

## 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

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

wobei:

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

$$\underline{U_t = \rho U_{t-1} + \varepsilon_t}, \quad \rho \neq 0$$

$$\text{Cor}( \downarrow U_t, U_{t-1} )$$

$$= \text{Cor}(\rho U_{t-1} + \varepsilon_t, U_{t-1})$$

$$= \text{Cor}(\rho U_{t-1}, U_{t-1}) + \underbrace{\text{Cor}(\varepsilon_t, U_{t-1})}_{(\varepsilon_t \text{ iid: } = 0)}$$

$$= \rho \underbrace{\text{Cor}(U_{t-1}, U_{t-1})}_{\text{Var}(U_{t-1})}$$

$$= \rho \cdot \text{Var}(U_{t-1}) \neq 0$$

## Die Kovarianz zweier Störterme bei AR(1)

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

$$\text{Cov}(u_t, u_{t-1}) = \rho \cdot \text{Var}(u_{t-1})$$

und entsprechend:

$$\text{Cov}(u_t, u_{t-s}) = \rho^s \cdot \text{Var}(u_{t-1})$$

$\xrightarrow{s \rightarrow \infty} 0 \quad \text{falls } |\rho| < 1$

# 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:

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

# Stationäre autoregressive Prozesse: Varianz

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

$$\begin{aligned}\underbrace{\text{Var}(u_t)}_{\sigma^2} &= \text{Var}(\rho u_{t-1} + \epsilon_t) \\ &= \underbrace{\text{Var}(\rho u_{t-1})}_{\rho^2} + \underbrace{2 \text{Cor}(\rho u_{t-1}, \epsilon_t)}_{=0} + \underbrace{\text{Var}(\epsilon_t)}_{\sigma_\epsilon^2} \\ &= \rho^2 \text{Var}(u_{t-1}) + \sigma_\epsilon^2\end{aligned}$$

Falls  $\{u_t\}$  stationär:  $\text{Var}(u_t) = \text{Var}(u_{t-1}) = \sigma^2$

$$\sigma^2 = \rho^2 \cdot \sigma^2 + \sigma_\epsilon^2 \Leftrightarrow \sigma^2(1 - \rho^2) = \sigma_\epsilon^2$$

Falls  $\rho^2 \neq 1$ :  $\text{Var}(u_t) = \frac{\sigma_\epsilon^2}{1 - \rho^2}$  notwendige Annahme:  $|\rho| < 1$

## Stationäre autoregressive Prozesse: Erwartungswert

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

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

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

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

Falls  $\{u_t\}$  stationär:  $E[u_t] = E[u_{t-1}] = \mu$

$$\mu = \rho \cdot \mu \quad \Leftrightarrow \quad \underline{\mu = 0} \quad \text{oder} \quad \cancel{\rho = 1}$$

$$E[u_t] = 0 \quad \text{für alle } t.$$

# Stationäre autoregressive Prozesse

Ein stationärer AR(1) Prozess

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

mit  $\epsilon_t$  iid,  $E[\epsilon_t] = 0$ ,  $Var(\epsilon_t) = \sigma_\epsilon^2$  erfüllt:

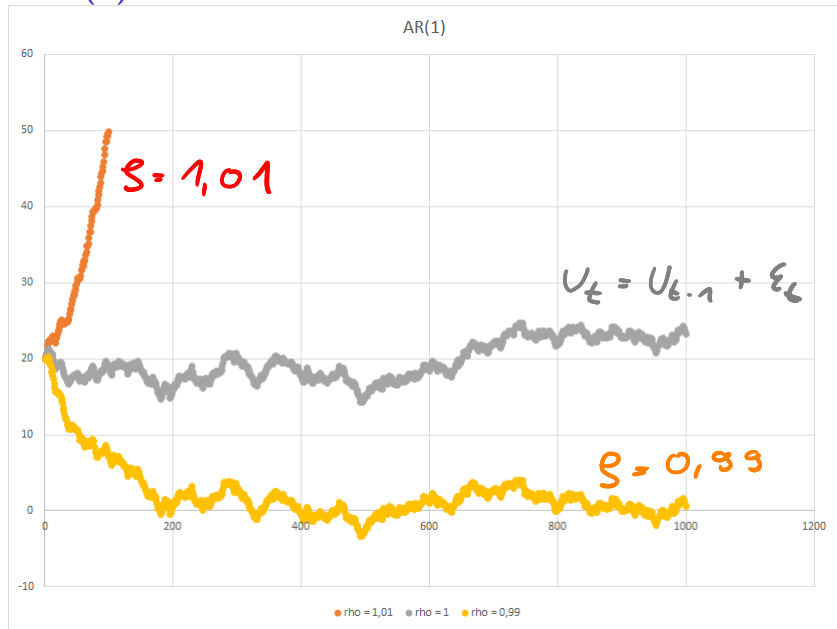
- ▶  $|\rho| < 1$
- ▶  $E[u_t] = 0 \quad \forall t$
- ▶  $Var(u_t) = \frac{\sigma_\epsilon^2}{1-\rho^2} \quad \forall t$

$$\begin{aligned} \text{Cor}(U_t, U_{t-1}) &= \rho \cdot \underbrace{\text{Var}(U_{t-1})}_{\downarrow \sigma^2} \\ &= \rho \cdot \frac{\sigma_\varepsilon^2}{1 - \rho^2} = \rho \cdot \sigma^2 \end{aligned}$$

$$\begin{aligned} \text{Corr}(U_t, U_{t-1}) &= \frac{\text{Cor}(U_t, U_{t-1})}{\sqrt{\text{Var}(U_t)} \cdot \sqrt{\text{Var}(U_{t-1})}} \\ &= \frac{\rho \cdot \sigma^2}{\sqrt{\sigma^2} \cdot \sqrt{\sigma^2}} = \rho \end{aligned}$$

$$\text{Analog: } \text{Corr}(U_t, U_{t-s}) = \dots = \rho^s$$

# Drei AR(1) Prozesse



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

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

$$\begin{aligned} |\rho| < 1 &\Rightarrow E\{u_t\} = 0 \\ \text{Var}(u_t) &= \sigma^2 \\ &= \frac{\sigma_\epsilon^2}{1-\rho^2} \end{aligned}$$

$$\Sigma(\mathbf{u}) = \begin{pmatrix} \text{Var}(u_1) & \text{Cov}(u_1, u_2) & \dots & \text{Cov}(u_1, u_n) \\ \text{Cov}(u_2, u_1) & \text{Var}(u_2) & \dots & \text{Cov}(u_2, u_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(u_n, u_1) & \text{Cov}(u_n, u_2) & \dots & \text{Var}(u_n) \end{pmatrix}$$

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

$$\begin{aligned} \text{Cov}(u_t, u_{t+s}) \\ = \rho^s \cdot \sigma^2 \end{aligned}$$

$\Omega$  ist also nur durch den einen Parameter  $\rho$  bestimmt!

# Geeignete Datentransformation: „Quasidifferenzierung“

Betrachte

$$\tilde{u}_t = u_t - \rho u_{t-1} = \cancel{\rho u_{t-1}} + \epsilon_t - \cancel{\rho u_{t-1}} = \epsilon_t, \quad t = 2, \dots, n$$

Der Transformierte Störterm  $\tilde{u}_t = \epsilon_t$  für  $t > 1$  ist iid, erfüllt also  $\mathbf{\Omega} = \mathbf{I}$  für die Stichprobe ohne die erste Beobachtung!

Definiere mit

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

die Transformierten Daten:

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

## Anmerkungen zu **P**

Offensichtlich ist **P** nicht symmetrisch.

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

Dennoch gilt:

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

Demnach ist mit

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

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

# FGLS bei Autokorrelation

## FGLS bei Autokorrelation

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

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

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

Alternativ

$$\hat{u}_t^2 = \alpha + \beta \hat{u}_{t-1}^2 + \epsilon_t$$

liefert den Schätzwert  $\hat{\rho}$ .

Falls  $u_t$  stationär:  $\hat{\alpha} = 0$

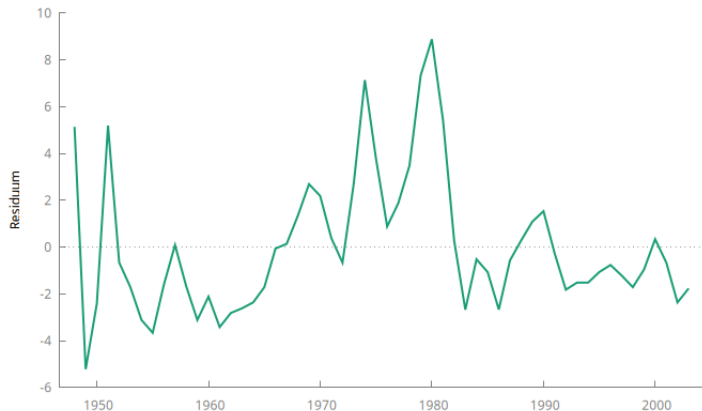
Mit diesem Schätzwert werden die Daten transformiert und die darauf angewandte OLS-Schätzung liefert den FGLS - Schätzer.

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



## Beispiel: Inflation und Arbeitslosigkeit (Philipps-Kurve)

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

$$\hat{u}_t = \underset{(0,31799)}{\overset{\alpha}{-0,111808}} + \underset{(0,10835)}{\overset{\beta}{0,572472}} \hat{u}_{t-1}$$

Der Schätzer für  $\rho$ , also  $\hat{\rho} = 0,572$ , ist signifikant von null verschieden ( $p$ -Wert:  $2,43 \cdot 10^{-6}$ )

Dies ist ein starker Hinweis auf Autokorrelation der Störterme.

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

## $t$ -Test auf Autokorrelation

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

$$\hat{\delta}_1 \rightsquigarrow \hat{\rho}$$

$$\tilde{u}_t = u_t - \hat{\rho} u_{t-1}$$

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

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

# Der Durbin-Watson Test

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

$$DW = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2}$$

Mit einigen Umformungen können wir zeigen:

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

*geschätzter  
Korrelations-  
koeffizient*

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

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

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:

$$H_0 : \rho = 0$$

$$H_0 : \rho = 0$$

$$H_1 : \rho < 0$$

$$\tilde{\rho} = 1 - \hat{\rho}$$

Alternativhypothese:

$$H_1 : \rho > 0$$

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

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

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

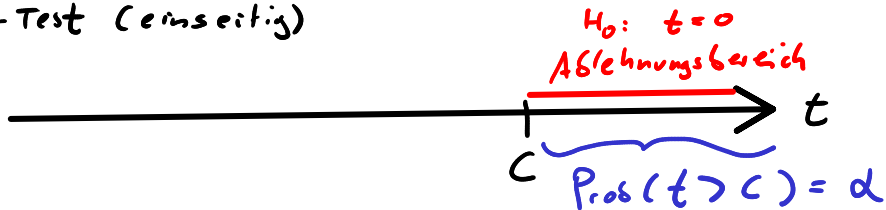
$DW > d_U$  :  $H_0$  kann nicht verworfen werden

$DW \in [d_L, d_U]$  : keine Aussage möglich *inconclusive*

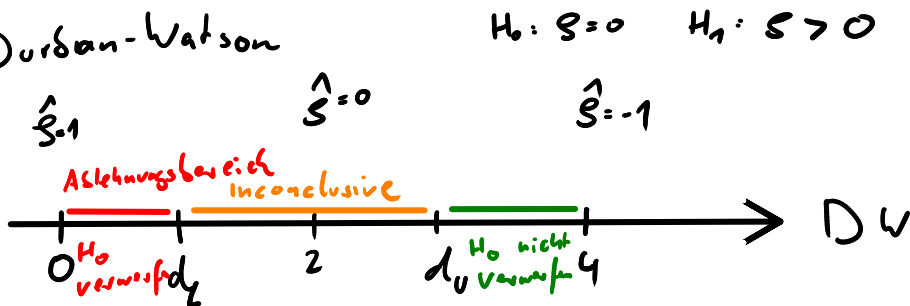
$DW < d_L$  :  $H_0$  muss zugunsten  $H_1$  verworfen werden

Vorteil gegenüber  $t$ -Test: Berücksichtigt die spezielle Verteilung von  $\hat{u}_t$  bei endlichen Stichproben.

t-Test (einseitig)



Durbin-Watson



## Breusch-Godfrey Test

nicht verwechseln mit  
Breusch-Pagan

1. Schätze  $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$  mit OLS und speichere die Residuen  $\hat{\mathbf{u}}$ .  
Hilfsregression
2. Regressiere  $\hat{u}_t$  auf  $\hat{u}_{t-1}$  und  $\mathbf{X}$  und berechne das Bestimmtheitsmaß  $R^2$  dieser Regression
3. Teststatistik:  $BG = (n-1)R^2 \sim_{asy} \chi_1^2$

(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  $\mathbf{X}$  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)).

$$\hat{\sigma}^2 = SSR / (n - k - 1)$$

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

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:**

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

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  $x_{t1}, \dots, x_{tk}$
- ▶  $g$  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.

# Zusammenfassung

- ▶ OLS ist bei Autokorrelation unverzerrt und konsistent, die Standardfehler sind aber falsch und dürfen nicht zur Inferenz benutzt werden.
- ▶ Ein AR(1)-Prozess  $u_t = \rho u_{t-1} + \epsilon_t$  ist stationär, falls  $|\rho| < 1$ .  $\rho$  bezeichnet hier die (Auto-)Korrelation zwischen zwei sukzessiven Störtermen.
- ▶ Durch die Quasidifferenzierung  $\tilde{u}_t = u_t - \hat{\rho}u_{t-1}$  wird die Autokorrelation von AR(1)-Prozessen neutralisiert.
- ▶ Der Durbin-Watson Test oder der Breusch-Godfrey Test können Autokorrelation feststellen.
- ▶ Mit Heteroskedastizitäts- und Autokorrelation konsistenten Standardfehlern (HAC) sind  $t$ - und  $F$ -Tests valide.