5. Heteroscedasticity

5.1 What is meant by heteroskedasticity and what can cause this model defect?

Solution:

The assumption A.2 was that the disturbance variables in each period have the same variance, i.e. are homoscedastic: \( \text{Var}(u_t) = \sigma^2 \) for \( t = 1, 2, \ldots, n \). One speaks of heteroscedasticity if the variance of the disturbance variable is not constant but changes in the individual periods or as a function of the cross-sectional units (e.g. companies, regions). Heteroscedasticity is also present if the variance of the disturbance variable varies in groups but is constant within the groups (e.g. dispersion differences between West and East, but not within West and East German regions).

Potential reasons for heteroscedasticity:

- for time series data:
  - If the exogenous variables are subject to a trend, this trend can also be reflected in the disturbance variable variance.

- for cross-section data:
  - In the case of groups of statistical units (e.g. households) which are in the upper range of one or more exogenous variables such as income, the dependent variable such as consumption often also shows a greater dispersion. In our example, it reflects the greater consumption possibilities of households with higher incomes. Greater consumption opportunities are shown by the fact that an above-average amount can be consumed or saved for a given income. This means, however, that the deviations (residuals) from the regression line tend to be greater than for low incomes, where the scope for consumption is limited. This expresses the fact that the disturbance variance is greater for high incomes than for low incomes.

5.2 Which inferential statistical consequences result for the OLS parameter estimators of a multiple regression model in the case of heteroscedasticity?

Solution:

The OLS estimators of the regression coefficients remain unbiased, but are no longer efficient. This means that the standard errors (= square root of variances) of the estimated regression coefficients estimators are distorted. As a result, the significance tests on the regression coefficients become invalid. In addition, the confidence intervals for the regression coefficients can no longer be calculated validly.
5.3 Explain the basic idea of the generalized least squares method (GLS method) to eliminate heteroscedasticity!

Solution:
If heteroscedasticity is present, then the main diagonal elements of the covariance matrix of the disturbance variables differ:

\[
\text{Cov}(u) = \text{E}(uu') = \begin{pmatrix}
\sigma_1^2 & \ldots & 0 & 0 \\
0 & \sigma_2^2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sigma_n^2
\end{pmatrix}
\]
or rather \( \sigma_i^2 = \sigma^2 \cdot \omega_i \cdot = \sigma^2 \cdot z_i^2 \) (\( \omega_i \): scaling variable for which a squared exogenous variable \( z_i^2 \) is usually selected).

\[
\text{Cov}(u) = \text{E}(uu') = \begin{pmatrix}
\omega_1 \cdot \sigma^2 & \ldots & 0 & 0 \\
0 & \omega_2 \cdot \sigma^2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \omega_n \cdot \sigma^2
\end{pmatrix} = \sigma^2 \cdot \begin{pmatrix}
\omega_1 & \ldots & 0 & 0 \\
0 & \omega_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \omega_n
\end{pmatrix}.
\]

A transformation matrix \( T \):

\[
T = \begin{pmatrix}
\frac{1}{\sqrt{\omega_1}} & \ldots & 0 & 0 \\
0 & \frac{1}{\sqrt{\omega_2}} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \frac{1}{\sqrt{\omega_n}}
\end{pmatrix}
\]
is determined so that applies:

\[
T' \cdot T = \Omega^{-1}.
\]

Multiply the regression equation from the left by \( T \) to obtain the transformed regression equation:

\[
\begin{pmatrix}
\hat{y}^* \\
\hat{X}^* \\
\hat{u}^*
\end{pmatrix} = \begin{pmatrix}
T \hat{X} \cdot \beta + T \hat{u}
\end{pmatrix}.
\]

The observed values are thus multiplied or weighted by the values of the transformation matrix (this is why the method is also called WLS estimation). As can be shown, the transformed disturbance satisfies the standard assumptions of the multiple regression model. How can \( T \) be determined?

Approach 1: Group-specific heteroskedasticity

OLS estimators are calculated separately for all m groups:
With the OLS estimators for all $\ell$ groups:

$$\hat{\beta}_\ell = (X'_\ell X_\ell)^{-1} X'_\ell y_\ell$$

the OLS residuals:

$$\hat{u}_\ell = y_\ell - X_\ell \hat{\beta}_\ell .$$

and the estimated disturbance variance (= residual variance):

$$\hat{\sigma}_\ell^2 = \frac{\hat{u}'_\ell \hat{u}_\ell}{n - k}$$

are determined for all $\ell$ groups. The estimated group-specific disturbance variations are used for the definition of the transformation matrix $T$. The transformation matrix $T$ is then determined for the group:

$$T = \begin{pmatrix}
1 & 0 & \ldots & 0 & \ldots & 0
\\
\frac{1}{\hat{\sigma}_1} & 0 & \ldots & 0 & \ldots & 0
\\
0 & \frac{1}{\hat{\sigma}_1} & \ldots & 0 & \ldots & 0
\\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots
\\
0 & 0 & \ldots & \frac{1}{\hat{\sigma}_1} & \ldots & 0
\\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots
\\
0 & 0 & \ldots & 0 & \ldots & \frac{1}{\hat{\sigma}_m}
\\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots
\\
0 & 0 & \ldots & 0 & \ldots & \frac{1}{\hat{\sigma}_m}
\end{pmatrix}$$

However, the approach is only applicable to a limited extent,

- since the $m$ groups must be known (e.g. west/east regime).
- The estimation will usually only be feasible with a small number of groups, because the estimation of the disturbance variance is not very precise if it is based on a too small number of observations.

**Approach 2: Heteroscedasticity as a function of the scaling variables $\omega$**

The scaling variable is defined by the power of an exogenous variable $z$, which is generally an explanatory variable $x_j$ of the regression model ($z$ could also be an exogenous variable that is not used in the model to explain the $y$-variables, such as a trend variable). In the simplest case, $z$ is equated with the $j$-th $x$ variable. Negative values are switched off from the outset when the $x$-values are squared:

$$\omega_\ell = z^2_\ell = x^2_{jt} \quad \text{bzw.} \quad \sigma^2_\ell = z^2_\ell \cdot \sigma^2 = x^2_{jt} \cdot \sigma^2 .$$

The elements are entered in the transformation matrix:
The transformed regression equation can then also be represented as follows:

\[
\frac{y_t}{x_{jt}} = \hat{\beta}_1 \frac{1}{x_{jt}} + \hat{\beta}_2 \frac{x_{jt}}{x_{jt}} + \hat{\beta}_j \frac{x_{jt}}{x_{jt}} + \ldots + \hat{\beta}_k \frac{x_{kt}}{x_{jt}} + \hat{\alpha}_t + \hat{\tilde{u}}_t
\]

\[
y^*_t = x^*_1 \quad x^*_2 \quad x^*_3 \quad x^*_4 = 1 \quad x^*_5 \quad \hat{\tilde{u}}_t
\]

5.4 Test the disturbance term of the natural gas demand model as a function of the gas price (= energy model Ia, data see Exercise 2.3) with the Goldfeld-Quandt test for heteroscedasticity (\(\alpha = 0.05\))!

**Solution:**

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Ordering the observation data according to the ascending values of the j-th regressor to which heteroscedasticity is attributed. The observation values are ordered according to the gas price</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>x_t</td>
</tr>
<tr>
<td>5</td>
<td>12,2</td>
</tr>
<tr>
<td>7</td>
<td>12,5</td>
</tr>
<tr>
<td>1</td>
<td>12,9</td>
</tr>
<tr>
<td>6</td>
<td>12,9</td>
</tr>
<tr>
<td>4</td>
<td>13,2</td>
</tr>
<tr>
<td>3</td>
<td>13,7</td>
</tr>
<tr>
<td>2</td>
<td>13,8</td>
</tr>
</tbody>
</table>

xt:
Step 2  **Splitting of the observation data into 2 equal groups**

Due to the small and odd number of observations, we set \( c = 1 \) so that the mean value \( (t = 6) \) of the ordered data set is not used. It is shown in gray in the following table. Above are the periods of the first sample and below the second sample.

<table>
<thead>
<tr>
<th></th>
<th>( t )</th>
<th>( x_t )</th>
<th>( y_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>5</td>
<td>12,2</td>
<td>1,2</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>12,5</td>
<td>1,4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>12,9</td>
<td>1,0</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>12,9</td>
<td>1,1</td>
</tr>
<tr>
<td>Sample 2</td>
<td>4</td>
<td>13,2</td>
<td>0,9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>13,7</td>
<td>0,7</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>13,8</td>
<td>0,8</td>
</tr>
</tbody>
</table>

Step 3  **Formulation of the hypothesis**:

\( H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2 \) und \( H_1 : \sigma_1^2 \neq \sigma_2^2 \)

Step 4  **Determination of the significance level \( \alpha \): \( \alpha = 0,05 \)**

Step 5  **Regression to be performed**

1. **sample**:

\[ y_{1t} = \beta_{11} + \beta_{12} \cdot x_{1t} + u_{1t} \]

a) **OLS estimator** \( \hat{\beta}_{11} \) and \( \hat{\beta}_{12} \)

<table>
<thead>
<tr>
<th>( t )</th>
<th>( x_t )</th>
<th>( y_t )</th>
<th>( x_t^2 )</th>
<th>( x_t \cdot y_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>12,2</td>
<td>1,2</td>
<td>148,84</td>
<td>14,64</td>
</tr>
<tr>
<td>7</td>
<td>12,5</td>
<td>1,4</td>
<td>156,25</td>
<td>17,50</td>
</tr>
<tr>
<td>1</td>
<td>12,9</td>
<td>1,0</td>
<td>166,41</td>
<td>12,90</td>
</tr>
<tr>
<td>( \Sigma )</td>
<td>37,6</td>
<td>3,6</td>
<td>471,50</td>
<td>45,04</td>
</tr>
</tbody>
</table>

**OLS estimator of regression coefficients** \( \beta_1 \) and \( \beta_2 \):

\[
\hat{\beta}_{12} = \frac{n_1 \cdot \sum x_{1t} y_{1t} - \sum x_{1t} \sum y_{1t}}{n_1 \cdot \sum x_{1t}^2 - (\sum x_{1t})^2} = \frac{3 \cdot 45,04 - 37,6 \cdot 3,6}{3 \cdot 471,5 - 37,6^2} = \frac{-0,24}{0,74} = -0,324
\]

\[
\hat{\beta}_{11} = \frac{\sum y_{1t}}{n_1} - \hat{\beta}_{12} \cdot \frac{\sum x_{1t}}{n_1} = \frac{3,6}{3} - (-0,324) \cdot \frac{37,6}{3} = 1,2 + 0,324 \cdot 12,533 = 5,261
\]

\[
= 4,061
\]

**OLS-estimated demand function for natural gas**:

\( \hat{y}_{1t} = 5,261 - 0,324 \cdot x_{1t}, \ t=5,7,1 \)

b) **regression values**
\[ t=5: \hat{y}_{15} = 5,261 - 0,324 \cdot (x_5=12,2) = 1,308 \]
\[ t=7: \hat{y}_{17} = 5,261 - 0,324 \cdot (x_7=12,5) = 1,211 \]
\[ t=1: \hat{y}_{11} = 5,261 - 0,324 \cdot (x_1=12,9) = 1,801 \]

c) residuals
\[ t=5: \hat{u}_{15} = y_{15} - \hat{y}_{15} = 1,2 - 1,308 = -0,108 \]
\[ t=7: \hat{u}_{17} = y_{17} - \hat{y}_{17} = 1,4 - 1,211 = 0,189 \]
\[ t=1: \hat{u}_{11} = y_{11} - \hat{y}_{11} = 1,0 - 1,081 = -0,081 \]
d) sum of squares of the residuals
\[
\hat{u}_1\hat{u}_1 = \begin{bmatrix} -0,108 \\ 0,189 \\ -0,081 \end{bmatrix} \begin{bmatrix} -0,108 \\ 0,189 \\ -0,081 \end{bmatrix} = 0,0539
\]

2. sample:
\[ y_{2t} = \beta_{21} + \beta_{22} \cdot x_{2t} + u_{2t} \]

a) OLS estimator \( \hat{\beta}_{11} \) and \( \hat{\beta}_{12} \)

<table>
<thead>
<tr>
<th>t</th>
<th>( x_t )</th>
<th>( y_t )</th>
<th>( x^2_t )</th>
<th>( x_t \cdot y_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>13,2</td>
<td>0,9</td>
<td>174,24</td>
<td>11,88</td>
</tr>
<tr>
<td>3</td>
<td>13,7</td>
<td>0,7</td>
<td>187,69</td>
<td>9,59</td>
</tr>
<tr>
<td>2</td>
<td>13,8</td>
<td>0,8</td>
<td>190,44</td>
<td>11,04</td>
</tr>
<tr>
<td>( \Sigma )</td>
<td>40,7</td>
<td>2,4</td>
<td>552,37</td>
<td>32,51</td>
</tr>
</tbody>
</table>

OLS estimator of regression coefficients \( \beta_1 \) und \( \beta_2 \):
\[
\hat{\beta}_{22} = \frac{n_2 \cdot \Sigma x_{2t} y_{2t} - \Sigma x_{2t} \Sigma y_{2t}}{n_2 \cdot \Sigma x^2_{2t} - (\Sigma x_{2t})^2} = \frac{3 \cdot 32,51 - 2,4 \cdot 40,7}{3 \cdot 552,37 - 40,7^2} = -0,15 \]
\[
\hat{\beta}_{21} = \frac{\Sigma y_{2t} \cdot \hat{\beta}_{22} - \Sigma x_{2t} \cdot \hat{\beta}_{22}}{n_2} = \frac{2,4 - (-0,242) \cdot 40,7}{3} = 0,8 + 0,242 \cdot 13,567 = 4,083
\]

OLS estimated demand function for natural gas:
\[ \hat{y}_{2t} = 4,083 - 0,242 \cdot x_{2t}, \ t=6,5,7 \]

b) regression values
\[ t=4: \hat{y}_{14} = 4,083 - 0,242 \cdot (x_4=13,2) = 0,889 \]
\[ t=3: \hat{y}_{13} = 4,083 - 0,242 \cdot (x_3=13,7) = 0,768 \]
\[ t=2: \hat{y}_{12} = 4,083 - 0,242 \cdot (x_2=13,8) = 0,743 \]
c) residuals
t=4: \[ \hat{u}_{24} = y_{24} - \hat{y}_{24} = 0.9 - 0.889 = 0.011 \]

\[ \hat{u}_{23} = y_{23} - \hat{y}_{23} = 0.7 - 0.768 = -0.068 \]

t=2: \[ \hat{u}_{22} = y_{22} - \hat{y}_{22} = 0.8 - 0.743 = 0.057 \]
d) sum of squares of the residuals

\[
\begin{pmatrix}
0.011 & -0.068 & 0.057 \\
-0.068 & 0.057 & 0.011
\end{pmatrix} = 0.0080
\]

**Step 6**

**Calculation of the test statistic:**

\[
1/GQ = \frac{\hat{u}_1^2 \hat{u}_1}{\hat{u}_2^2 \hat{u}_2} = \frac{0.0539}{0.0080} = 6.738
\]

**Step 7**

**Critical value (\(\alpha=0.05\)):**

\[
F_{(n-c)/2-k; (n-c)/2-k;1-\alpha} = F_{(7-1)/2-2; (7-1)/2-2;0.95} = F_{1;1;0.95} = 161
\]

**Step 8**

**Test decision:**

\[
1/GQ = 6.738 < F_{1;1;0.95} = 161 \rightarrow H_0 \text{ can not be rejected}
\]

The Goldfeld-Quandt test does not indicate heteroskedasticity of the disturbance term.

---

**5.5**

Estimate the demand function for natural gas depending on the gas price (= energy model Ia) using the specification \( \sigma_t^2 = \sigma^2 \cdot \text{GASP}_t^2 \) of the error term variance with the weighted least squares method (WLS method)!

**Note:** Calculate the WLS estimators for \( \beta_1 \) and \( \beta_2 \) using the formulas for the OLS estimators with \( y_t \rightarrow y_t^* \), \( x_t \rightarrow x_t^* \). Use the transformed variables \( (x_{1t}^*)^2 \) and \( x_{1t}^* \cdot y_t^* \) with three decimal places!

**Solution:**

The specification of the disturbance variance in the form of

\[
\sigma_t^2 = \sigma^2 \cdot \text{GASP}_t^2
\]

leads to the transformed econometric demand model for natural gas (\( \text{GASV}_t = y_t \), \( \text{GASP}_t = x_t \)):

\[
\left( \begin{array}{c}
\text{GASV}_t^* \\
\text{GASP}_t^*
\end{array} \right) = \beta_1 \cdot \left( \begin{array}{c}
\text{GASV}_t \\
\text{GASP}_t
\end{array} \right) + \beta_2 \cdot \text{GASP}_t + \hat{u}_t^* = \beta_1 \cdot \left( \begin{array}{c}
\text{GASV}_t \\
\text{GASP}_t
\end{array} \right) + \beta_2 + u_t^*
\]

When calculating the WLS estimators for \( \beta_1 \) and \( \beta_2 \), the formulas of the OLS estimators can be used in the following form using the transformed variables:

\[
y_t \rightarrow y_t^*, \quad x_t \rightarrow x_{1t}^*
\]
WLS estimator of the regression coefficients $\beta_1$ and $\beta_2$:

$$\hat{\beta}_1,\text{WLS} = \frac{n \cdot \sum x_{1t}^* y_{1t}^* - \sum x_{1t}^* \sum y_{1t}^*}{n \cdot (\sum x_{1t}^*)^2 - (\sum x_{1t}^*)^2} = \frac{7 \cdot 0.0428 - 0.539 \cdot 0.550}{7 \cdot 0.0416 - 0.539^2} = 4.632$$

$$\hat{\beta}_2,\text{WLS} = \frac{\sum y_{1t}^*}{n} - \hat{\beta}_1,\text{WLS} \cdot \frac{\sum x_{1t}^*}{n} = \frac{0.550}{7} - 4.632 \cdot \frac{0.539}{7} = 0.079 - 4.632 \cdot 0.077 = -0.278$$

WLS-estimated demand function for natural gas:

$$GASV_t = 4.632 - 0.278 \cdot GASP_t$$

The White test has produced the following result for the Energy Model III:

White Heteroskedasticity Test:

<table>
<thead>
<tr>
<th>Obs*R-squared</th>
<th>5.696925</th>
<th>Probability</th>
<th>0.457982</th>
</tr>
</thead>
</table>

Test Equation: Dependent Variable: RESID^2
Method: Least Squares
Sample: 1980 1995, Included observations: 16

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-29.74777</td>
<td>58.70411</td>
<td>-0.506741</td>
<td>0.6245</td>
</tr>
<tr>
<td>GASPR</td>
<td>-25.04399</td>
<td>28.17182</td>
<td>-0.888973</td>
<td>0.3972</td>
</tr>
<tr>
<td>GASPR^2</td>
<td>17.07002</td>
<td>15.58422</td>
<td>1.095340</td>
<td>0.3018</td>
</tr>
<tr>
<td>FERNWPR</td>
<td>-21.22605</td>
<td>89.98156</td>
<td>-0.235893</td>
<td>0.8188</td>
</tr>
<tr>
<td>FERNWPR^2</td>
<td>8.174083</td>
<td>44.31804</td>
<td>0.184441</td>
<td>0.8578</td>
</tr>
<tr>
<td>VEINKR</td>
<td>0.068311</td>
<td>0.062335</td>
<td>1.095866</td>
<td>0.3016</td>
</tr>
<tr>
<td>VEINKR^2</td>
<td>-2.23E-05</td>
<td>2.03E-05</td>
<td>-1.097094</td>
<td>0.3011</td>
</tr>
</tbody>
</table>

R-squared    | 0.356058    | Mean dependent var | 0.376351 |
Adjusted R-squared | -0.073237 | S.D. dependent var | 0.682307 |
S.E. of regression | 0.706851 | Akaike info criterion | 2.443642 |
Sum squared resid | 4.496743 | Schwarz criterion | 2.781649 |
Log likelihood  | -12.54913 | F-statistic | 0.829402 |
Durbin-Watson stat | 2.894434 | Prob(F-statistic) | 0.575525 |
a) Explain the basic idea of the White test with reference to the auxiliary regression carried out!

**Solution:**

Basic idea of the White test:

Heteroskedasticity, which is based on the explanatory variables, the squared values of the explanatory variables (variance concept) is modelled and tested.

b) Enter the test decision of the White test shown here using the critical value and the p-value (α=0,10)! Interpret the result!

**Solution:**

Null hypothesis of the White test:

\[ H_0: \sigma_t^2 = \sigma^2 \text{ for all } t \] (homoskedasticity)

Test statistic:

\[ W = n \cdot R^2 = 16 \cdot 0.356058 = 5.697 \] (R² from auxiliary regression)

Critical value (α=0,10): \( \chi^2_{r;0.90} = \chi^2_{6;0.90} = 10.6 \)

Test decision:

\[ W = 5.697 < \chi^2_{6;0.90} = 10.6 \Rightarrow H_0 \text{ maintain} \]

or

\[ p = 0.458 > \alpha=0.10 \Rightarrow H_0 \text{ maintain} \]

Interpretation:

The White test did not reveal any heteroskedasticity.

c) Enter the auxiliary regression of the full White test for the energy model III and interpret the corresponding abridged output:

**Solution:**

Full auxiliary regression of the White test:

\[
\hat{u}_t^2 = \alpha_0 + \alpha_1 \cdot \text{GASPR}_t + \alpha_2 \cdot \text{FERNWPR}_t + \alpha_3 \cdot \text{VEINKR}_t + \alpha_11 \cdot \text{GASPR}_t^2 \\
+ \alpha_{22} \cdot \text{FERNWPR}_t^2 + \alpha_{33} \cdot \text{VEINKR}_t^2 + \alpha_{12} \cdot \text{GASPR}_t \cdot \text{FERNWPR}_t \\
+ \alpha_{13} \cdot \text{GASPR}_t \cdot \text{VEINKR}_t + \alpha_{23} \cdot \text{FERNWPR}_t \cdot \text{VEINKR}_t + v_t
\]

with \( \hat{u}_t \) as the residual of the demand function for natural gas

Test statistic: \( W = n \cdot R^2 = 8.074 \)
Test decision:
Since \( p = 0.527 \) is greater than all common significance levels \( (\alpha = 0.01, \alpha = 0.05, \alpha = 0.10) \), the null hypothesis of a homoscedastic disturbance variance, \( H_0: \sigma_t^2 = \sigma^2 \) for all \( t \), cannot be rejected.

5.7 From the regression of natural gas demand on the gas price (= energy model Ia, data see Assignm. 2.3), the following OLS residuals result:

<table>
<thead>
<tr>
<th>( t )</th>
<th>( \hat{u}_t )</th>
<th>( x_t - \bar{x} )</th>
<th>( (x_t - \bar{x})^2 )</th>
<th>( \hat{u}_t^2 )</th>
<th>( (x_t - \bar{x})^2 \cdot \hat{u}_t^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.9</td>
<td>-0.062</td>
<td>-0.129</td>
<td>0.016641</td>
<td>0.003844</td>
</tr>
<tr>
<td>2</td>
<td>13.8</td>
<td>0.069</td>
<td>0.771</td>
<td>0.594441</td>
<td>0.004761</td>
</tr>
<tr>
<td>3</td>
<td>13.7</td>
<td>-0.067</td>
<td>0.671</td>
<td>0.450241</td>
<td>0.004489</td>
</tr>
<tr>
<td>4</td>
<td>13.2</td>
<td>-0.051</td>
<td>0.171</td>
<td>0.029241</td>
<td>0.002601</td>
</tr>
<tr>
<td>5</td>
<td>12.2</td>
<td>-0.119</td>
<td>-0.829</td>
<td>0.687241</td>
<td>0.014161</td>
</tr>
<tr>
<td>6</td>
<td>12.9</td>
<td>0.038</td>
<td>-0.129</td>
<td>0.016641</td>
<td>0.001444</td>
</tr>
<tr>
<td>7</td>
<td>12.5</td>
<td>0.191</td>
<td>-0.529</td>
<td>0.279841</td>
<td>0.036481</td>
</tr>
<tr>
<td>Nder</td>
<td>91.2</td>
<td>-0.001 ( \approx ) 0</td>
<td>-0.003 ( \approx ) 0</td>
<td>2.074287</td>
<td>0.067781</td>
</tr>
</tbody>
</table>

\[ \bar{x} = \frac{1}{7} \sum_{t=1}^{7} x_t = \frac{1}{7} \cdot 91.2 = 13.029 \]

- Standard error \( \hat{\beta}_2 \) of homoscedastic interference terms

\[ \hat{\sigma}^2 = \frac{1}{n-k} \sum \hat{u}_t^2 = \frac{1}{7-2} \cdot 0.067781 = 0.013556 \]

\[ \text{Var}(\hat{\beta}_2) = \frac{1}{\sum (x_t - \bar{x})^2} \cdot \hat{\sigma}^2 = \frac{1}{2.074287} \cdot 0.013556 = 0.006535 \]

\[ \hat{\sigma}_{\hat{\beta}_2} = \sqrt{0.006535} = 0.080839 \]

- White's heteroskedasticity-consistent standard error of \( \hat{\beta}_2 \):
\[
\text{Var}(\hat{\beta}_2) = \frac{\sum (x_t - \bar{x})^2 \hat{u}_t^2}{\sum (x_t - \bar{x})^2} = 0.024956, \quad 2.074287^2 = 0.005800
\]

\[\hat{\sigma}_{\hat{\beta}_2} = \sqrt{0.005800} = 0.076158\]

b) Specify White's heteroskedasticity-consistent estimators of the standard errors of the two regression coefficients numerically using the variance-covariance matrix \(^\wedge\text{Cov}(\hat{\beta})\)!

\textbf{Solution:}

White's heteroskedasticity-consistent estimators of the variance-covariance matrix \(^\wedge\text{Cov}(\hat{\beta})\):

\[\text{Cov}(\hat{\beta}) = (X'X)^{-1}X^2 \text{diag}(\hat{u}_1^2, \hat{u}_2^2, ..., \hat{u}_n^2)X(X'X)^{-1}\]

observation matrix \(X\):

\[
X = \begin{bmatrix}
1 & 12.9 \\
1 & 13.8 \\
1 & 13.7 \\
1 & 13.2 \\
1 & 12.2 \\
1 & 12.9 \\
1 & 12.5 \\
\end{bmatrix}
\]

product matrix \(X'X\):

\[
X'X = \begin{bmatrix}
1 & 12.9 \\
1 & 13.8 \\
1 & 13.7 \\
1 & 13.2 \\
1 & 12.2 \\
1 & 12.9 \\
1 & 12.5 \\
\end{bmatrix} = \begin{bmatrix}
1 & 12.9 \\
1 & 13.8 \\
1 & 13.7 \\
1 & 13.2 \\
1 & 12.2 \\
1 & 12.9 \\
1 & 12.5 \\
\end{bmatrix} = \begin{bmatrix}
7 & 119.2 \\
91.2 & 1190.28 \\
\end{bmatrix}
\]

Inverse of the product matrix \(X'X\):

\[
(X'X)^{-1} = \begin{bmatrix}
81.975207 & -6.280992 \\
-6.280992 & 0.482094 \\
\end{bmatrix}
\]
Estimated variance-covariance matrix of the error terms:

\[
\widehat{\text{Cov}}(u) = \text{diag}(\hat{u}_1^2, \hat{u}_2^2, \ldots, \hat{u}_n^2)
\]

\[
= \begin{bmatrix}
0.003844 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.004761 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.004689 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.002601 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.014161 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.001444 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.036481 \\
\end{bmatrix}
\]

White's heteroskedasticity-consistent estimators variance-covariance matrix \(\widehat{\text{Cov}}(\hat{\beta})\):

\[
\widehat{\text{Cov}}(\hat{\beta}) = (X'X)^{-1}X\text{diag}(\hat{u}_1^2, \hat{u}_2^2, \ldots, \hat{u}_n^2)X(X'X)^{-1}
\]

\[
= \begin{bmatrix}
81.975207 & -6.280992 \\
-6.280992 & 0.482094 \\
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
12.9 & 13.8 & 13.7 & 13.2 & 12.2 & 12.9 & 12.5 \\
\end{bmatrix} \cdot \begin{bmatrix}
0.003844 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.004761 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.004689 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.002601 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.014161 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.001444 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.036481 \\
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 & 12.9 \\
1 & 13.8 \\
1 & 13.7 \\
1 & 13.2 \\
1 & 12.2 \\
1 & 12.9 \\
1 & 12.5 \\
\end{bmatrix} \cdot \begin{bmatrix}
81.975207 & 6.280992 \\
6.280992 & 0.482094 \\
\end{bmatrix} = \begin{bmatrix}
1.032797 & 0.077475 \\
0.077475 & 0.005817 \\
\end{bmatrix}
\]

Heteroscedasticity-consistent standard errors of regression coefficients:

\[
\hat{\sigma}_{\hat{\beta}_1} = \sqrt{1.032797} = 1.016266
\]

\[
\hat{\sigma}_{\hat{\beta}_2} = \sqrt{0.005817} = 0.076269
\]
5.8 Discuss the advantages and disadvantages of using heteroskedasticity consistent (HC) standard errors instead of GLS estimation (generalized least squares estimation) of the regression model!

Solution:
While GLS estimation requires a hypothesis on the occurrence of heteroskedasticity, the use of heteroskedasticity-consistent (HC) standard errors does not require a specific form of heteroskedasticity. Rather, the squares of the OLS residuals, as estimators of the heteroscedastic variances of the disturbance, are prominently used. Since often only clues about the cause of heteroskeasticity are available instead of precise knowledge, econometric practice prefers to use heteroskedasticity-consistent (HC) standard errors. Regardless of the choice of either method, a GLS estimate or the use of heteroskedasticity-consistent (HC) standard error, the estimated regression coefficients are not unexpected. For both methods, it can be shown that the bias of estimated standard errors of regression coefficients is close to zero as the sample size increases. By using adjustment factors for finite samples, the validity of the significance test and confidence intervals is to be established even for small sample sizes. In contrast, the best precision (efficiency) of the estimated regression coefficients cannot be achieved by using heteroskedasticity-consistent (HC) standard errors, but only by a generalized least squares estimate (GLS estimate). However, the accuracy advantage only occurs if the source and form of heteroskedasticity are known and can therefore be explicitly taken into account in the GLS estimate.

5.9 Why is the original White estimator usually used in a modified form in an HC estimation of the variance-covariance matrix Cov(\(\hat{\beta}\))? Explain the three modifications of the White estimator available in the mathematical and statistical program R!

Solution:
The original White estimator for Cov(\(\hat{\beta}\)) is consistent, i.e. it approaches the variance-covariance matrix of the estimated regression coefficients better and better as the sample size grows. Consistency includes a statement about the asymptotic behavior of the estimator, but does not yet say anything about its properties in finite samples. Simulation studies suggest that the original white correction underestimates the "true" variances of the estimated regression coefficients in small samples. For this reason, various finiteness corrections have been proposed to allow the White estimator to be used even in the case of small samples if heteroskedasticity is present after modification. In program R, the White estimator is available in the car and sandwich packages under the names "hc0" and "HC0" respectively. The different modifications of the White estimator in the two R packages, "hc1", "hc2", "hc3" and "hc4" or "HC1", "HC2", "HC3" and "HC4", contain different forms of finiteness correction. Here the explanation is limited to the first three modifications of the White estimator.

Der HC1-Schätzer ("hc1" bzw. "HC1") geht davon aus, dass Erwartungstreue oft durch Ersetzung der Beobachtungszahl n durch die Anzahl der Freiheitsgrade hergestellt.
werden kann. Da einer Berechnung der Varianzen $\hat{V}(\hat{\beta}_j)$ aufgrund der Schätzung der Regressionskoefizienten $\beta_j, 1=1,2,\ldots,k$, $k$ Restriktionen zugrunde liegen, wird nach diesem Prinzip der Faktor 1/n in der $\hat{S}_0$-Matrix durch den Faktor 1/(n-k) ersetzt.

The HC1 estimator ("hc1" or "HC1") assumes that expectation fidelity can often be achieved by replacing the observation number $n$ with the number of degrees of freedom. Since a calculation of the variances $\hat{V}(\hat{\beta}_j)$ is based on the estimation of the regression coefficients $\beta_j, 1=1,2,\ldots,k$, $k$ restrictions, the factor 1/n in the matrix is replaced by the factor 1/(n-k) according to this principle.

The HC2 estimator assumes a differentiation between the variance of the disturbances $u_t$, $\text{Var}(u_t) = \sigma^2_t$ and the variance of the OLS residues $\hat{u}_t$, $\text{Var}(\hat{u}_t) = \sigma^2_t(1-h_{tt})$. The quantities $h_{tt}$ are the main diagonal elements of the Hat-Matrix (Projection matrix) $H = X(X'X)^{-1}X'$. Because of $1/n \leq h_{tt} \leq 1$ is the variances $\text{Var}(\hat{u}_t)$ of the OLS residuals generally smaller than the variances $\sigma^2_t$ of the disturbance variables. Accordingly, if the quantities $\hat{u}_t^2/(1-h_{tt})$ are used instead of the squared residuals $\hat{u}_t^2$ to estimate the variances $\sigma^2_t$ of the disturbance variables, a reduction of the underestimation in finite samples can be expected.

With the HC3 estimator, the denominator size $\hat{u}_t^2$ of the HC2 estimator is used in square form analogous to the counter size $\hat{u}_t^2/(1-h_{tt})^2$. Formally, however, this correction is more difficult to justify with the so-called jackknife technique. In simulation studies, the HC3 estimator shows a superior performance in the case of small samples, which has led to this estimator being the default setting in the R functions car and sandwich.

5.10 Some modifications of the White estimator make use of the hat matrix (projection matrix) $H = X(X'X)^{-1}X'$.

a) Demonstrate that the vector of the regression values $\hat{y}$ can be calculated from the vector of the dependent variable $y$ using the hat matrix $H$!

Solution:

With the OLS estimator $\hat{\beta}$,

1) $\hat{\beta} = (X'X)^{-1}X'y$

the vector of the regression values can be modified by $\hat{y}$

2) $\hat{y} = X \cdot \hat{\beta}$

If one inserts (1) into (2), one obtains the relation
\[ \hat{y} = X(X'X)^{-1}X' y, \]

which is calculated using the definition of the hat-matrix \( H \),

\[ H = X(X'X)^{-1}X', \]

write in the form of

(3) \( \hat{y} = H \cdot y \)

b) Show the relationship between the vector of the OLS residuals, \( \hat{u} \), and the vector of the disturbances, \( u \), using the hat matrix \( H \!

**Solution:**

The vector of the OLS residues, \( \hat{u} \), is the difference between the vector of the dependent variable, \( y \), and the vector of the regression values, \( \hat{y} \):

(4) \( \hat{u} = y - \hat{y} \).

If equation (3) from part a) is used here, the result is the equation

\[ \hat{u} = y - H \cdot y, \]

which, after omitting \( y \) from the relationship

\[ \hat{u} = (I - H) \cdot y \]

with \( I \) as nxn unit matrix.