universität innsbruck

Heteroskedastizität

Grundlagen der Ökonometrie

herbert.stocker@uibk.ac.at

www.hsto.info/econometrics

Heteroskedastizität

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \ldots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ **Heteroskedastizität**

Heteroskedastizität

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \mathrm{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \ldots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ **Heteroskedastizität**
2. unabhängig verteilt sind, d.h. dass es keine stochastische Abhängigkeit zwischen den einzelnen ε gibt!
   Wenn *nicht* $\Rightarrow$ *Autokorrelation*

Heteroskedastizität

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \ldots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ **Heteroskedastizität**
2. unabhängig verteilt sind, d.h. dass es keine stochastische Abhängigkeit zwischen den einzelnen ε gibt!
   Wenn *nicht* $\Rightarrow$ *Autokorrelation*

Beide Verletzungen verursachen die gleichen Probleme:

- OLS *Koeffizienten* sind erwartungstreu und konsistent, *aber nicht effizient*!

Heteroskedastizität

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \ldots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ **Heteroskedastizität**
2. unabhängig verteilt sind, d.h. dass es keine stochastische Abhängigkeit zwischen den einzelnen ε gibt!
   Wenn *nicht* $\Rightarrow$ *Autokorrelation*

Beide Verletzungen verursachen die gleichen Probleme:

- OLS *Koeffizienten* sind erwartungstreu und konsistent, *aber nicht effizient*!
- OLS *Standardfehler* sind weder erwartungstreu noch effizient noch konsistent!!!

Heteroskedastizität

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \ldots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ **Heteroskedastizität**
2. unabhängig verteilt sind, d.h. dass es keine stochastische Abhängigkeit zwischen den einzelnen ε gibt!
   Wenn *nicht* $\Rightarrow$ *Autokorrelation*

Beide Verletzungen verursachen die gleichen Probleme:

- OLS *Koeffizienten* sind erwartungstreu und konsistent, *aber nicht effizient*!
- OLS *Standardfehler* sind weder erwartungstreu noch effizient noch konsistent!!!
- $\Rightarrow$ Hypothesentests, Konfidenzintervalle etc. ungültig!!!

Heteroskedastizität

Zur Erinnerung:

$$\begin{aligned}
\operatorname{var}(\widehat{\beta}_2) &:= \mathrm{E}[\widehat{\beta}_2 - \mathrm{E}(\widehat{\beta}_2)]^2 \\
&= \mathrm{E}\left(w_1^2\varepsilon_1^2 + w_2^2\varepsilon_2^2 + \cdots + w_n^2\varepsilon_n^2 + \cdots\right. \\
&\qquad \left.\cdots + 2w_1w_2\varepsilon_1\varepsilon_2 + \cdots + 2w_{n-1}w_n\varepsilon_{n-1}\varepsilon_n\right) \\
&= \underbrace{\mathrm{E}\left(\sum_{i=1}^{n} w_i^2\varepsilon_i^2\right)}_{= \sigma^2 \sum_i w_i^2 \text{ wenn homoskedastisch}} + \underbrace{\mathrm{E}\left(\sum_{i=1}^{n}\sum_{\substack{j=2\\ j>i}}^{n} 2w_iw_j\varepsilon_i\varepsilon_j\right)}_{= 0 \text{ wenn keine Autokorrelation}}
\end{aligned}$$

Nur wenn $\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2)$

ist OLS-Formel für Standardfehler $\widehat{\text{se}}(\widehat{\beta}_2) = \sqrt{\dfrac{\hat{\sigma}^2}{\sum_i (x_i - \bar{x})^2}}$ gültig!!!

Heteroskedastizität

in Matrixschreibweise:

$$\operatorname{var}(\hat{\boldsymbol{\beta}}) := \mathrm{E}\left(\left[\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}})\right]\left[\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}})\right]'\right) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\,\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right)\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

mit Varianz-Kovarianzmatrix der Störterme $\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right)$; Dimension $(n \times n)$

$$\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right) = \begin{pmatrix} \operatorname{var}(\varepsilon_1) & \operatorname{cov}(\varepsilon_1\varepsilon_2) & \dots & \operatorname{cov}(\varepsilon_1\varepsilon_n) \\ \operatorname{cov}(\varepsilon_2\varepsilon_1) & \operatorname{var}(\varepsilon_2) & \dots & \operatorname{cov}(\varepsilon_2\varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}(\varepsilon_n\varepsilon_1) & \operatorname{cov}(\varepsilon_n\varepsilon_2) & \dots & \operatorname{var}(\varepsilon_n) \end{pmatrix} := \operatorname{var}(\boldsymbol{\varepsilon})$$

Heteroskedastizität

in Matrixschreibweise:

$$\mathrm{var}(\hat{\boldsymbol{\beta}}) := \mathrm{E}\left(\left[\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}})\right]\left[\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}})\right]'\right) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\,\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right)\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

mit Varianz-Kovarianzmatrix der Störterme $\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right)$; Dimension $(n \times n)$

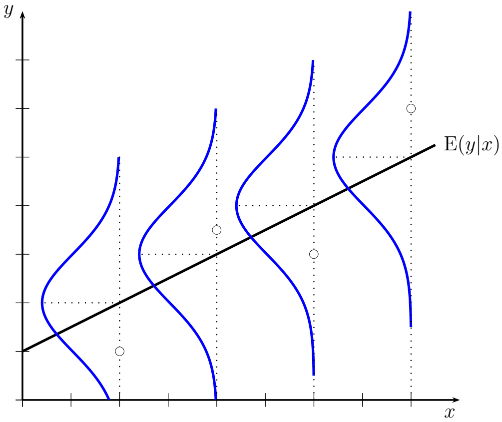
$$\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right) = \begin{pmatrix} \mathrm{var}(\varepsilon_1) & \mathrm{cov}(\varepsilon_1\varepsilon_2) & \dots & \mathrm{cov}(\varepsilon_1\varepsilon_n) \\ \mathrm{cov}(\varepsilon_2\varepsilon_1) & \mathrm{var}(\varepsilon_2) & \dots & \mathrm{cov}(\varepsilon_2\varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(\varepsilon_n\varepsilon_1) & \mathrm{cov}(\varepsilon_n\varepsilon_2) & \dots & \mathrm{var}(\varepsilon_n) \end{pmatrix} := \mathrm{var}(\boldsymbol{\varepsilon})$$

Nur wenn $\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right) = \sigma^2\boldsymbol{I}$ ist

$$\mathrm{var}_{\mathsf{OLS}}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

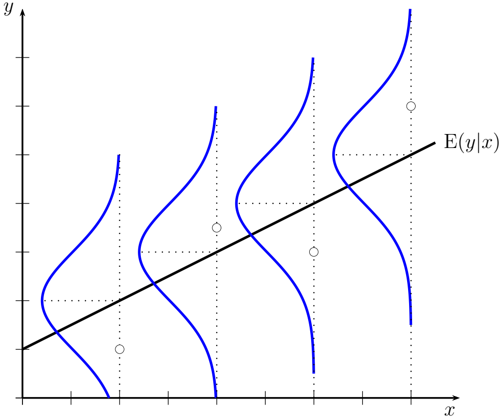
Heteroskedastizität

Homoskedastizität: $\mathrm{E}(\varepsilon_i^2|\boldsymbol{X}) = \sigma^2$

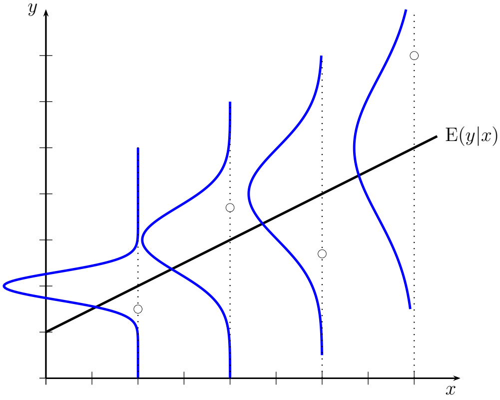


Heteroskedastizität

Homoskedastizität: $\mathrm{E}(\varepsilon_i^2|\boldsymbol{X}) = \sigma^2$

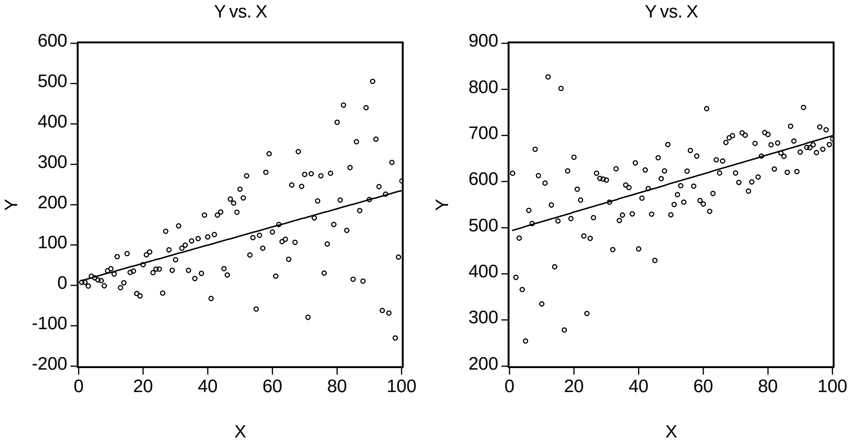


Heteroskedastizität: $\mathrm{E}(\varepsilon_i^2|\boldsymbol{X}) = \sigma_i^2$



Heteroskedastizität

Residuenplot:



Tests auf Heteroskedastizität

Breusch-Pagan-Godfrey Test:

Nullhypothese *keine* Heteroskedastizität: $\mathrm{H}_0 \colon \sigma_i^2 = \sigma^2 \qquad \forall\, i$

Quadrierte Residuen werden für Test verwendet:

1. Schätze die OLS-Regression

$$y_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_{i2} + \cdots + \widehat{\beta}_k x_{ik} + \hat{\varepsilon}_i \quad \rightarrow \quad \hat{\varepsilon}_i^2$$

Tests auf Heteroskedastizität

Breusch-Pagan-Godfrey Test:

Nullhypothese *keine* Heteroskedastizität: $\mathrm{H}_0 \colon \sigma_i^2 = \sigma^2 \qquad \forall\, i$

Quadrierte Residuen werden für Test verwendet:

1. Schätze die OLS-Regression
$$y_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_{i2} + \cdots + \widehat{\beta}_k x_{ik} + \hat{\varepsilon}_i \quad \rightarrow \quad \hat{\varepsilon}_i^2$$
2. Hilfsregression:
$$\hat{\varepsilon}_i^2 = \hat{\gamma}_1 + \hat{\gamma}_2 z_{i2} + \cdots + \hat{\gamma}_l z_{il} + \nu_i \quad \rightarrow \quad R^2_{\hat{\varepsilon}^2}$$
mit z_h: Variablen, die Heteroskedastizität erklären können;
Default: x-Variablen!

Tests auf Heteroskedastizität

Breusch-Pagan-Godfrey Test:

Nullhypothese *keine* Heteroskedastizität: $\mathrm{H}_0 : \sigma_i^2 = \sigma^2 \qquad \forall\, i$

Quadrierte Residuen werden für Test verwendet:

1. Schätze die OLS-Regression

$$y_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_{i2} + \cdots + \widehat{\beta}_k x_{ik} + \hat{\varepsilon}_i \quad \rightarrow \quad \hat{\varepsilon}_i^2$$

2. Hilfsregression:

$$\hat{\varepsilon}_i^2 = \hat{\gamma}_1 + \hat{\gamma}_2 z_{i2} + \cdots + \hat{\gamma}_l z_{il} + \nu_i \quad \rightarrow \quad R^2_{\hat{\varepsilon}^2}$$

   mit z_h: Variablen, die Heteroskedastizität erklären können;
   Default: x-Variablen!

3. *Lagrange-Multiplier* Teststatistik:

$$\text{LM} = n R^2_{\hat{\varepsilon}^2} \sim \chi^2_{l-1}$$

   Alternativ: F-Test der $\mathrm{H}_0 : \gamma_2 = \gamma_3 = \cdots = \gamma_l = 0$

Tests auf Heteroskedastizität

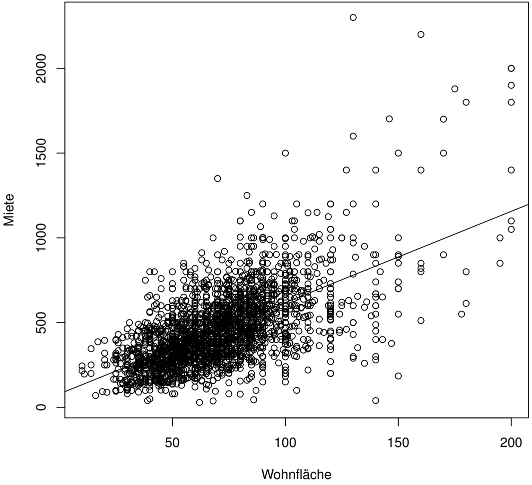
R-Beispiel: Mietausgaben

```
# Mietpreise in Österreich (EU-Silc 2012)
rm(list=ls())
df <- read.csv("http://www.hsto.info/econometrics/dl/mieten2012.csv",
      header = TRUE)
eq <- lm(Miete ~ WFl + WR, data = df)
# Breusch Pagan Test
library(AER)
bptest(eq)
## BP = 294.85, df = 2, p-value < 2.2e-16

# oder sehr ausführlich (nur zur Demonstration)
ehatsq <- (resid(eq))^2
eq2 <- lm(ehatsq ~ WFl + WR, data = df)
BP <- nobs(eq2)*summary(eq2)$r.squared
p_value <- 1-pchisq(BP, length(coefficients(eq2)) - 1)
## BP =  294.8472  p-value =  0
```

Tests auf Heteroskedastizität

R-Beispiel: Mietausgaben, BP = 294.8472 p-value = 0.00



Tests auf Heteroskedastizität

R-Beispiel: Mietausgaben, BP = 294.8472 p-value = 0.00

Interpretation:

Wenn die Störterme (DGP!) homoskedastisch sind (d.h. die H_0 wahr ist) und alle Annahmen erfüllt sind (insbesondere die PRF richtig spezifiziert ist) würden wir bei einer neuerlichen Durchführung des Zufallsexperiments (Stichprobenziehung) nur in einer verschwindend geringen Wahrscheinlichkeit einen ähnlich extremen Wert der Breusch-Pagan Teststatistik erwarten ...

$\Rightarrow$ wir verwerfen die H_0 homoskedastischer Störterme auf einem Signifikanzniveau von 1%.

Tests auf Heteroskedastizität

White-Test: $\mathrm{H}_0: \quad \sigma_i^2 = \sigma^2$ für alle i

Für die Hilfsregression werden zusätzlich die Quadrate und Kreuzprodukte aller erklärenden Variablen berücksichtigt, d.h. für

$$y_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_{i2} + \widehat{\beta}_3 x_{i3} + \widehat{\beta}_4 x_{i4} + \hat{\varepsilon}_i$$

wird die Hilfsregression

$$\begin{aligned}\hat{\varepsilon}_i^2 \quad = \quad & \hat{\gamma}_1 + \hat{\gamma}_2 x_{i2} + \hat{\gamma}_3 x_{i3} + \hat{\gamma}_4 x_{i4} + \\ & \hat{\gamma}_5 x_{i2}^2 + \hat{\gamma}_6 x_{i3}^2 + \hat{\gamma}_7 x_{i4}^2 + \\ & \hat{\gamma}_8 x_{i2} x_{i3} + \hat{\gamma}_9 x_{i2} x_{i4} + \hat{\gamma}_{10} x_{i3} x_{i4}\end{aligned}$$

geschätzt, oder alternativ

$$\hat{\varepsilon}_i^2 = \hat{\delta}_1 + \hat{\delta}_2 \widehat{y_i} + \hat{\delta}_3 \widehat{y}_i^2 \qquad \text{mit } \mathrm{H}_0\colon \delta_2 = \delta_3 = 0 \quad \Rightarrow \text{ F oder LM-Test}$$

Heteroskedastiekonsistente (robuste) Standardfehler

Robuste Standardfehler:

Idee: Schätzung der Varianz der Störterme aus den Residuen (White 1980, Eickert 1963, Huber 1967).
Ermöglicht *konsistente* Schätzung von Standardfehlern (kleine Stichprobeneigenschhaften unbekannt).

Trade-off

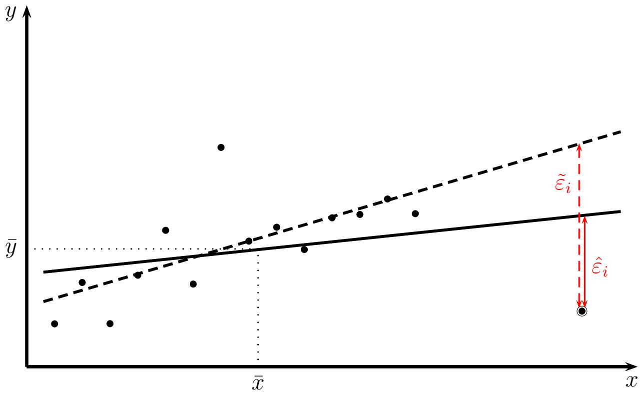
- *bei Homoskedastizität:* OLS Standardfehler sind erwartungstreu, effizient und konsistent!
  Robuste Standardfehler sind konsistent, aber nicht effizient (d.h. ungenau)!
- *bei Heteroskedastizität:* OLS Standardfehler sind weder erwartungstreu noch effizient noch konsistent!!!
  Robuste Standardfehler sind konsistent!

Für große Stichproben werden häufig robuste Standardfehler geschätzt.

Heteroskedastiekonsistente (robuste) Standardfehler

Verschiedene Varianten robuster Standardfehler: HC0, HC1, HC2, HC3, ...

R-Default: **HC3**, berücksichtigt auch *Leverage*



Heteroskedastiekonsistente (robuste) Standardfehler

Beispiel: Mietpreise

```
eq <- lm(Miete ~ WFl + WR, data = df)
# Robuste Standardfehler (AER package)
library(AER)
rob.HC3 <- sqrt(diag(vcovHC(eq, type = "HC3")))

stargazer(eq, eq, se=list(OLS.se, rob.HC3),
          # title="Mietpreise in Österreich",
          no.space=TRUE, align=TRUE, omit.stat=c("LL","ser","f", "rsq"),
          column.labels=c("OLS", "HC3"),
          covariate.labels=c("Interzept", "Wohnfläche", "Wohnräume")
          dep.var.caption="", intercept.bottom=FALSE,
          model.numbers=FALSE,
          type="text"
)
```

Heteroskedastiekonsistente (robuste) Standardfehler

Beispiel: Mietpreise

	Miete	
	OLS	HC3
Interzept	98.651^{***}	98.651^{***}
	(10.852)	(14.251)
Wohnfläche	6.404^{***}	6.404^{***}
	(0.217)	(0.375)
Wohnräume	-31.945^{***}	-31.945^{***}
	(5.252)	(6.489)
Observations	2,441	2,441
Adjusted R^2	0.391	0.391
Note:	*p$<$0.1; **p$<$0.05; ***p$<$0.01	

Heteroskedastiekonsistente (robuste) Standardfehler

- Große Stichproben (z.B. mehrere hundert Beobachtungen): häufig werden nur robuste Standardfehler ausgewiesen;

Heteroskedastiekonsistente (robuste) Standardfehler

- Große Stichproben (z.B. mehrere hundert Beobachtungen): häufig werden nur robuste Standardfehler ausgewiesen;
- Kleine Stichproben: *trade-off*, Test auf Heteroskedastizität, Vergleich robuste vs. OLS Standardfehler (größere verwenden?)

Autokorrelation

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \dots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ *Heteroskedastizität*, v.a. bei Querschnittsdaten
2. unabhängig verteilt sind, d.h. dass es keine stochastische Abhängigkeit zwischen den einzelnen ε gibt!
   Wenn *nicht* $\Rightarrow$ **Autokorrelation**, v.a. bei Zeitreihendaten!

Autokorrelation

Gauss-Markov Annahme A4

$$\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2) \qquad \forall\, i \in \{1, \ldots, n\}$$

impliziert, dass *alle* ε_i

1. identisch verteilt sind, d.h. alle ε_i den gleichen Erwartungswert und die gleiche Varianz haben!
   Wenn *nicht* $\Rightarrow$ *Heteroskedastizität*, v.a. bei Querschnittsdaten
2. unabhängig verteilt sind, d.h. dass es keine stochastische Abhängigkeit zwischen den einzelnen ε gibt!
   Wenn *nicht* $\Rightarrow$ **Autokorrelation**, v.a. bei Zeitreihendaten!

Beide Verletzungen verursachen die gleichen Probleme:

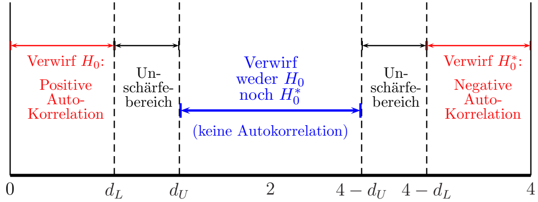
- OLS *Koeffizienten* sind erwartungstreu und konsistent, *aber nicht effizient*!
- OLS *Standardfehler* sind weder erwartungstreu noch effizient noch konsistent!!!
- $\Rightarrow$ Hypothesentests, Konfidenzintervalle etc. ungültig!!!

Autokorrelation

Klassische Teststatistik: **Durbin-Watson Test**

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

Diese Teststatistik kann Werte zwischen 0 und 4 annehmen, wobei Werte nahe bei Null auf positive Autokorrelation und Werte nahe bei 4 auf negative Autokorrelation hindeuten.
Im Idealfall sollten die Werte der DW Statistik möglichst nahe bei 2 liegen.



Autokorrelation

Zeitreihen:

- Zusätzliche Gefahr bei Zeitreihen: **Nicht-Stationarität** (wenn Autokorrelationskoeffizient exakt gleich Eins, $\rightarrow$ *random walk*) führt zu Scheinkorrelation (*‘spurious corralation’*)
  Tests: z.B. Dickey-Fuller Test, KPSS Test, ...

Autokorrelation

Zeitreihen:

- Zusätzliche Gefahr bei Zeitreihen: **Nicht-Stationarität** (wenn Autokorrelationskoeffizient exakt gleich Eins, $\rightarrow$ *random walk*)
  führt zu Scheinkorrelation (*‘spurious corralation’*)
  Tests: z.B. Dickey-Fuller Test, KPSS Test, ...
- Robust: **HAC-Standardfehler**: (*‘heteroscedasticity and autocorrelation consistent standard errors’*), sind konsistent.

Autokorrelation

Zeitreihen:

- Zusätzliche Gefahr bei Zeitreihen: **Nicht-Stationarität** (wenn Autokorrelationskoeffizient exakt gleich Eins, $\rightarrow$ *random walk*)
  führt zu Scheinkorrelation (*'spurious corralation'*)
  Tests: z.B. Dickey-Fuller Test, KPSS Test, ...
- Robust: **HAC-Standardfehler**: (*'heteroscedasticity and autocorrelation consistent standard errors'*), sind konsistent.
- $\rightarrow$ Vertiefungskurs!