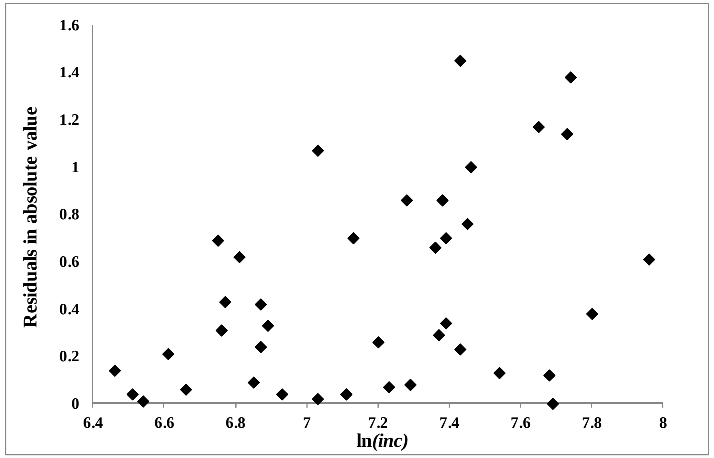
Test	*Statistic*	*Table values*
Breusch-Pagan	$BP = nR^2_{ra}$ =1.05	$\chi_2^{2(0.10)}$ =4.61
White	$W = nR^2_{ra}$ =2.64	$\chi_2^{2(0.10)}$ =4.61

EXAMPLE 6.8 Is there heteroskedasticity in demand of hostel services? (file hostel)

$$\ln(hostel) = \beta_1 + \beta_2 \ln(inc) + \beta_3 secstud + \beta_4 terstud + \beta_5 hhsize + u$$

$$\widehat{\ln(hostel)}_i = \underset{(2.26)}{-16.37} + \underset{(0.324)}{2.732}\ln(inc)_i + \underset{(0.258)}{1.398}\, secstud_i + \underset{(0.333)}{2.972}\, terstud_i - \underset{(0.088)}{0.444}\, hhsize_i$$

$$R^2 = 0.921 \quad n = 40$$



GRAPHIC 6.3. Scatter plot between the residuals in absolute value and the variable ln(*inc*) in the hostel model.

TABLE 6.8. Tests of heteroskedasticity in the model of demand for hostel services.

Test	*Statistic*	*Table values*
Breusch-Pagan-Godfrey	$BPG = nR^2_{ra} = 7.83$	$\chi_2^{2(0.05)} = 5.99$
White	$W = nR^2_{ra} = 12.24$	$\chi_2^{2(0.01)} = 9.21$

6.5 Heteroskedasticity

EXAMPLE 6.9 Heteroskedasticity consistent standard errors in the models explaining the market value of Spanish banks (Continuation of example 6.7) (file bolmad95)

Non consistent

$$\widehat{marktval} = \underset{(30.85)}{29.42} + \underset{(0.127)}{1.219}\,bookval \qquad \widehat{\ln(marktval)} = \underset{(0.265)}{0.676} + \underset{(0.062)}{0.9384}\ln(bookval)$$

White procedure

$$\widehat{marktval} = \underset{(18.67)}{29.42} + \underset{(0.249)}{1.219}\,bookval \qquad \widehat{\ln(marktval)} = \underset{(0.3218)}{0.676} + \underset{(0.0698)}{0.9384}\ln(bookval)$$

EXAMPLE 6.10 Application of weighted least squares in the demand of hotel services (Continuation of example 6.8) (file hostel)

$$\widehat{|\hat{u}_i|} = \underset{(0.143)}{0.0239} + \underset{(2.73)}{0.0003}\, inc \qquad R^2 = 0.1638$$

$$\widehat{|\hat{u}_i|} = \underset{(-1.34)}{-0.4198} + \underset{(2.82)}{0.0235}\sqrt{inc} \qquad R^2 = 0.1733$$

$$\widehat{|\hat{u}_i|} = \underset{(5.39)}{0.8857} - \underset{(-2.87)}{532.1}\frac{1}{inc} \qquad R^2 = 0.1780$$

$$\widehat{|\hat{u}_i|} = \underset{(-2.46)}{-2.7033} + \underset{(2.88)}{0.4389}\ln(inc) \qquad R^2 = 0.1788$$

WLS estimation

$$\widehat{\ln(hostel)}_i = \underset{(2.15)}{-16.21} + \underset{(0.309)}{2.709}\ln(inc)_i + \underset{(0.247)}{1.401}\, secstud_i + \underset{(0.326)}{2.982}\, terstud_i - \underset{(0.085)}{0.445}\, hhsize_i$$

$$R^2 = 0.914 \qquad n = 40$$

6.6 Autocorrelation

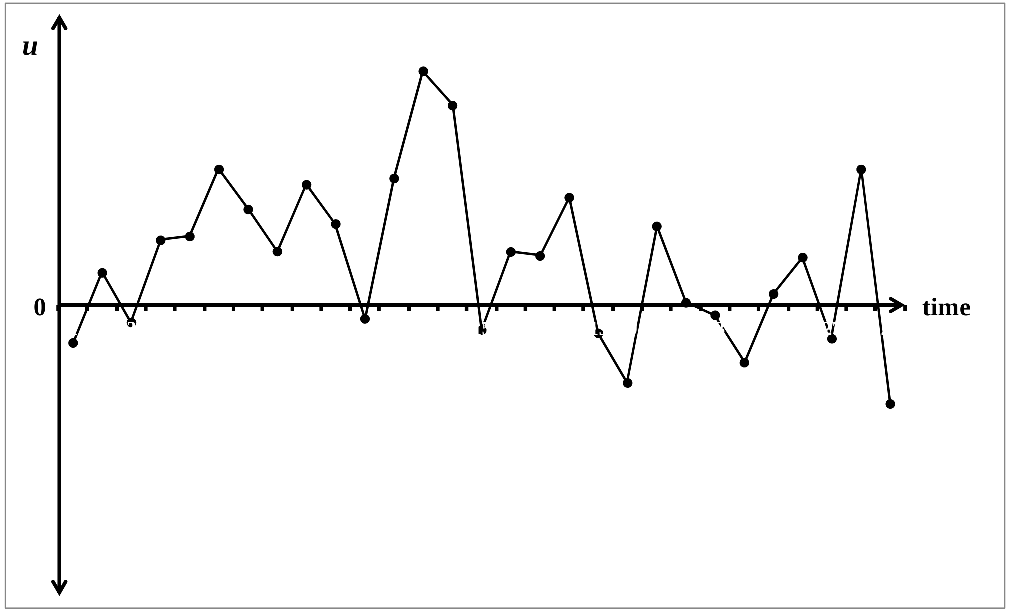


FIGURE 6.3. Plot of non-autocorrelated disturbances.

6.6 Autocorrelation

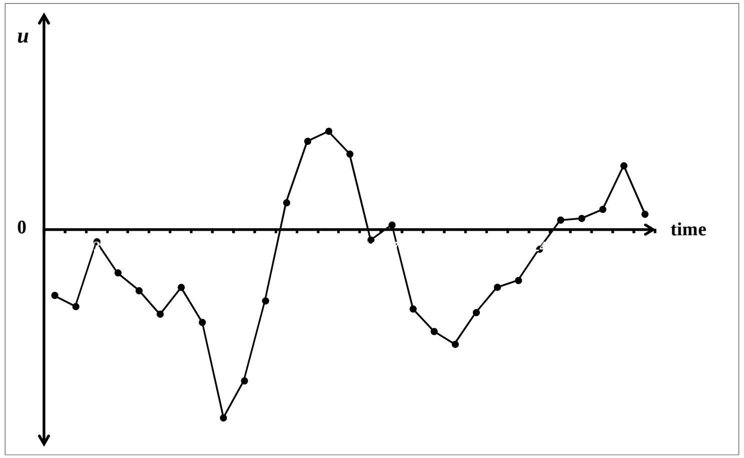


FIGURE 6.4. Plot of positive autocorrelated disturbances.

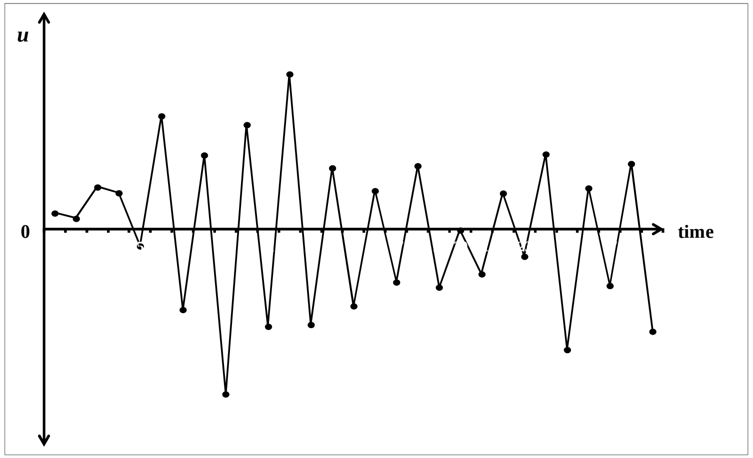


FIGURE 6.5. Plot of negative autocorrelated disturbances.

6.6 Autocorrelation

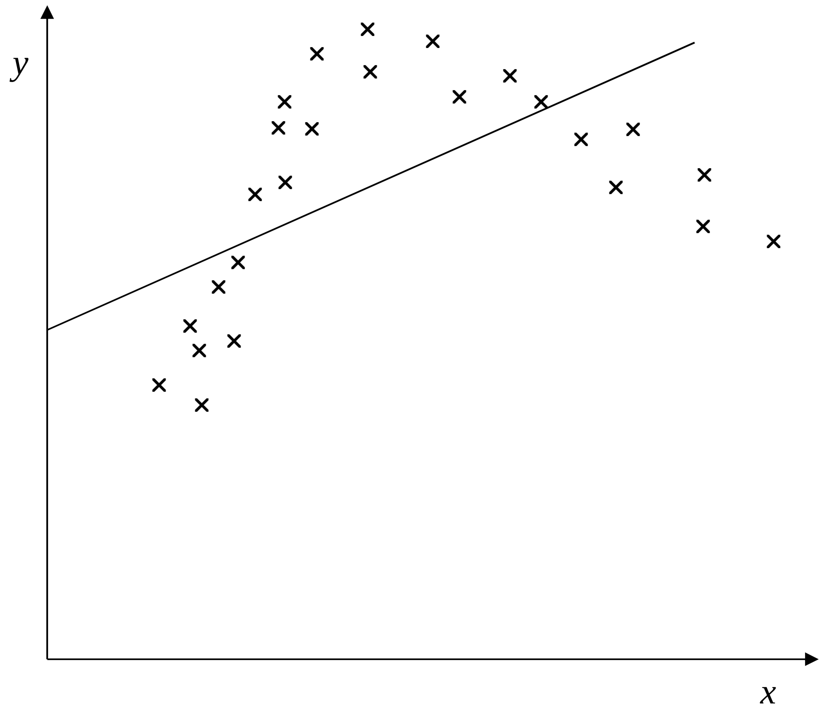
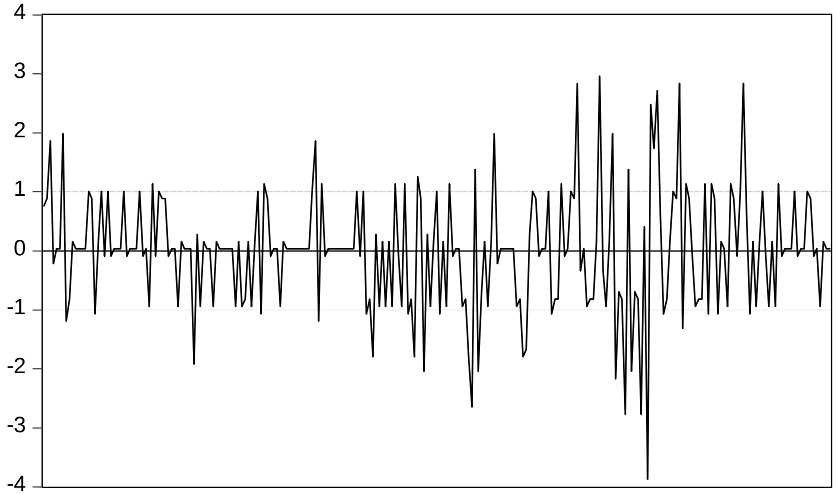


FIGURE 6.6. Autocorrelated disturbances due to a specification bias.

EXAMPLE 6.11 Autocorrelation in the model to determine the efficiency of the Madrid Stock Exchange (file bolmadef)

d_L=1.664; d_U=1.684

Since DW=2.04>d_U, we do not reject the null hypothesis that the disturbances are not autocorrelated for a significance level of α =0.01, i.e. of 1%.



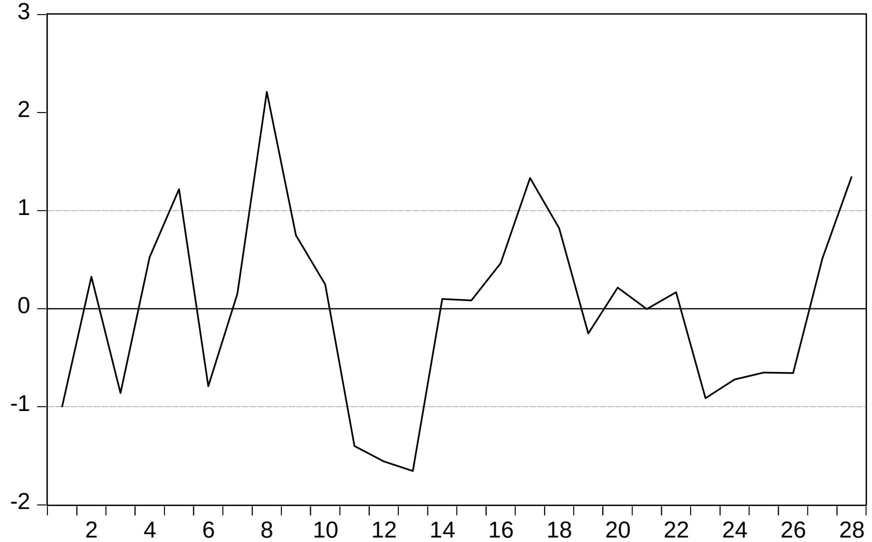
GRAPHIC 6.4. Standardized residuals in the estimation of the model to determine the efficiency of the Madrid Stock Exchange.

EXAMPLE 6.12 Autocorrelation in the model for the demand for fish (file fishdem)

For n=28 and k'=3, and for a significance level of 1%:

$$d_L=0.969; \qquad d_U=1.415$$

Since $d_L<1.202<d_U$, there is not enough evidence to accept the null hypothesis, or to reject it.



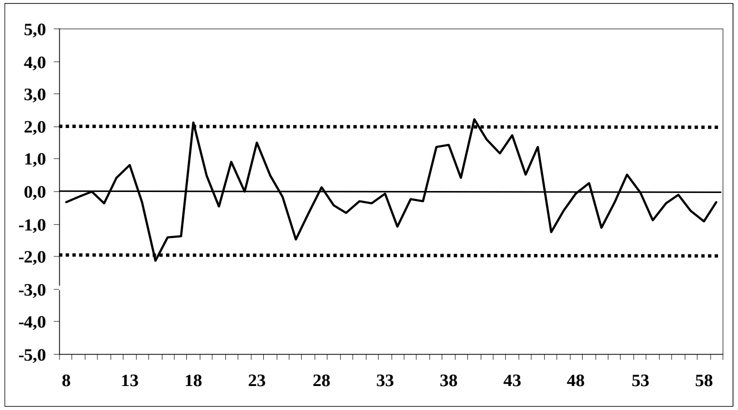
GRAPHIC 6.5. Standardized residuals in the model on the demand for fish.

6.6 Autocorrelation

EXAMPLE 6.13 Autocorrelation in the case of Lydia E. Pinkham (file pinkham)

$$h=\hat{\rho}\sqrt{\frac{n}{1-n\widehat{\text{var}}\left(\hat{\beta}_j\right)}}=\left[1-\frac{d}{2}\right]\sqrt{\frac{n}{1-n\widehat{\text{var}}\left(\hat{\beta}_j\right)}}=\left[1-\frac{1.2012}{2}\right]\sqrt{\frac{53}{1-53\times 0.0814^2}}=3.61$$

Given this value of h, the null hypothesis of no autocorrelation is rejected for α=0.01 or, even, for α=0.001, according to the table of the normal distribution.



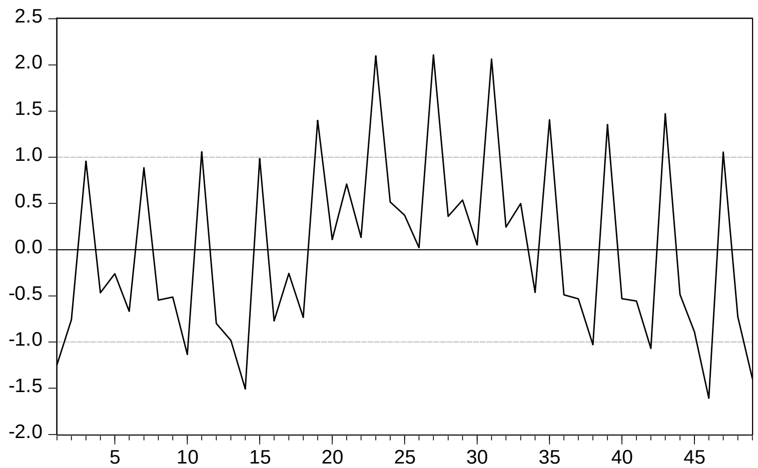
GRAPHIC 6.6. Standardized residuals in the estimation of the model of the Lydia E. Pinkham case.

6.6 Autocorrelation

EXAMPLE 6.14 Autocorrelation in a model to explain the expenditures of residents abroad (file qnatacsp)

$$\widehat{\ln(turimp_t)} = \underset{(3.43)}{-17.31} + \underset{(0.276)}{2.0155}\ln(gdp_t)$$

$$R^2 = 0.531 \qquad DW = 2.055 \qquad n = 49$$



GRAPHIC 6.7. Standardized residuals in the estimation of the model explaining the expenditures of residents abroad.

For a $AR(4)$ scheme, is equal to $BG = nR_{ar}^2 = 36.35$. Given this value of BG, the null hypothesis of no autocorrelation is rejected for α=0.01, since $\chi_5^{2(\alpha)} = 15.09$.

6.6.4 *HAC* standard errors

EXAMPLE 6.15 *HAC* standard errors in the case of Lydia E. Pinkham (Continuation of example 6.13) (file pinkham)

TABLE 6.9.The *t* statistics, conventional and *HAC*, in the case of Lydia E. Pinkham.

regressor	t conventional	t HAC	ratio
intercept	2.644007	1.779151	1.49
advexp	3.928965	5.723763	0.69
sales(-1)	7.45915	6.9457	1.07
$d1$	-1.499025	-1.502571	1
$d2$	3.225871	2.274312	1.42
$d3$	-3.019932	-2.658912	1.14