6.4 Generalized Least Squares estimation

6.4.1 Autocorrelation and GLS estimation

A **generalized least squares estimator (GLS estimator)** for the vector of the regression coefficients, $\boldsymbol{\beta}$, can be be determined with the help of a specification of the variance-covariance matrix of the error terms, Cov ($\mathbf{u}$), similar to what was shown for heteroscedasticity in section 5.3.1. Also in the presence of autocorrelation will it be possible to decompose Cov($\mathbf{u}$) into a product consisting of a scalar factor σ^2 and a matrix $\boldsymbol{\Omega}$, as indicated in (5.18):

(5.18) Cov ($\mathbf{u}$) = E ($\mathbf{u}\mathbf{u}'$) = $\sigma^2 \cdot \boldsymbol{\Omega}$.

While the covariance matrix Cov($\mathbf{u}$) and thus also the matrix $\boldsymbol{\Omega}$ in the presence of heteroscedasticity is a diagonal matrix, i.e. it has different elements along the main diagonal while the off-diagonal elements are 0, this is not the case in the presence of autocorrelation.

The main diagonal of Cov($\mathbf{u}$) contains the variances σ^2_t of the disturbance variables u_t, which, however, in the case of homoskedasticity are all identically equal to σ^2:

$$\sigma^2_1 = \sigma^2_2 = ... = \sigma^2_n = \sigma^2 .$$

The off-diagonal elements are the autocovariances of the j-th order, σ_j, j=t-s, $t \neq s$, of the disturbance variable u. Due to the definition of theoretical the autocorrelation coefficient of the j-th order,

$$(6.18) \qquad \rho_j = \frac{\mathrm{Cov}(u_t, u_s)}{\sqrt{\mathrm{Var}(u_t)} \cdot \sqrt{\mathrm{Var}(u_s)}} = \frac{\sigma_j}{\sigma^2}, \; j = t - s; t \neq s,$$

they can be expressed as the product of the variance of the disturbance variables, σ^2, and the autocorrelation coefficient ρ_j.

$$\sigma_j = \sigma^2 \cdot \rho_j, \; j \neq 0\text{th}$$

Definition (6.18) is grounded on the assumption of homoscedasticity. In addition it is assumed that the dependence between the disturbance variables of the different periods solely depends on the time interval j=t-s.

If the disturbance variable u of the multiple regression model

$$(2.1b) \qquad y_t = \mathbf{x}_t' \cdot \boldsymbol{\beta} + u_t$$

follows a first-order autoregressive process (Markov process),

$$(6.6) \qquad u_t = \phi \cdot u_{t-1} + v_t, \;\; |\phi| < 1, \;\; t = 2,\ldots,n.$$

then the theoretical autocorrelation coefficients are given by

$$(6.19) \qquad \rho_j = \phi^j, \; j = 0,1,2,\ldots$$

(for the derivation see excursion). After that, the strength of the connections between the temporally separated values of the disturbance variables u weakens exponentially with increasing time difference. If the autoregressive parameter ϕ for example equals 0,6, then the autocorrelation of immediate subsequent disturbance terms amounts to 0,6. In the event of a time interval of two periods it amounts to 0,36 (= 0,6²), for a time interval of 3 periods 0,216 (= 0,6³), etc.

The **variance-covariance matrix** of the disturbance terms for a **first-order auto-regressive process (Markov process)** is given by

(6.20)

$$\text{Cov}(\mathbf{u}) = \sigma^2 \mathbf{\Omega} = \begin{pmatrix} \sigma^2 & \phi\sigma^2 & \cdots & \phi^{n-1}\sigma^2 \\ \phi\sigma^2 & \sigma^2 & \cdots & \phi^{n-2}\sigma^2 \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1}\sigma^2 & \phi^{n-2}\sigma^2 & \cdots & \sigma^2 \end{pmatrix} = \sigma^2 \cdot \begin{pmatrix} 1 & \phi & \cdots & \phi^{n-1} \\ \phi & 1 & \cdots & \phi^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \cdots & 1 \end{pmatrix}$$

If one now inserts in (6.6) the OLS residuals after an estimation of the multiple regression model (2.1b), then the autoregressive parameter ϕ can be estimated by means of the auxiliary regression of $\hat{u}_t$ on $\hat{u}_{t-1}$:

(6.9) $\qquad \hat{u}_t = \phi \cdot \hat{u}_{t-1} + v_t , \quad |\phi| < 1 ,$

One now replaces the unknown autoregressive parameter ϕ in the matrix $\boldsymbol{\Omega}$ by the OLS estimator

(6.10) $\qquad \hat{\phi} = \dfrac{\dfrac{1}{n-1} \sum\limits_{t=2}^{n} \hat{u}_t \hat{u}_{t-1} - \bar{\hat{u}}_0 \bar{\hat{u}}_1}{\dfrac{1}{n-1} \sum\limits_{t=2}^{n} \hat{u}_{t-1}^2 - \bar{\hat{u}}_1^2}$.

With the estimated $\boldsymbol{\Omega}$-matrix

(6.21) $\qquad \hat{\boldsymbol{\Omega}} = \begin{pmatrix} 1 & \hat{\phi} & \cdots & \hat{\phi}^{n-1} \\ \hat{\phi} & 1 & \cdots & \hat{\phi}^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\phi}^{n-1} & \hat{\phi}^{n-2} & \cdots & 1 \end{pmatrix}$

the GLS estimator for $\boldsymbol{\beta}$, specified in (5.27b), can be determined:

(5.27c) $\qquad \hat{\boldsymbol{\beta}}_{GLS} = \left(\mathbf{X}' \hat{\boldsymbol{\Omega}}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \hat{\boldsymbol{\Omega}}^{-1} \mathbf{y}$.

With this, the principle of GLS estimation in the presence of autocorrelated disturbance terms has been outlined. In the econometric practice, however, one usually does not directly apply the GLS estimator (5.27c), but prefers instead methods of variable transformation. These approaches shall now be outlined and illustrated.

6.4.2 Cochran-Orcutt method

The starting point of the **Cochran-Orcutt method** is the multiple regression model

(2.1b) $\qquad y_t = \mathbf{x}_t' \cdot \boldsymbol{\beta} + u_t$

with the first-order autoregressive disturbance process (Markov-process)

(6.6) $\qquad u_t = \phi \cdot u_{t-1} + v_t$,

If one lags (2.1b) by one period, i.e. one considers the multiple regression model in the period t-1, then one obtains

$$y_{t\text{-}1} = \mathbf{x}_{t\text{-}1}' \cdot \boldsymbol{\beta} + u_{t\text{-}1}$$

and after multiplication by the autoregressive parameter ϕ

(6.22) $\qquad \phi y_{t-1} = \phi\left(\mathbf{x}_{t-1}' \boldsymbol{\beta}\right) + \phi u_{t-1}$,

If one now subtracts (6.22) from the original model (2.1b), then the transformed model

$$y_t - \phi y_{t-1} = \mathbf{x}_t' \boldsymbol{\beta} - \phi \mathbf{x}_{t-1}' \boldsymbol{\beta} + u_t - \phi u_{t-1}$$

or

(6.23) $\qquad y_t - \phi y_{t-1} = \left(\mathbf{x}_t' - \phi \mathbf{x}_{t-1}'\right)\boldsymbol{\beta} + v_t$

,

is obtained.

In (6.23) the disturbance variable v_t (= u_t - $\phi\, u_{t-1}$) satisfies the standard assumptions of an econometric one equation model. To estimate the **transformed regression model** with the OLS method, the autoregressive parameter must be known. As can be shown, this form of model estimation describes a GLS estimation.

The **Cochran-Orcutt approach** is composed of several steps.

<u>Step 1</u>

OLS estimation of the original regression model (2.1b), from which, with the estimation vector $\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'y}$, the OLS residuals emerge

$$\hat{u}_t = y_t - \mathbf{x}_t'\hat{\beta} \ .$$

<u>Step 2</u>

Replacing the disturbance variables u_t with the OLS residuals $\hat{u}_t$ and the implementation of the auxiliary regression

(6.9) $\qquad\qquad \hat{u}_t = \phi\hat{u}_{t-1} + v_t,$

from which one obtains the OLS estimator $\hat{\phi}$ for the autoregressive parameter ϕ following (6.10).

Step 3

Transformation of the variables:

(6.24a)
$$y_t^* = y_t - \hat{\phi} \cdot y_{t-1}, \quad t = 2,3,\ldots,n$$
$$\mathbf{x}_t^{*\prime} = \mathbf{x}_t^{\prime} - \hat{\phi} \cdot \mathbf{x}_{t-1}^{\prime}$$

Step 4

OLS estimation of the transformed regression model

(6.25)
$$y_t^* = \mathbf{x}_t^{*\prime} \boldsymbol{\beta} + v_t$$

leads to the GLS estimator of the Cochran-Orcutt approach:

(6.26a)
$$\hat{\boldsymbol{\beta}}_{GLS}^{CO} = \left(\underset{kx(n-1)}{\mathbf{X}^{*\prime}} \underset{(n-1)xk}{\mathbf{X}^*} \right)^{-1} \underset{kx(n-1)}{\mathbf{X}^{*\prime}} \underset{(n-1)x1}{\mathbf{y}^*}$$

with

$$\underset{(n-1)xk}{\mathbf{X}^*} = \begin{bmatrix} \mathbf{x}_2^{*\prime} \\ \mathbf{x}_3^{*\prime} \\ \vdots \\ \mathbf{x}_n^{*\prime} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_2^{\prime} - \hat{\phi} \cdot \mathbf{x}_1^{\prime} \\ \mathbf{x}_3^{\prime} - \hat{\phi} \cdot \mathbf{x}_2^{\prime} \\ \vdots \\ \mathbf{x}_n^{\prime} - \hat{\phi} \cdot \mathbf{x}_{n-1}^{\prime} \end{bmatrix} \quad \text{and} \quad \underset{(n-1)x1}{\mathbf{y}^*} = \begin{bmatrix} y_2^* \\ y_3^* \\ \vdots \\ y_n^* \end{bmatrix} = \begin{bmatrix} y_2 - \hat{\phi} \cdot y_1 \\ y_3 - \hat{\phi} \cdot y_2 \\ \vdots \\ y_n - \hat{\phi} \cdot y_{n-1} \end{bmatrix},$$

It is in this connection to be noted that the transformed observation matrix $\mathbf{X}^*$ is a (n-1)xk matrix and the transformed vector of the dependent variable, $\mathbf{y}^*$, is a (n-1)x1 vector. By the transformations (6.24a) are the number of observations reduced from n to n-1.

The Cochran-Orcutt approach can also be iterated, i.e. with the GLS-estimator (6.26a) one obtains the GLS-residuals with which an "improved" estimator of the autoregressive parameter ϕ can be determined. This may be improved anew in a 2nd, 3rd, ... iteration step until the process converges.

Excursion

GLS estimator (6.26a) of the Cochran-Orcutt method using the transformation matrix **T**:

(6.26b) $\hat{\boldsymbol{\beta}}_{\substack{GLS \\ kx1}}^{CO} = \left(\underset{kx(n-1)}{\mathbf{X}^{*\prime}} \underset{(n-1)xk}{\mathbf{X}^{*}}\right)^{-1} \underset{kx(n-1)}{\mathbf{X}^{*\prime}} \underset{(n-1)x1}{\mathbf{y}^{*}}$

$= (\underset{kxn}{\mathbf{X}^{\prime}} \underset{nxn}{\hat{\boldsymbol{\Omega}}^{-1}} \underset{nxk}{\mathbf{X}})^{-1} \underset{kxn}{\mathbf{X}^{\prime}} \underset{nxn}{\hat{\boldsymbol{\Omega}}^{-1}} \underset{nx1}{\mathbf{y}}$

with $\hat{\boldsymbol{\Omega}}^{-1} = \mathbf{T}^{\prime}\mathbf{T}$

$$\underset{(n-1)xn}{\mathbf{T}} = \begin{bmatrix} -\hat{\phi} & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\hat{\phi} & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\hat{\phi} & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\hat{\phi} & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\hat{\phi} & 1 \end{bmatrix}$$

□

Example:

The Durbin-Watson test and the Breusch-Godfrey test have shown that the disturbance term of the Keynesian consumption function is autocorrelated. Specifically both tests show an autocorrelation of first order. To take account of autocorrelation in the econometric estimation of the consumption function, we make use of the **Cochran-Orcutt approach**.

Step 1

We first estimate the Keynesian consumption function with the OLS method

$$\hat{C}_t = \underset{(-2,215)}{-38{,}522} + \underset{(75{,}379)}{0{,}934} \cdot Y_t^v ,$$

$$R^2 = 0{,}997 , \quad DW = 0{,}488 ,$$

and determine the OLS residuals:

$$\hat{u}_t = C_t - (-38{,}522 + 0{,}934 \cdot Y_t^v)$$

Step 2

To get an estimator for the autoregressive parameter ϕ, we regress the OLS residuals $\hat{u}_t$ of the Keynesian consumption function on the one-period lagged OLS residuals $\hat{u}_{t-1}$:

$$\hat{\hat{u}}_t = \underset{(4{,}533)}{0{,}705} \cdot \hat{u}_{t-1} ,$$

$$R^2 = 0{,}547 , \quad DW = 1{,}482 .$$

With the estimated autoregressive parameter

$$\hat{\phi} = 0,705$$

we form the generalized differences

$$C_t^* = C_t - 0,705 \cdot C_{t-1}$$

and

$$Y_t^{*v} = Y_t^v - 0,705 \cdot Y_{t-1}^v$$

An OLS estimation of the transformed regression model

$$\hat{C}_t^* = \beta_1 \cdot (1 - 0,705) + \beta_2 \cdot Y_t^{v^*}$$

gives the GLS estimator on the basis of the Cochran-Orcutt method:

$$\hat{\beta}_1^* = \hat{\beta}_1 \cdot (1 - 0,705) = -1,541 \quad (t = -0,133)$$

and

$$\hat{\beta}_2^* = \hat{\beta}_2 = 0,913 \quad (t = 34,589)$$

While the marginal consumption propensity directly matches with the estimator $\hat{\beta}_2^*$ of the transformed regression equation, the estimator for autonomous consumption is dervied from the transformation

$$\hat{\beta}_1 = \hat{\beta}_1^* / (1 - 0,705) = -5,231$$

For autonomous consumption (= constant term) one obtains therefore by taking into account the autocorrelation of the residuals a considerably smaller value as in the simple OLS estimation; in addition, the marginal consumption ratio (= slope measure) falls by about 2 percentage points.

The **Durbin-Watson test** discloses that with the assistance of the Cochran-Orcutt approach a clean up of the autocorrelation has succesfully been carried out, because the resulting DW statistic of 1,492 is now located in the acceptance region ($\alpha = 0,05$; n = 18; k = 2):

$$DW = 1,492 \in [d_0 = 1,39;\ 4 - d_0 = 2,61]$$

The **coefficient of determination** $R^2_{GLS\ *}$ of 0,987 relates solely to the goodness of fit of the transformed regression model, similar to what applied in the employment of the weighted least squares method in the case of heteroscedasticity. Here as there a very analogous a coefficient of determination R^2_{GLS} for the GLS estimated consumption function can be determined, which at a value of 0,954 remains below that of the OLS estimation. □

6.4.3 Prais-Winston method

The **Prais-Winston method** differs solely from the Cochran-Orcutt method in that now the GLS estimation of the coefficient vector $\boldsymbol{\beta}$ is carried out not with n-1 but with all n available observations of the variables. It improves thus the efficiency of the Cochran-Orcutt approach, which can be especially advantageous in small samples, that are typically related to annual data.

The OLS estimation of the transformed regression model

(6.25) $$y_t^* = \mathbf{x}_t^{*'} \boldsymbol{\beta} + v_t$$

With the Prais-Winston method is carried out after an extended variable trans-formation:

(6.24b)

$$y_1^* = \sqrt{1 - \hat{\phi}^2} \cdot y_1 \quad \text{und} \quad \mathbf{x}_1^* = \sqrt{1 - \hat{\phi}^2} \cdot \mathbf{x}_1$$

$$y_t^* = y_t - \hat{\phi} \cdot y_{t-1} \quad \text{und} \quad \mathbf{x}_t^* = \mathbf{x}_t - \hat{\phi} \cdot \mathbf{x}_{t-1}, \quad t = 2,3,\ldots, n$$

The complete transformation (6.24b) can be motivated by the fact that the transformation matrix in the case of a first-order autocorrelation of the disturbance terms is hereby exactly specified after the estimation of the autoregressive para-meter ϕ. From this point of view, the Cochran-Orcutt method constitutes a simplified GLS method.

The GLS estimator β of the Prais-Winston method is with (6.24b) and (6.25) given by

(6.27a) $\qquad \hat{\boldsymbol{\beta}}^{PW}_{\substack{GLS \\ kx1}} = \left(\underset{kxn \ nxk}{\mathbf{X}^{*\prime} \, \mathbf{X}^{*}} \right)^{-1} \underset{kxn \ nx1}{\mathbf{X}^{*\prime} \, \mathbf{y}^{*}}$

. $\mathbf{X}^{*}$ is herein the transformed observation matrix of dimension nxk and $\mathbf{y}^{*}$ is the transformed vector of the dependent variable of dimension nx1.

Example:

The Keynesian consumption function shall now be estimated with the **Prais-Winston method**. Steps 1 and 2 do not herein differ from those of the Cochran-Orcutt approach.

<u>Step 1</u>

We first estimate the Keynesian consumption function with the OLS method

$$\hat{C}_t = - 38{,}522 + \underset{(75{,}379)}{0{,}934} \cdot Y_t^v,$$
$$\phantom{\hat{C}_t = }\underset{(-2{,}215)}{}$$

$$R^2 = 0{,}997, \quad DW = 0{,}488,$$

and determine the OLS residuals:

$$\hat{u}_t = C_t - (-38{,}522 + 0{,}934 \cdot Y_t^v)$$

<u>Step 2</u>

To get an estimator for the autoregressive parameter ϕ, we regress the OLS residuals $\hat{u}_t$ of the Keynesian consumption function on the one-period lagged OLS residuals $\hat{u}_{t-1}$:

$$\hat{\hat{u}}_t = \underset{(4{,}533)}{0{,}705} \cdot \hat{u}_{t-1},$$

$$R^2 = 0{,}547, \quad DW = 1{,}482.$$

Step 3

With the estimated autoregressive parameter

$$\hat{\phi} = 0{,}705$$

we form the generalized differences

$$C_1^* = \sqrt{1 - 0{,}705^2} \cdot C_t = 0{,}709 \cdot C_t \quad \text{und} \quad Y_1^{v^*} = \sqrt{1 - 0{,}705^2} \cdot Y_t^v = 0{,}709 \cdot Y_t^v$$

and

$$C_t^* = C_t - 0{,}705 \cdot C_{t-1} \quad \text{und} \quad Y_t^{v^*} = Y_t^v - 0{,}705 \cdot Y_{t-1}^v \quad \text{für } t = 2{,}3{,}\ldots{,}n = 19$$

Step 4

An OLS estimation of the transformed regression model

$$\hat{C}_t^* = \beta_1^* + (\beta_2^* = \beta_2) \cdot Y_t^{v^*}$$

gives the Prais-Winston method's GLS estimator:

$$\hat{\beta}_1^* = 12{,}998 \quad \left(t = 2{,}23\right)$$

and

$$\hat{\beta}_2^* = \hat{\beta}_2 = 0{,}879 \quad \left(t = 70{,}24\right)$$

The GLS estimator of the marginal propensity to consume based on the Prais-Winston method is lower than the original OLS estimator and is also below that of the Cochran-Orcutt approach. The absolute term of the original consumption function can be determined here by using a weighting procedure:

$$\hat{\beta}_1 = \frac{1}{n} \frac{\hat{\beta}_1^*}{\sqrt{1-0{,}705^2}} + \frac{n-1}{n} \frac{\hat{\beta}_1^*}{1-0{,}705} = \frac{1}{19} \frac{12{,}998}{0{,}709} + \frac{18}{19} \frac{12{,}998}{0{,}295} = 42{,}707$$

According to the standardized testing approach, the **Durbin-Watson statistic** of 1.222 lies (without regard to the specific x values) in the uncertainty area ($\alpha = 0{,}05$; n = 19; k = 2):

$$DW = 1{,}222 \in [d_u = 1{,}18; d_o = 1{,}40]$$

However, if one carries out the DW test, for example with the program R, by taking account of the exact x-values, no significance at the 5% level (p = 0.086) is obtained. Subsequently, the null hypothesis of an absence of a first-order autocorrelation of the disturbance variables cannot be rejected.

With a coefficient of determination $R^2_{GLS\,*}$ of 0,997 in the transformed regression model of the consumption function the goodness of fit is slightly better than with the Cochran-Orcutt method.

GLS estimator (6.27a) of the Prais-Winston approach using the transformation matrix **T:**

$$(6.27b) \qquad \hat{\boldsymbol{\beta}}_{\substack{GLS \\ kx1}}^{PW} = \left(\underset{kxn}{\mathbf{X}^{*\prime}} \underset{nxk}{\mathbf{X}^{*}}\right)^{-1} \underset{kxn}{\mathbf{X}^{*\prime}} \underset{nx1}{\mathbf{y}^{*}} = (\underset{kxn}{\mathbf{X}^{\prime}} \underset{nxn}{\hat{\boldsymbol{\Omega}}^{-1}} \underset{nxk}{\mathbf{X}})^{-1} \underset{kxn}{\mathbf{X}^{\prime}} \underset{nxn}{\hat{\boldsymbol{\Omega}}^{-1}} \underset{nx1}{\mathbf{y}}$$

with $\qquad \hat{\boldsymbol{\Omega}}^{-1} = \mathbf{T}^{\prime}\mathbf{T}$

● Variance-covariance-matrix of disturbance variable u by first-order autocorrelation:

$$(6.20)$$

$$\mathrm{Cov}(\mathbf{u}) = \sigma^2 \boldsymbol{\Omega} = \begin{pmatrix} \sigma^2 & \phi\sigma^2 & \cdots & \phi^{n-1}\sigma^2 \\ \phi\sigma^2 & \sigma^2 & \cdots & \phi^{n-2}\sigma^2 \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1}\sigma^2 & \phi^{n-2}\sigma^2 & \cdots & \sigma^2 \end{pmatrix} = \sigma^2 \cdot \begin{pmatrix} 1 & \phi & \cdots & \phi^{n-1} \\ \phi & 1 & \cdots & \phi^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \cdots & 1 \end{pmatrix}$$

with $\qquad \boldsymbol{\Omega} = \begin{pmatrix} 1 & \phi & \cdots & \phi^{n-1} \\ \phi & 1 & \cdots & \phi^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \cdots & 1 \end{pmatrix}$

Estimated inverse of **Ω**:

$$
\underset{n \times n}{\hat{\boldsymbol{\Omega}}}^{-1} = \mathbf{T'T} =
\begin{bmatrix}
1 & -\hat{\phi} & 0 & \cdots & 0 & 0 \\
-\hat{\phi} & 1+\hat{\phi}^2 & -\hat{\phi} & \cdots & 0 & 0 \\
0 & -\hat{\phi} & 1+\hat{\phi}^2 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1+\hat{\phi} & -\hat{\phi} \\
0 & 0 & 0 & \cdots & -\hat{\phi} & 1
\end{bmatrix}
$$

Decomposition of the inverse **Ω**$^{-1}$ in the form of **Ω**$^{-1}$ = **T'T** implies

$$
\underset{nxn}{\mathbf{T}} =
\begin{bmatrix}
\sqrt{1-\hat{\phi}^2} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
-\hat{\phi} & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & -\hat{\phi} & 1 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & -\hat{\phi} & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & -\hat{\phi} & 1
\end{bmatrix}
$$

□

6.5 Heteroscedasticity and Autocorrelation Consistent (HAC) Variance-Covariance Matrix

As an alternative to using the generalized method of least squares (GLS method), for example, in the form of the Cochran-Orcutt or Prais-Winston a correction of the estimator of the variance-covariance matrix $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ by

(5.33) $\qquad \mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X'X})^{-1} \mathbf{X' \Omega X} (\mathbf{X'X})^{-1}$

lends itself because of the unbiasedness of the OLS estimator $\hat{\boldsymbol{\beta}}$ in the case of auto-correlation. In (5.33) the generalized variance-covariance matrix of the disturbance terms defined in (6.1), $\mathrm{Cov}(\mathbf{u}) = \sigma^2 \cdot \boldsymbol{\Omega}$, is used.

Autocorrelation-consistent estimators of $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$, however, typically control for simul-taneously for potentially existing heteroscedasticity, so that they are based on the general form of the variance-covariance matrix of the disturbance terms:

(6.28) $\qquad \mathrm{Cov}(\mathbf{u}) = \sigma^2 \cdot \boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$

Using the definition

$$(6.29) \quad \mathbf{S}_X = \sigma^2 (\mathbf{X'\Omega X})/n = \left(\sum_t \sum_s \sigma_{ts} \mathbf{x}_t \mathbf{x}_s' \right)/n \; .$$

in which the vector $\mathbf{x}_t$ depicts the values of the exogenous variables in period t (= t-th row of the observation matrix $\mathbf{X}$), the variance-covariance matrix $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ can be presented in the form

$$(5.33a) \quad \mathrm{Cov}(\hat{\boldsymbol{\beta}}) = n(\mathbf{X'X})^{-1} \mathbf{S}_X (\mathbf{X'X})^{-1}$$

Given the quantities n and $\mathbf{X,}$ an estimation of $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ is carried out with an estimator for $\mathbf{S}_X$. To this end, Newey and West (1987) have developed the **heteroskedasticity and autocorrelation consistent (HAC) estimator**

$$(6.30) \; \hat{\mathbf{S}}_X = \left(\sum_{t=1}^{n} \hat{u}_t^2 \mathbf{x}_t \mathbf{x}_t' \right)/n + \left(\sum_{j=1}^{p} \sum_{t=j+1}^{n} (1 - w_j) \hat{u}_t \hat{u}_{t-j} (\mathbf{x}_t \mathbf{x}_{t-j}' + \mathbf{x}_{t-j} \mathbf{x}_t') \right)/n$$

with the weights

$$(6.31) \qquad\qquad w_j = \frac{j}{p+1} \; .$$

.

In (3.31) p specifies the maximum lag that needs to be considered for the control of autocorrelation. Folowing Greene (2003), a common value of p is given by the choice of $p \approx n^{1/4}$.

The consistency of the HAC estimator of the variance-covariance matrix $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ with (30.06) is an asymptotic property. An adjustment for small samples has been proposed using the factor n/(n - k) analogously to the HC1-estimator with heteroscedasticity.

When considering exclusively first-order autocorrelation (p=1), the HAC-estimator (6.31) of Newey and West simplifies to

$$\hat{\mathbf{S}}_X = \left(\sum_{t=1}^{n} \hat{u}_t^2 \mathbf{x}_t \mathbf{x}_t' \right)/n + \left(0.5 \cdot \sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1} (\mathbf{x}_t \mathbf{x}_{t-1}' + \mathbf{x}_{t-1} \mathbf{x}_t') \right)/n$$

In the special case of p=0, autocorrelation is no longer controlled for, so that (6.30) translates into White's heteroskedasticity-consistent (HC) estimator:

$$\hat{\mathbf{S}}_X = \left(\sum_{t=1}^{n} \hat{u}_t^2 \mathbf{x}_t \mathbf{x}_t' \right)/n .$$

Newey/West HAC estimator (6.30) for $\mathrm{Var}(\hat{\beta}_2)$ in the case of simple regression:

$$\widehat{\mathrm{Var}}(\hat{\beta}_2)_{\mathrm{HAC}} = \frac{1}{[\sum_{t=1}^{n}(x_t - \bar{x})^2]^2} \left[\sum_{t=1}^{n} \hat{u}_t^2 (x_t - \bar{x})^2 + \sum_{j=1}^{p} \sum_{t=j+1}^{n} (1 - w_j) \hat{u}_t \hat{u}_{t-j} (x_t - \bar{x})(x_{t-j} - \bar{x}) \right]$$

Example:

Concerning the OLS estimated Keynesian consumption function for the period 1994-2012 for Germany, both tests for autocorrelation, the Durbin-Watson test and the Breusch-Godfrey test, result in an unambiguous rejection of the null hypothesis of an absence of autocorrelation of the disturbance term. The conventionally performed significance tests thereby loose their validity. One possible remedy is offered by the application of the generalized least squares method (GLS-method) in the form of the Cochran-Orcutt or Prais-Winston approaches. Both transformation approaches aim to override first-order autocorrelation in the estimation of the influence of disposable income on private consumption.

An alternative is to correct the standard errors of the OLS estimated regression coefficients as in the case of heteroscedasticity. With the heteroscedasticity- and autocorrelation-consistent Newey-West estimator (6.30) for S_x, the HAC-estimator for $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ arises,

$$
\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}})_{\mathrm{NW}} = \begin{array}{c} \text{Const. term} \qquad \text{Disp. inc.} \\ \begin{bmatrix} 439{,}447514 & -0{,}295162 \\ -0{,}295162 & 0{,}000201 \end{bmatrix} \begin{array}{l} \text{Const. term} \\ \text{Disp. inc.} \end{array} \end{array}
$$

using the automatic selectivity procedure of Newey and West. Here it has been calculated with the function NeweyWest of the R package sandwich :

[R command:> = covbKons.nw NeweyWest (Konsumfkt9412.lm, prewhite = FALSE)]

Taking into account the adjustment factor n/(n-k) for finite samples, the HAC estimator according to Newey and West is

$$\text{Const. term} \quad \text{Disp. inc}$$

$$\overset{\wedge}{C\,o\,v}(\hat{\boldsymbol{\beta}})_{\text{NWa}} = \begin{bmatrix} 491{,}147221 & -0{,}329887 \\ -0{,}329887 & 0.000224 \end{bmatrix} \begin{array}{l} \text{Const. term} \\ \text{Disp. inc.} \end{array}$$

[R command ::> = covbKons.nwa NeweyWest (Konsumfkt9412.lm, prewhite = FALSE, adjust = TRUE)]

The selection procedure according to Newey and West displays a relevance with respect to 1st and 2nd order autocorrelation in the residuals of the consumption function, which has been taken into account in the determination of both HAC-estimators.

The results of the HAC estimation

HAC standard error (with adjustment): $\sqrt{\overset{\wedge}{V\,a\,r}(\hat{\beta}_1)_{\text{NWa}}} = 22{,}16184$ and $\sqrt{\overset{\wedge}{V\,a\,r}(\hat{\beta}_2)_{\text{NWa}}} = 0{,}0149808$

suggest that the conventional standard errors of the OLS estimated regression coefficients

conventional Standard Error: $\sqrt{\overset{\wedge}{V\,a\,r}(\hat{\beta}_1)} = 17{,}39254$ and $\sqrt{\overset{\wedge}{V\,a\,r}(\hat{\beta}_2)} = 0{,}012385$

lead to a significant underestimation of the corresponding population parameters. With the conventional estimation such an underestimation of the "true" standard error can always be expected in economic time series, which are characterized by slow changes in the trend over time.

Compact can the significance tests of the OLS estimated regression coefficients be retrieved using the Newton/West-HAC estimator of the standard error, for example, with the function coeftest of the R-package lmtest. For the Keynesian consumption function in the period 1994-2012 one obtains:

t test of coefficients (Newey/West HAC std. error with adjustment):

	Estimate	Std. Error	t value	Pr (> | t |)
(Intercept)	-38.522194	22.161842	-1.7382	0.11
Yvr	0.933535	0.014981	62.3155	2.25e-15 ***

Signif. codes: 0 ' *** ' 0.001 ' ** ' 0.01 ' * ' 0.05 ' . ' 0.1 ' ' 1

[R command:> = coeftestKons.NWt coeftest (Konsumfkt9412.lm, df = nk, vcov = NeweyWest (Konsumfkt9412.lm, prewhite = FALSE, adjust = TRUE))]

When using the HAC standard error according to Newey and West the constant term of the (long-term) consumption function is - as theoretically is to be expected - no longer significant.