6.4 Verallgemeinerte Kleinst-Quadrate-Schätzung

6.4.1 Autokorrelation und GLS-Schätzung

Ein **verallgemeinerter Kleinst-Quadrate-Schätzer (GLS-Schätzer)** für den Vektor der Regressionskoeffizienten, **β**, ließe sich unter Spezifikation der Varianz-Kovarianz-Matrix der Störterme, Cov(**u**), wie bei der Heteroskedastizität in Abschn. 5.3.1 aufgezeigt bestimmen. Auch bei Autokorrelation lässt Cov(**u**) in ein Produkt aus einem skalaren Faktor σ^2 und einer Matrix **Ω** wie in (5.18) angegeben zerlegen:

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

Während die Kovarianzmatrix Cov(**u**) und damit auch die Matrix **Ω** bei Heteroskedastizität eine Diagonalmatrix ist, d.h. unterschiedliche Hauptdiagonalelemente hat, während die Nichthauptdiagonalelemente 0 sind, ist dies bei Autokorrelation nicht der Fall.

Die Hauptdiagonale von Cov(**u**) enthält die Varianzen σ^2_t der Störvariablen u_t, die aber bei Homoskedastizität alle identisch gleich σ^2 sind:

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

Die Nichthauptdiagonalelemente sind die Autokovarianzen j-ter Ordnung, σ_j, j=t-s, t≠s, der Störvariablen u. Sie lassen sich aufgrund der Definition der theoretischen Autokorrelationskoeffizienten j-ter Ordnung,

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

als Produkt aus der Varianz der Störvariablen, σ^2, und dem Autokorrelations-koeffizienten ρ_j darstellen.

$$\sigma_j = \sigma^2 \cdot \rho_j,\ \ j \neq 0.$$

Der Definition (6.18) liegt die Annahme der Homoskedastizität zugrunde. Außer-dem ist davon ausgegangen, dass die Abhängigkeit zwischen der Störvariablen unterschiedlicher Perioden allein vom zeitlichen Abstand j=t-s abhängt.

Folgt die Störvariable u des multiplen Regressionsmodells

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

einem autoregressiven Prozess 1. Ordnung (Markov-Prozess),

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

dann sind die theoretische Autokorrelationskoeffizienten durch

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

gegeben (zur Herleitung s. Exkurs). Danach schwächt sich die Stärke des Zusammenhangs zwischen zeitlich auseinander liegenden Werten der Störvariablen u mit zunehmender zeitlicher Differenz exponentiell ab. Ist der autoregressive Parameter ϕ z.B. gleich 0,6, dann beträgt die Autokorrelation zeitlich unmittelbar aufeinander folgender Störterme 0,6. Bei einem zeitlichen Abstand von zwei Perioden beträgt sie 0,36 (=0,6²), bei einem zeitlichen Abstand von 3 Perioden 0,216 (=0,6³), usw.

Die **Varianz-Kovarianz-Matrix** der Störterme ist bei einem **autoregressiven Prozess 1.Ordnung (Markov-Prozess)** durch

(6.20)

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

gegeben. Setzt man nun die OLS-Residuen nach einer Schätzung des multiplen Regressionsmodells (2.1b) in (6.6) ein, dann kann der autoregressive Parameter ϕ durch die Hilfsregression von $\hat{u}_t$ auf $\hat{u}_{t-1}$ geschätzt werden:

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

Man ersetzt nun den unbekannten autoregressiven Parameter ϕ in der Matrix $\mathbf{\Omega}$ durch den OLS-Schätzer

(6.10) $$\hat{\phi} = \frac{\frac{1}{n-1}\sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1} - \bar{\hat{u}}_0\bar{\hat{u}}_1}{\frac{1}{n-1}\sum_{t=2}^{n}\hat{u}_{t-1}^2 - \bar{\hat{u}}_1^2} .$$

Mit der geschätzten $\mathbf{\Omega}$-Matrix

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

lässt sich dann der in (5.27b) angegebene GLS-Schätzer für $\boldsymbol{\beta}$ bestimmen:

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

Damit ist der Prinzip der GLS-Schätzung bei autokorrelierten Störvariablen aufgezeigt. In der ökonometrischen Praxis wendet man jedoch den GLS-Schätzer (5.27c) in der Regel nicht direkt an, sondern bevorzugt Methoden der Variablentransformation. Diese Verfahren sollen mit folgenden aufgezeigt und illustriert werden.

6.4.2 Cochran-Orcutt-Methode

Ausgangspunkt der **Cochran-Orcutt-Methode** ist das multiple Regressionsmodell

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

mit dem autoregressiven Störprozess 1. Ordnung (Markov-Prozess)

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

Verzögert man (2.1.b) um eine Periode, d.h. betrachtet man das multiple Regressionsmodell in der Periode t-1, dann erhält man

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

und nach Multiplikation mit dem autoregressiven Parameter ϕ

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

Subtrahiert man nun (6.22) vom Ausgangsmodell (2.1b), dann ergibt sich

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

oder

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

In (6.23) erfüllt die Störvariable v_t (=$u_t - \phi\ u_{t-1}$) die Standardannahmen eines ökonometrischen Eingleichungsmodells. Um das **transformierte Regressions-modell** mit der OLS-Methode zu schätzen, muss der autoregressive Parameter bekannt sein. Wie gezeigt werden kann, stellt diese Form der Modellschätzung eine GLS-Schätzung dar.

Das **Cochran-Orcutt-Verfahren** setzt sich aus mehreren Schritten zusammen.

<u>1. Schritt</u>

OLS-Schätzung des originären Regressionsmodells (2.1b), woraus mit dem Schätzvektor $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ die OLS-Residuen

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

hervorgehen.

<u>2. Schritt</u>

Ersetzen der Störvariablen u_t durch die OLS-Residuen $\hat{u}_t$ und Durchführung der Hilfsregression

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

womit man den OLS-Schätzer $\hat{\phi}$ für den autoregressiven Parameter ϕ nach (6.10) erhält.

3. Schritt

Transformation der Variablen:

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

$$\mathbf{x}_t^{*'} = \mathbf{x}_t' - \hat{\phi} \cdot \mathbf{x}_{t-1}'$$

4. Schritt

OLS-Schätzung des transformierten Regressionsmodells

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

führt zu dem GLS-Schätzer nach dem Cochran-Orcutt-Verfahren:

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

mit

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

Zu beachten ist hierbei, dass die transformierte Beobachtungsmatrix **X*** eine (n-1)xk-Matrix ist und der transformierte Vektor der abhängigen Variablen, **y***, ein (n-1)x1-Vektor ist. Durch die Transformationen (6.24a) ist die Beobachtungszahl von n auf n-1 reduziert worden.

Das Cochran-Orcutt-Verfahren lässt sich auch iterieren, d.h. mit dem GLS-Schätzer (6.26a) erhält man GLS-Residuen, mit denen ein “verbesserten” Schätzer des autoregressiven Parameters ϕ ermittelt werden kann. Dieser kann in einem 2., 3., … Iterationsschritt erneut verbessert werden bis das Verfahren konvergiert.

<u>Exkurs</u>

GLS-Schätzer (6.26a) des Cochran-Orcutt-Verfahrens unter Verwendung der Transformationsmatrix **T:**

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

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

mit
$$\hat{\boldsymbol{\Omega}}^{-1} = \mathbf{T}'\mathbf{T}$$

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

□

Beispiel:

Der Durbin-Watson-Test und der Breusch-Godfrey-Test haben gezeigt, dass die Störgröße der Keynesschen Konsumfunktion autokorreliert ist. Speziell zeigen beide Tests eine Autokorrelation 1. Ordnung an. Um der Autokorrelation bei der ökonometrischen Schätzung der Konsumfunktion Rechnung zu tragen, wenden wir das **Cochran-Orcutt-Verfahren** an.

<u>1. Schritt</u>

Wir schätzen zunächst die Keynessche Konsumfunktion mit der OLS-Methode,

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

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

und ermitteln die OLS-Residuen:

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

<u>2. Schritt</u>

Um einen Schätzer für den autoregressiven Parameter ϕ zu erhalten, regressieren wir die OLS-Residuen $\hat{u}_t$ der Keynesschen Konsumfunktion auf die um eine Periode verzögerten OLS-Residuen $\hat{u}_{t-1}$:

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

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

3. Schritt

Mit dem geschätzten autoregressiven Parameter

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

bilden wir die verallgemeinerten Differenzen

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

und

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

4. Schritt

Eine OLS-Schätzung des transformierten Regressionsmodells

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

ergibt die GLS-Schätzer nach der Cochran-Orcutt-Methode:

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

und

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

Während die marginale Konsumquote unmittelbar mit dem Schätzer $\hat{\beta}_2^*$ der transformierten Regressionsgleichung übereinstimmt, ergibt sich der Schätzer für den autonomen Konsum aus der Transformation

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

Für den autonomen Konsum (=absolutes Glied) erhält man bei Berücksichtigung der Autokorrelation der Residuen mithin einen absolut erheblich kleineren Wert als bei der einfachen OLS-Schätzung; außerdem sinkt die marginale Konsumquote (=Steigungsmaß um gut 2 Prozentpunkte.

Der **Durbin-Watson-Test** weist aus, dass mit Hilfe des Cochran-Orcutt-Verfahrens eine Bereinigung der Autokorrelation gelungen ist, da die sich ergebende DW-Statistik von 1,492 nun im Annahmebereich liegt (α=0,05; n=18; k=2):

$$\mathrm{DW} = 1{,}492 \in \left[\mathrm{d}_0 = 1{,}39;\ 4 - \mathrm{d}_0 = 2{,}61\right]$$

Der **Determinationskoeffizient** $R^2_{GLS^*}$ von 0,987 bezieht sich hier ähnlich wie bei der Anwendung der gewichteten Kleinst-Quadrate-Methode im Falle der Heteroskedastizität allein auf die Güte der Anpassung des transformierten Regressions-Modells. Hier lässt sich ganz analog wie dort ein Determinationskoeffizent R^2_{GLS} für die GLS-geschätzte Konsumfunktion bestimmen, der bei einem Wert von 0,954 unterhalb desjenigen der OLS-Schätzung bleibt. □

6.4.3 Prais-Winston-Methode

Die **Prais-Winston-Methode** unterscheidet sich von der Cochran-Orcutt-Methode allein darin, dass bei ihr die GLS-Schätzung des Koeffizientenvektors $\boldsymbol{\beta}$ nicht mit mit n-1 sondern allen n vorliegenden Beobachtungen der Variablen erfolgt. Sie verbessert damit die Effizienz des Cochran-Orcutt-Verfahrens, was Insbesondere bei kleinen Stichproben, die vor allem bei Jahresdaten auftreten können, vorteilhaft sein kann.

Die OLS-Schätzung des transformierten Regressionsmodells

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

erfolgt bei der Prais-Winston-Methode nach Durchführung der gegenüber (6.24a) erweiteren Variablentransformationen

(6.24b) $$y_1^* = \sqrt{1-\hat{\phi}^2}\cdot y_1 \quad \text{und} \quad \mathbf{x}_1^* = \sqrt{1-\hat{\phi}^2}\cdot \mathbf{x}_1$$

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

Die vollständige Transformation (6.24b) lässt sich damit begründen, dass die Transformationsmatrix im Falle einer Autokorrelation 1. Ordnung der Störterme nach Schätzung des autoregressiven Parameters ϕ genau hierdurch spezifiziert ist. Aus dieser Sicht stellt sich die Cochran-Orcutt-Methode als vereinfachte GLS-Methode dar.

Der GLS-Schätzer für β der Prais-Winston-Methode ist mit (6.24b) und (6.25) durch

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

gegeben. $\mathbf{X}^*$ ist hierin die transformierte Beobachtungsmatrix der Dimension nxk und $\mathbf{y}^*$ der transformierte Vektor der abhängigen Variablen der Dimension nx1.

Beispiel:

Die Keynessche Konsumfunktion soll jetzt mit der **Prais-Winston-Methode** geschätzt werden. Die Schritte 1 und 2 unterscheiden sich hierbei nicht von denen des Cochran-Orcutt-Verfahrens.

<u>1. Schritt</u>

Wir schätzen zunächst die Keynessche Konsumfunktion mit der OLS-Methode,

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

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

und ermitteln die OLS-Residuen:

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

<u>2. Schritt</u>

Um einen Schätzer für den autoregressiven Parameter ϕ zu erhalten, regressieren wir die OLS-Residuen $\hat{u}_t$ der Keynesschen Konsumfunktion auf die um eine Periode verzögerten OLS-Residuen $\hat{u}_{t-1}$:

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

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

3. Schritt:

Mit dem geschätzten autoregressiven Parameter

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

bilden wir die verallgemeinerten Differenzen

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

und

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

4. Schritt

Eine OLS-Schätzung des transformierten Regressionsmodells

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

ergibt die GLS-Schätzer nach der Prais-Winston-Methode:

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

und

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

Der GLS-Schätzer der marginalen Konsumquote der Prais-Winston-Methode ist niedriger als der originäre OLS-Schätzer und liegt auch unterhalb desjenigen des Cochran-Orcutt-Verfahrens. Das absolute Glied der originären Konsumfunktion könnte man hier unter Verwendung einer Gewichtung bestimmen:

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

Die **Durbin-Watson-Statistik** von 1,222 liegt nach dem standardmäßigen Testverfahren (ohne Berücksichtigung der konkreten x-Werte) im Unschärfebereich (α=0,05; n=19; k=2):

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

Führt man den DW-Test jedoch z.B. mit dem Programm R unter Berücksichtigung der exakten x-Werte durch, ergibt sich bei einem p-Wert von 0,086 keine Signifikanz auf dem 5%-Niveau. Die Nullhypothese fehlender Autokorrelation 1. Ordnung der Störvariablen lässt sich danach nicht ablehnen.

Mit einem Bestimmtheitsmaß $R^2_{GLS^*}$ von 0,997 im transformierten Regressions-Modell der Konsumfunktion ist die Güte der Anpassung etwas besser als bei Cochran-Orcutt-Methode.

<u>Exkurs</u>

GLS-Schätzer (6.27a) des Prais-Winston-Verfahrens unter Verwendung der Transformationsmatrix **T:**

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

mit $$\hat{\boldsymbol{\Omega}}^{-1} = \mathbf{T}'\mathbf{T}$$

- Varianz-Kovarianz-Matrix der Störvariablen u bei Autokorrelation 1. Ordnung

(6.20)

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

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

Geschätzte Inverse von $\mathbf{\Omega}$:

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

Zerlegung der Inversen $\mathbf{\Omega}^{-1}$ in der Form $\mathbf{\Omega}^{-1} = \mathbf{T}'\mathbf{T}$ impliziert

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

□

6.5 Heteroskedastizitäts- und Autokorrelations-konsistente (HAC) Varianz-Kovarianz-Matrix

Als Alternative zur Anwendung der verallgemeinerten Methode der kleinsten Quadrate (GLS)-Methode) z.B. in Form der Cochran-Orcutt- oder Prais-Winston-Methode bietet sich bei vorhandener Autokorrelation aufgrund der bestehenden Erwartungstreue der OLS-geschätzten Regressionskoeffizienten eine Korrektur des Schätzers der Varianz-Kovarianz-Matrix für $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ an, die durch

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

mit der in (6.1) definierten Varianz-Kovarianz-Matrix der Störgrößen. $\mathrm{Cov}(\mathbf{u}) = \sigma^2 \cdot \boldsymbol{\Omega}$, gegeben ist. Autokorrelations-konsistente Schätzer von $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ kontrollieren jedoch typischerweise gleichzeitig für potenziell vorhandene Heteroskedastizität, so dass ihnen die allgemeine Form der Varianz-Kovarianz-Matrix der Störgrößen.

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

zugrunde liegt.

Unter Verwendung der Definition

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

in der der Vektor $\mathbf{x}_t$ die Werte der exogenen Variablen in der Periode t wiedergibt (= t-te Zeile der Beobachtungsmatrix **X**), lässt sich die Varianz-Kovarianz-Matrix $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ in der Form

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

darstellen.

Bei gegebenen Größen n und **X** erfolgt eine Schätzung von $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ über einen Schätzer für $\mathbf{S}_x$. Newey und West (1987) haben hierfür den **Heteroskedastizitäts- und Autokorrelations-konsistente (HAC) Schätzer**

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

mit den Gewichten

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

entwickelt.

In (3.31) gibt p den maximalen Lag an, der für die Kontrolle von Autokorrelation zu berücksichtigen ist. Ein gängiger Wert für p ist nach Greene (2003) durch die Wahl von $p \approx n^{1/4}$ gegeben.

Die Konsistenz des HAC-Schätzers der Varianz-Kovarianz-Matrix $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ mit (6.30) ist eine asymptotische Eigenschaft. Eine Anpassung für kleine Stichproben ist unter Verwendung des Faktors n/(n-k) analog zum HC1-Schätzers bei Heteroskedastizität vorgeschlagen worden.

Bei ausschließlicher Berücksichtigung von Autokorrelation 1. Ordnung (p=1) vereinfacht sich der HAC-Schätzer (6.31) von Newey und West zu

$$\hat{\mathbf{S}}_x = \left(\sum_{t=1}^{n} \hat{u}_t^2 \mathbf{x}_t \mathbf{x}_t'\right)/n + \left(0{,}5 \cdot \sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1} (\mathbf{x}_t \mathbf{x}_{t-1}' + \mathbf{x}_{t-1} \mathbf{x}_t')\right)/n$$

Im Spezialfall p=0 wird nicht mehr für Autokorrelation kontrolliert, so dass (6.30) in den Heteroskedastizitäts-konsistenten (HC) Schätzer von White übergeht:

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

Newey/West HAC-Schätzer (6.30) für $\mathrm{Var}(\hat{\beta}_2)$ im Fall der einfachen Regression:

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

Beispiel:

Beide Tests auf Autokorrelation, der Durbin-Watson-Test und der Breusch-Godfrey-Test, führen bei der OLS-geschätzten Keynesschen Konsumfunktion im Zeitraum 1994-2012 für Deutschland zu einer eindeutigen Ablehnung der Nullhypothese fehlender Autokorrelation der Störgröße. Die konventionell durchgeführten Signifikanztests verlieren damit ihre Gültigkeit. Eine Möglichkeit der Abhilfe bietet die Anwendung der verallgemeinerten Kleinst-Quadrate-Methode (GLS-Methode) in Form des Cochran-Orcutt- oder Prais-Winston-Verfahrens. Beide Transformationsverfahren zielen darauf ab, Autokorrelation 1. Ordnung bei der Schätzung des Einflusses des verfügbaren Einkommens auf den privaten Konsum auszuschalten.

Als Alternative hierzu bietet es sich an, bei Autokorrelation wie im Falle der Heteroskedastizität die Standardfehler der OLS-geschätzten Regressionskoeffizienten zu korrigieren. Mit dem Heteroskedastizitäts- und Autokorrelations-konsistenten Newey-West-Schätzer (6.30) für S_x ergibt sich unter Verwendung der automatischen Selektionsprozedur nach Newey und West der HAC-Schätzer für $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$,

$$\hat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}})_{NW} = \begin{array}{cc} \text{Abs. Glied} & \text{Verf. Eink.} \\ \begin{bmatrix} 439{,}447514 & -0{,}295162 \\ -0{,}295162 & 0{,}000201 \end{bmatrix} & \begin{array}{l} \text{Abs. Glied} \\ \text{Verf. Eink.} \end{array} \end{array}$$

der hier mit der Funktion NeweyWest des R-Pakets sandwich berechnet worden ist:

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

Unter Berücksichtigung des Adjustierungsfaktors n/(n-k) für endliche Stichproben lautet der HAC-Schätzer nach Newey und West

$$\hat{Cov}(\hat{\boldsymbol{\beta}})_{NWa} = \begin{array}{c} \begin{array}{cc} \text{Abs. Glied} & \text{Verf. Eink.} \end{array} \\ \begin{bmatrix} 491{,}147221 & -0{,}329887 \\ -0{,}329887 & 0.000224 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Abs. Glied} \\ \text{Verf. Eink.} \end{array}$$

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

Die Selektionsprozedur nach Newey und West zeigt eine Relevanz von Autokorrelation 1. und 2. Ordnung in den Residuen der Konsumfunktion an, was bei der Ermittlung der beiden HAC-Schätzer berücksichtigt worden ist.

Die Ergebnisse der HAC-Schätzung,

HAC-Standardfehler (mit Adjustierung): $\sqrt{\hat{Var}(\hat{\beta}_1)_{NWa}} = 22{,}16184$ und $\sqrt{\hat{Var}(\hat{\beta}_2)_{NWa}} = 0{,}0149808$

lassen darauf schließen, dass die konventionellen Standardfehler der OLS-geschätzten Regressionskoeffizienten,

Konventionelle Standardfehler: $\sqrt{\hat{Var}(\hat{\beta}_1)} = 17{,}39254$ und $\sqrt{\hat{Var}(\hat{\beta}_2)} = 0{,}012385$

zu einer deutlichen Unterschätzung der entsprechenden Parameter der Grundgesamtheit führen. Eine solche Unterschätzung der „wahren“ Standardfehler ist bei der konventionellen Schätzung stets bei ökonomischen Zeitreihen zu erwarten, die ein durch langsame trendmäßige Veränderungen im Zeitverlauf gekennzeichnet sind.

Kompakt lassen sich die Signifikanztests der OLS-geschätzten Regressionskoeffizienten unter Verwendung des Newton/West-HAC-Schätzers der Standardfehler z.B. mit der Funktion coeftest des R-Pakets lmtest abrufen. Für die Keynessche Konsumfunktion im Zeitraum 1994-2012 erhält man :

```
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-Befehl: > coeftestKons.NWt = coeftest(Konsumfkt9412.lm, df=n-k, vcov=NeweyWest(Konsumfkt9412.lm, prewhite=FALSE, adjust= TRUE))]

Bei Verwendung des HAC-Standardfehlers nach Newey und West ist das absolute solute Glied der (langfristigen) Konsumfunktion - wie theoretisch zu erwarten – nicht mehr signifikant.