Kapitel 12:

# Serielle Korrelation in Zeitreihen

Moodle

Lehrbuch

# Das klären wir in diesem Kapitel:

Eigenschaften von OLS mit seriell korrelierten Fehlern

Autoregressive Prozesse erster Ordnung

FGLS bei Autokorrelation

Serielle Korrelation: Robuste Fehler

# Eigenschaften von OLS mit seriell korrelierten Fehlern

# Eigenschaften von OLS mit seriell korrelierten Fehlern

- OLS ist unverzerrt und konsistent.
- Das Bestimmtheitsmaß $R^2$ ist valide.
- OLS Standardfehler und darauf basierende Tests sind nicht valide.
- OLS ist nicht effizient.

# Autoregressive Prozesse erster Ordnung

# AR(1) serielle Korrelation

AR(1): Autoregressiver Prozess 1. Ordnung

$$u_t = \rho u_{t-1} + \epsilon_t \ , t = 1, \ldots, n$$

wobei:

- $u_0$ ein passender Startwert
- $\rho \in \mathbb{R}$: unbekannter und fester Parameter
- $\epsilon_t$ iid Zufallsvariable mit $E[\epsilon_t|X] = 0$ und $Var(\epsilon_t|X) = \sigma_\epsilon^2$

$$u_t = \rho u_{t-1} + \varepsilon_t \qquad t = 1, \dots, n$$

$$E[u_t] = E[\rho u_{t-1} + \varepsilon_t]$$

$$= E[\rho u_{t-1}] + \underbrace{E[\varepsilon_t]}_{=0}$$

$$= \rho E[u_{t-1}]$$

$$E[u_1] = \rho E[u_0] = \rho u_0$$

$$E[u_2] = \rho E[u_1] = \rho \cdot \rho u_0 = \rho^2 \cdot u_0$$

$$\vdots$$

$$E[u_t] = \rho^{t+1} \cdot u_0$$

$$U_t = \varrho U_{t-1} + \varepsilon_t \qquad t = 1, \ldots, n$$

$$Var(U_t) = Var(\varrho U_{t-1} + \varepsilon_t)$$

$$= Var(\varrho U_{t-1}) + 2\underbrace{Cov(\varrho U_{t-1}, \varepsilon_t)}_{=0} + \underbrace{Var(\varepsilon_t)}_{\sigma_\varepsilon^2}$$

$$(\ )^2$$

$$= \varrho^2 Var(U_{t-1}) + \sigma_\varepsilon^2$$

$$E[v_t] = \rho E[v_{t-1}]$$

$$Var(v_t) = \rho^2 Var(v_{t-1}) + \sigma_\varepsilon^2$$

$$U_t = \rho U_{t-1} + \varepsilon_t$$

$$Cov(U_t, U_{t-1}) = Cov(\rho U_{t-1} + \varepsilon_t, U_{t-1})$$

$$= Cov(\rho U_{t-1}, U_{t-1}) + \underbrace{Cov(\varepsilon_t, U_{t-1})}_{=0}$$

$$= \rho\, Cov(U_{t-1}, U_{t-1})$$

$$= \rho\, Var(U_{t-1})$$

$$U_t = \rho U_{t-1} + \varepsilon_t$$

$$= \rho(\rho U_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$= \rho^2 U_{t-2} + \rho\, \varepsilon_{t-1} + \varepsilon_t$$

$$Cov(U_t, U_{t-2}) = Cov(\rho^2 U_{t-2} + \rho\, \varepsilon_{t-1} + \varepsilon_t, U_{t-2})$$

$$= Cov(\rho^2 U_{t-2}, U_{t-2}) + \underbrace{Cov(\rho\, \varepsilon_{t-1}, U_{t-2})}_{=0} + \underbrace{Cov(\varepsilon_t, U_{t-2})}_{=0}$$

$$= \rho^2 Cov(U_{t-2}, U_{t-2})$$

$$= \rho^2 Var(U_{t-2})$$

$$Cor(U_t, U_{t-1}) = \rho \, Var(U_{t-1})$$

$$Cor(U_t, U_{t-2}) = \rho^2 \, Var(U_{t-2})$$

$$\vdots$$

$$Cor(U_t, U_{t-s}) = \rho^s \, Var(U_{t-s})$$

$$E[U_t] = \rho \, E[U_{t-1}]$$

$$Var(U_t) = \rho^2 Var(U_{t-1}) + \sigma_\varepsilon^2$$

# Die Kovarianz zweier Störterme bei AR(1)

Es gilt

$$Cov(u_t, u_{t-1}) = Cov(\rho u_{t-1} + \epsilon_t, u_{t-1}) = \rho Var(u_{t-1}) + \underbrace{Cov(\epsilon_t, u_{t-1})}_{=0}$$

und entsprechend:

$$Cov(u_t, u_{t-s}) = \rho^s Var(u_{t-s})$$

# Stationäre autoregressive Prozesse

Ein AR(1)-Prozess

$$u_t = \rho u_{t-1} + \epsilon_t$$

ist **stationär**, falls der Erwartungswert $E[u_t]$ und die Varianz $Var(u_t)$ für alle Elemente von $\{u_t\}$ identisch sind:

- $E[u_t] = \mu \ \forall \ t$
- $Var(u_t) = \sigma^2 \ \forall \ t$

$$\underbrace{E[U_t]}_{=\mu} = \rho \underbrace{E[U_{t-1}]}_{=\mu}$$

Stationarität: $\mu = \rho \cdot \mu$

$\rho = 1$

oder

$\mu = 0$

$$\underbrace{Var(U_t)}_{=\sigma^2} = \rho^2 \underbrace{Var(U_{t-1})}_{=\sigma^2} + \sigma_\varepsilon^2$$

Stationarität

$$\sigma^2 = \rho^2 \sigma^2 + \sigma_\varepsilon^2$$

$$\Leftrightarrow \sigma^2(1-\rho^2) = \sigma_\varepsilon^2 > 0 \Rightarrow \rho^2 \neq 1$$

$$\Rightarrow \mu = 0$$

falls $\rho^2 \neq 1$
$\Leftrightarrow$ $$\sigma^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$\mu = 0 \quad , \quad \varrho \neq 1$

$$\Rightarrow \quad E[U_t] = 0 \qquad \forall\ t = 1, \dots, n$$

$$Var(U_t) = \underbrace{\left(\frac{\sigma_\varepsilon^2}{1-\varrho^2}\right)}_{= \sigma^2} \qquad \forall\ t = 1, \dots, n$$

falls $|\varrho| > 1$

$\Rightarrow \sigma^2 < 0$ ↯

$$Cor(U_t, U_{t-s}) = \varrho^s \cdot \sigma^2$$

$\Rightarrow |\varrho| < 1$

Korrelationskoeffizient von $U_t$ & $U_{t-s}$

$$\frac{Cor(U_t, U_{t-s})}{\sqrt{Var(U_t)}\sqrt{Var(U_{t-s})}} = \frac{\rho^s \overbrace{Var(U_{t-s})}^{\sigma^2}}{\sqrt{\sigma^2}\sqrt{\sigma^2}}$$

$$= \rho^s$$

$\Rightarrow \rho$ ist korr.koeff. von $U_t$ & $U_{t-1}$

# Stationäre autoregressive Prozesse: $|\rho| < 1$

Für $u_t = \rho u_{t-1} + \epsilon_t$ mit $\rho \neq 0$ muss also gelten:

$$\underbrace{E[u_t]}_{\mu} = E[\rho u_{t-1} + \epsilon_t] = \rho \underbrace{E[u_{t-1}]}_{\mu} + \underbrace{E[\epsilon_t]}_{0} ,$$

also $\mu = 0$ oder $\rho = 1$.

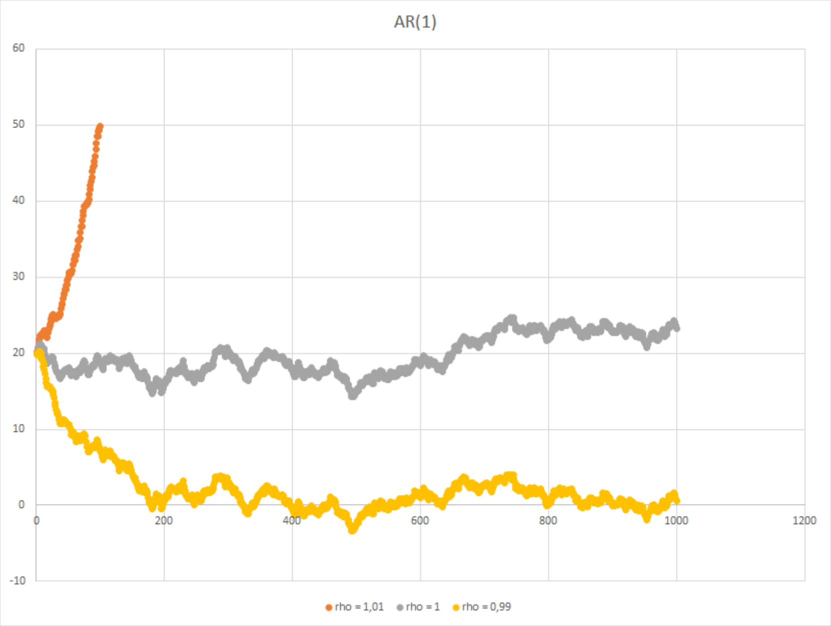
Ebenso:

$$\underbrace{Var(u_t)}_{\sigma^2} = Var(\rho u_{t-1} + \epsilon_t) = \rho^2 \underbrace{Var(u_{t-1})}_{\sigma^2} + 2\rho \underbrace{Cov(u_{t-1}, \epsilon_t)}_{0} + \underbrace{Var(\epsilon_t)}_{\sigma_\epsilon^2}$$

$$\Rightarrow \sigma^2 = \rho^2 \sigma^2 + \sigma_\epsilon^2 \Leftrightarrow \sigma^2 = \frac{\sigma_\epsilon^2}{1 - \rho^2}$$

Damit die Varianz wohldefiniert und positiv ist, muss zusätzlich $|\rho| < 1$ gelten.

# Drei AR(1) Prozesse



# Varianz-Kovarianz-Matrix von stationären AR(1)

$|\rho| < 1$ $\quad u_t = \rho u_{t-1} + \varepsilon_t$

Mit $Var(u_t) = \frac{\overbrace{\sigma_\epsilon^2}^{Var(\varepsilon)}}{1-\rho^2} = \sigma^2$ gilt:

$$\Sigma(\mathbf{u}) = \begin{pmatrix} \overbrace{Var(u_1)}^{=\sigma^2} & \overbrace{Cov(u_1, u_2)}^{=\rho \cdot \sigma^2} & \dots & \overbrace{Cov(u_1, u_n)}^{\rho^{n-1} \cdot \sigma^2} \\ Cov(u_2, u_1) & Var(u_2) & \dots & Cov(u_2, u_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(u_n, u_1) & Cov(u_n, u_2) & \dots & Var(u_n) \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} 1 & \rho & \dots & \rho^{n-1} \\ \rho & 1 & \dots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \dots & 1 \end{pmatrix}$$

$\Omega$ ist in diesem Fall also nur durch einen Parameter ($\rho$) bestimmt!

# Geeignete Datentransformation: „Quasidifferenzierung“

Betrachte

$$\tilde{u}_t = u_t - \rho u_{t-1} = \underbrace{\rho u_{t-1} + \epsilon_t}_{u_t} - \rho u_{t-1} = \epsilon_t, \;\; t = 1, \ldots, n$$

Der Transformierte Störterm $\tilde{u}_t = \epsilon_t$ ist iid, erfüllt also $\mathbf{\Omega} = \mathbf{I}$!

Definiere mit

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\rho & 1 \end{pmatrix}$$

die Transformierten Daten:

$$\tilde{\mathbf{y}} = \mathbf{P}\mathbf{y} \text{ und } \tilde{\mathbf{X}} = \mathbf{P}\mathbf{X}$$

Falls $P$ Wurzel von $\Omega^{-1}$:

$$P \cdot P' = \Omega^{-1}$$

$$P'\Omega \cdot P = I$$

$$P'\Omega = \underbrace{\begin{pmatrix} 1 & -\varrho & 0 \\ 0 & 1 & -\varrho \\ 0 & 0 & 1 \end{pmatrix}}_{P'} \begin{pmatrix} 1 & \varrho & \varrho^2 \\ \varrho & 1 & \varrho \\ \varrho^2 & \varrho & 1 \end{pmatrix} = \begin{pmatrix} 1-\varrho^2 & \varrho-\varrho & \varrho^2-\varrho^2 \\ \varrho-\varrho^3 & 1-\varrho^2 & \varrho-\varrho \\ \varrho^2 & \varrho & 1 \end{pmatrix}$$

$$P'\Omega P = \begin{pmatrix} 1-\varrho^2 & 0 & 0 \\ \varrho(1-\varrho^2) & 1-\varrho^2 & 0 \\ \varrho^2 & \varrho & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\varrho & 1 & 0 \\ 0 & -\varrho & 1 \end{pmatrix} = \begin{pmatrix} 1-\varrho^2 & 0 & 0 \\ \varrho-\varrho^3-\varrho+\varrho^3 & 1-\varrho^2 & 0 \\ \varrho^2-\varrho^2 & \varrho-\varrho & 1 \end{pmatrix}$$

$$P'\Omega P = \begin{pmatrix} 1-\rho^2 & 0 & 0 \\ 0 & 1-\rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\sigma^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$\sigma^2 \cdot P'\Omega P = \sigma_\varepsilon^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1-\rho^2} \end{pmatrix}$$

# Anmerkungen zu $\mathbf{P}$

Offensichtlich ist **P** nicht symmetrisch.

Die Matrix **P** ist *eine* Wurzel von $\mathbf{\Omega}^{-1}$, aber nicht die eindeutige pds-Wurzel.

Dennoch gilt:

$$\mathbf{P\Omega P}' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1-\rho^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1-\rho^2 \end{pmatrix}$$

Demnach ist mit

$$\frac{\sigma_\epsilon^2}{1-\rho^2}\mathbf{P\Omega P}' = \sigma_\epsilon^2 \begin{pmatrix} \frac{1}{1-\rho^2} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Annahme **MLR 5** bis auf die 1. Zeile & Spalte erfüllt.

# FGLS bei Autokorrelation

# FGLS bei Autokorrelation

$u_t = \rho u_{t-1} + \varepsilon_t$

In der Praxis ist $\rho$ nicht bekannt. Indem wir $\rho$ durch eine Schätzung $\widehat{\rho}$ ersetzen, erhalten wir FGLS.

Zur Schätzung von $\widehat{\rho}$ können wir die OLS-Residuen $\widehat{u}_t$ des ursprünglichen Modells verwenden. Die OLS-Hilfsregression

$\bar{\hat{u}} = 0$

$$\widehat{u}_t = \rho \widehat{u}_{t-1} + \epsilon_t$$

liefert den Schätzwert $\widehat{\rho}$. $= \dfrac{\sum_{t=2}^{n} (\hat{u}_t - \bar{\hat{u}})\, \hat{u}_{t-1}}{\sum_{t=2}^{n} (\hat{u}_{t-1} - \bar{\hat{u}})^2}$

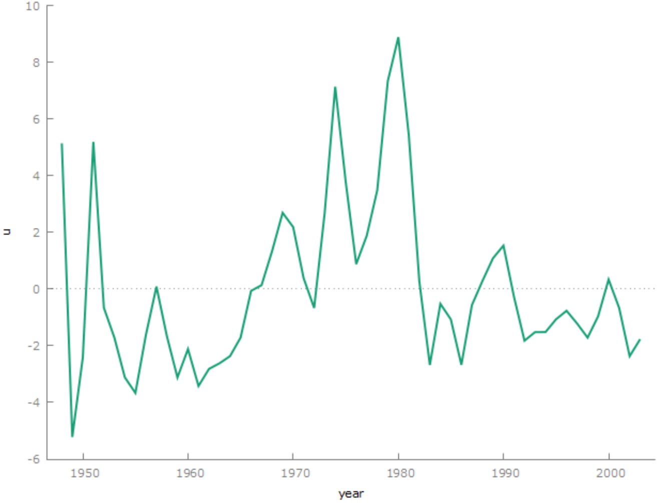
Mit diesem Schätzwert werden die Daten transformiert und die darauf angewandte OLS-Schätzung liefert den FGLS - Schätzer.

Durch die Quasidifferenzierung geht eine Beobachtung verloren. Diese spezielle FGLS - Variante heißt nach ihren Erfindern **Cochrane - Orcutt - Methode**.

# Beispiel: Inflation und Arbeitslosigkeit (Philipps-Kurve)

Benutze die Daten `philipps.xls`.

Regressiere *inf* auf *unem* und erhalte folgende Residuen:



# Beispiel: Inflation und Arbeitslosigkeit (Philipps-Kurve)

$\hat{u}_t = \rho \hat{u}_{t-1} + v_t$

Regressiere die Residuen auf ihren ersten Lag ($\hat{u}_t$ auf $\hat{u}_{t-1}$):

$$\rightarrow \widehat{\hat{u}_t} = \underset{(0{,}107465)}{\overbrace{0,572055}^{\hat{\rho}}}\hat{u}_{t-1}$$

Der Schätzer für $\rho$, also $\hat{\rho} = 0,572$, ist signifikant von null verschieden ($p$-Wert: $2 \cdot 10^{-6}$)

Dies ist ein starker Hinweis auf Autokorrelation der Störterme.

Der $t$-Test ist aber nur asympotisch valide (bei großen Stichproben).

# $t$-Test auf Autokorrelation

1. OLS-Regression des Modells $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$
2. $\rightarrow$ Residuen $\hat{\mathbf{u}} = \mathbf{y} - \underbrace{\mathbf{X}\hat{\boldsymbol{\beta}}}_{} = \hat{y}$
3. OLS-Regression des Modells $\hat{u}_t = \delta_0 + \delta_1 \hat{u}_{t-1} + \epsilon_t$
4. $\rightarrow \hat{\delta}_1 = \frac{\sum_{t=2}^{n} u_t u_{t-1}}{\sum_{t=2}^{n} \hat{u}_{t-1}^2}$, $se(\hat{\delta}_1)$ $= \hat{\sigma} \cdot \sqrt{\frac{1}{\sum_{t=2} \hat{u}_{t-1}^2}}$
5. $H_0 : \delta_1 = 0, \; H_1 : \delta_1 \neq 0$
6. $\rightarrow t = \frac{\hat{\delta}_1}{se(\hat{\delta}_1)}$
7. Lehne $H_0$ ab, falls $|t| > c_{n-K-1,1-\frac{\alpha}{2}}$.

Problem: Der OLS-Standardfehler $se(\hat{\delta}_1)$ ist unter Verletzung von MLR 5 falsch, der $t$-Test ist daher nicht valide.

„Lösung“: Der $t$-Test ist wenigstens asympotisch valide.

# Der Durbin-Watson Test

Die Durbin-Watson (DW) Statistik ist definiert durch:

$a^2 - 2ab + b^2$

$$DW = \frac{\sum_{t=2}^{n} \left(\hat{u}_t - \hat{u}_{t-1}\right)^2}{\underbrace{\sum_{t=1}^{n} \hat{u}_t^2}_{SSR}} \approx 2 - 2\,\frac{\frac{1}{n}\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1}}{\frac{1}{n}\sum_{t=1}^{n} \hat{u}_t^2}$$

Mit einigen Umformungen können wir zeigen:

$$DW \approx 2(1 - \hat{\rho})$$

$\hat{\rho}$: empirische Korrelationskoeffizient

Ist $\hat{\rho} = 1$, so ist $DW \approx 0$.

Ist $\hat{\rho} = 0$, so ist $DW \approx 2$. ← empirisch unkorreliert

Ist $\hat{\rho} = -1$, so ist $DW \approx 4$.

Demnach gilt immer:

$$0 < DW < 4$$

# Durbin-Watson Test auf positive Autokorrelation

Sei $u_t$ ein AR(1)-Prozess.

Nullhypothese:

$$H_0 : \rho = 0$$

Alternativhypothese:

$$H_1 : \rho > 0$$

Für die DW-Statistik gibt es zwei kritische Werte: $d_U$ und $d_L$.

Diese beiden Werte hängen vom Signifikanz-Niveau $\alpha$ ab.

$$\begin{aligned}
DW > d_U & \quad : H_0 \text{ kann nicht verworfen werden} \\
DW \in [d_L, d_U] & \quad : \text{keine Aussage möglich} \\
DW < d_L & \quad : H_0 \text{ muss zugunsten } H_1 \text{ verworfen werden}
\end{aligned}$$

# Breusch-Godfrey Test

1. Schätze $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{u}$ mit OLS und speichere die Residuen $\hat{\boldsymbol{u}}$.
2. Regressiere $\hat{u}_t$ auf $\hat{u}_{t-1}$ **und** $X$ und berechne das Bestimmtheitsmaß $R^2$ dieser Regression
3. Teststatistik: $BG = (n-1)R^2 \sim_{asy} \chi^2_1$

(In gretl:
Tests $\rightarrow$ Autokorrelation. Dort wird $BG = nR^2$ berechnet.)

Der Breusch-Godfrey ist im Gegensatz zum einfachen $t$-Test auch dann valide, falls die Regressoren $X$ nicht strikt exogen sind.

# Serielle Korrelation: Robuste Fehler

# Serielle Korrelation: Robuste Fehler

- Im Fall seriell korrelierter Störterme überschätzen die OLS-Standardfehler statistische Signifikanz, da weniger unabhängige Variation besteht.
- Serielle Korrelation-robuste Standardfehler können mit den OLS-Residuen berechnet werden.
- Das ist nützlich, da FGLS strikt Exogene Regressoren benötigt und eine sehr spezifische Form der Autokorrelation voraussetzt (AR(1) oder allgemeiner AR(q)).

$$\widetilde{se}(\hat{\beta}_j) = \left(\frac{se(\hat{\beta}_j)}{\hat{\sigma}}\right)^2 \sqrt{\hat{v}}$$

Die OLS-Standardfehler $se(\hat{\beta}_j)$ werden normiert und dann mit einem Faktor $\sqrt{\hat{v}}$ „aufgeblasen".

# Serielle Korrelation: Robuste Fehler

Der serielle Korrelation-robuste Standardfehler heißt auch **Newey-West Standardfehler**:

$$\widetilde{se}(\hat{\beta}_j) = \left(\frac{se(\hat{\beta}_j)}{\hat{\sigma}}\right)^2 \sqrt{\hat{v}}$$

Der Faktor $\hat{v}$ berechnet sich wie folgt, wobei $g \in \mathbb{N}$:

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + 2 \sum_{h=1}^{g} \left(1 - \frac{h+1}{g}\right) \sum_{t=h+1}^{n} \hat{a}_t \hat{a}_{t-1}$$

wobei

- $\hat{a}_t = \hat{r}_t \hat{u}_t$ ist und $\hat{r}_t$ das Residuum der Regression von $x_{tj}$ auf $x_{t1}, \ldots, x_{tk}$
- $g$ die nächste ganz Zahl an $\frac{3}{4} n^{\frac{1}{3}}$

... zum Glück brauchen wir uns diese Formel nicht zu merken!

(gretl berechnet diese robusten Standardfehler per Mausklick.)

# Diskussion: robuste Standardfehler

- Diese Standardfehler sind ebenfalls robust gegenüber Heteroskedastizität und werden demnach **Heteroskedastizitäts und Autokorrelation konsistent (HAC)** genannt.
- Da gewöhnlich Stichproben von Querschnittsdaten größer sind als von Zeitreihen werden Heteroskedastizitäts-robuste Standardfehler (aus Kapitel 8) häufiger benutzt.
- Die Newey-West Standardfehler sind bei starker Autokorrelation und kleinen Stichproben nicht sehr solide.