5. Lineare Regressionsmodelle mit Zeitreihendaten

5.1 Vorüberlegungen

Besonderheiten von Zeitreihendaten im Vergleich zu Querschnittsdaten:

- Bei Zeitreihendaten werden eine oder mehrere Variablen über mehrere aufeinander folgende Zeitperioden erhoben, so dass im Gegensatz zu Querschnittsdaten eine zeitliche Anordnung vorliegt
- Dabei können zwar zurückliegende Werte zukünftige Werte beeinflussen, jedoch nicht umgekehrt (im Unterschied zum „Star Trek Universum")
- Querschnittsdaten werden häufig als Stichprobe vom Umfang n aus der Grundgesamtheit betrachtet, so dass man bei unterschiedlichen Stichproben in der Regel verschiedene Werte von abhängigen und erklärenden Variablen (z.B. Bildung, Erfahrung, Geschlecht, Lohn) erhält. Aber auch ökonomische Zeitreihendaten können als Zufallsvariablen aufgefasst werden, da sie im Vorfeld nicht vorhersehbar sind.
- Formal werden die Sequenzen der Zufallsvariablen über die Zeit als stochastische Prozesse oder Zeitreihenprozesse aufgefasst
- Wenn Zeitreihendaten gesammelt werden, erhält man eine Realisation des stochastischen Prozesses, wobei nur eine einzige spezifische Realisation untersucht werden kann

- Falls bestimmte Bedingungen in der Vergangenheit unterschiedlich gewesen wären, hätte man aber eine andere Realisation erhalten. Dies ist der Grund, warum Zeitreihendaten als Realisationen von Zufallsvariablen aufgefasst werden können.
- Beispiele: Makroökonomische Daten (z.B. Einkommen, Konsum, Investitionen, Geldangebot, Preisindizes, Inflationsraten), Finanzmarktdaten (z.B. Aktienkurse, Aktienrenditen)
- Im Gegensatz zur Querschnittsanalyse werden bei der Betrachtung von Zeitreihendaten üblicherweise nicht die Indizes i = 1,…, n für die einzelnen Beobachtungen verwendet, sondern t = 1,…, n für den stochastischen Prozess $\{(x_{t1}, x_{t2},\dots, x_{tk}, y_t)\}$, wobei n jetzt die Anzahl der Zeitperioden ist

Damit ergibt sich die entsprechende Formulierung eines multiplen linearen Regressionsmodells:

$$(5.1) \quad y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + u_t \quad \text{für } t = 1,\dots, n$$

Dabei stellt $\{u_t: t = 1,\dots, n\}$ die Sequenz von Störtermen dar und x_{tj} (in Analogie zu linearen Regressionsmodellen mit Querschnittsdaten) den Wert der erklärenden Variablen j = 1,…, k in Zeitperiode t. Im Folgenden beinhalten der k-dimensionale Vektor $x_t = (x_{t1},\dots, x_{tk})$ die erklärenden Variablen in t sowie die (nxk)-dimensionale Matrix X sämtliche erklärende Variablen über alle Perioden, wobei x_t die t-te Zeile von X darstellt.

Beispiel: Erklärung von Mordraten

Mit Hilfe eines linearen Regressionsmodells soll für eine Stadt der Effekt der Mordaufklärungsrate (convrte), der lokalen Arbeitslosenrate (unem) und des Bevölkerungsanteils von Männern zwischen 18 und 25 Jahren (yngmle) auf die Anzahl der Morde pro 10000 BewohnerInnen (mrdrte) in einem Jahr t untersucht werden. Dabei liegen Daten für n = 8 Jahre vor, so dass für X gilt:

t	convrte	unem	yngmle
1	0,46	0,074	0,12
2	0,42	0,071	0,12
3	0,42	0,063	0,11
4	0,47	0,062	0,09
5	0,48	0,060	0,10
6	0,50	0,059	0,11
7	0,55	0,058	0,12
8	0,56	0,059	0,13

5.2 Eigenschaften von OLS-Schätzern

→ Die unbekannten Parameter von multiplen linearen Regressionsmodellen können grundsätzlich ebenfalls mit der OLS-Methode geschätzt werden. Zur Ableitung der Eigenschaften von OLS-Schätzern bei endlichen Stichprobenumfängen sowie der asymptotischen Eigenschaften werden erneut verschiedene Annahmen getroffen, die sich teilweise von den Annahmen bei Querschnittsanalysen unterscheiden.

Annahmen zur Betrachtung des Erwartungswerts von OLS-Schätzern:

- Annahme B1: Linearität in den Parametern
  Der Zeitreihenprozess $\{(x_{t1},\dots, x_{tk}, y_t): t = 1,\dots, n\}$ folgt dem linearen Modell
  $y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + u_t$
- Annahme B2: Keine perfekte Kollinearität
  In der Stichprobe (und daher auch im zugrundeliegenden Zeitreihenprozess) ist keine der erklärenden Variablen konstant und es besteht keine exakte lineare Beziehung zwischen den erklärenden Variablen
- Annahme B3: Bedingter Erwartungswert von u_t ist null
  Für jede Zeitperiode t ist der bedingte Erwartungswert von u_t, gegeben die erklärenden Variablen für alle Perioden t = 1,…, n, null, d.h.

$$(5.2) \quad E(u_t|X) = 0 \quad t = 1,\dots, n$$

Zu Annahme B3 als wesentliche Annahme:

- Sie impliziert, dass der Störterm u_t in einer Zeitperiode t mit jeder erklärenden Variablen x_{tj} (j = 1,…, k) in jeder Periode t = 1,…, n unkorreliert ist. Falls u_t unabhängig von X ist und $E(u_t) = 0$, dann gilt automatisch (5.2).
- Im folgenden Fall sind die x_{tj} lediglich kontemporär exogen:

$$(5.3)\quad E(u_t|x_{t1},\ldots, x_{tk}) = E(u_t|x_t) = 0 \quad \text{für } t = 1,\ldots, n$$

  In diesem Fall sind die u_t und die erklärenden Variablen kontemporär unkorreliert, d.h. $Corr(x_{tj}, u_t) = 0$ für alle j = 1,…, k.
- Annahme B3 erfordert jedoch mehr als kontemporäre Exogenität, da u_t mit x_{sj} selbst für $s \neq t$ unkorreliert sein muss. In diesem Fall spricht man von strikter Exogenität der erklärenden Variablen, wenngleich die kontemporäre Exogenität bereits für einige Eigenschaften ausreichend ist.
- Bei Querschnittsanalysen ist die Annahme der Unkorreliertheit des Störtermes u_i mit den erklärenden Variablen x_{sj} für andere Beobachtungen ($s \neq i$) nicht notwendig, wenn eine Zufallsstichprobe unterstellt wird, da diese automatisch die Unabhängigkeit zwischen u_i und x_{sj} ($s \neq i$) impliziert
- Bei Regressionsanalysen mit Zeitreihendaten ist die Annahme von Zufallsstichproben dagegen fast nie passend, so dass hier mit B3 eine spezielle Annahme getroffen werden muss. Wenn aber der Störterm u_t doch mit einer erklärenden Variablen in einer beliebigen Zeitperiode korreliert ist, wird gegen Annahme B3 verstoßen.

- Annahme B3 ist in Regressionsanalysen mit Zeitreihendaten aber sehr häufig unrealistisch und gilt nur in recht wenigen Fällen. Dennoch soll diese Annahme zunächst betrachtet werden.

Auch unter den drei Annahmen B1 bis B3 sind alle mit der OLS-Methode geschätzten Parameter sowohl unter der Bedingung von X und auch ohne Bedingung erwartungstreu, d.h. der Erwartungswert der Schätzer ist gleich den unbekannten Regressionsparametern:

$$(5.4) \quad E(\hat{\beta}_j) = \beta_j \quad \text{für } j = 0, 1, \ldots, k$$

Annahmen zur Betrachtung der Varianz von OLS-Schätzern:

- Annahmen B1 bis B3 (die bei der Betrachtung des Erwartungswertes von OLS-Schätzern getroffen werden)
- Annahme B4: Homoskedastizität
  Die bedingte Varianz des Fehlerterms u_t ist konstant über alle Zeitperioden $t = 1, \ldots, n$, d.h. es gilt $Var(u_t|X) = Var(u_t) = \sigma^2$. Falls dies nicht zutrifft, liegt wie bei Querschnittsanalysen Heteroskedastizität vor.
- Annahme B5: Keine Autokorrelation
  Unter der Bedingung von X sind die Störterme für beliebige Zeitperioden unkorreliert, d.h. es gilt $Corr(u_t,u_s|X) = 0$ für alle $t \neq s$. Bei dieser Annahme wird zur einfacheren Interpretation oft von der Bedingung von X abstrahiert:

  $$(5.5) \quad Corr(u_t, u_s) = 0 \quad \text{für alle } t \neq s$$

Zu Annahme B5:

- Falls B5 nicht gilt, liegt bei den Störtermen Autokorrelation bzw. serielle Korrelation über die Zeit vor
- Wenn z.B. bei positivem u_{t-1} im Durchschnitt auch u_t in der nächsten Zeitperiode positiv ist, folgt $Corr(u_t,u_{t-1}) > 0$, so dass die Störterme autokorreliert sind. Daraus folgt, dass ein sehr hoher Wert der abhängigen Variablen (z.B. ein Zinssatz) in t-1 im Durchschnitt (bei konstanten Werten der erklärenden Variablen) zu überdurchschnittlich hohen Werten in der folgenden Periode t führt.
- Zu beachten ist, dass Annahme B5 keine Implikationen für intertemporale Korrelationen in den erklärenden Variablen hat, d.h. solche Korrelationen sind verträglich mit B5
- Bei Querschnittsanalysen ist Annahme B5 nicht notwendig, wenn eine Zufallsstichprobe unterstellt wird, da diese automatisch die Unabhängigkeit der Störterme u_i und u_h für zwei beliebige Beobachtungen i und h impliziert
- Somit wird Autokorrelation vornehmlich als Problem von Zeitreihenanalysen betrachtet, so dass die Annahmen B1 bis B5 in diesem Kontext passende Gauss-Markov-Annahmen darstellen (allerdings können diese Annahmen vereinzelt auch bei Querschnittsanalysen von Bedeutung sein, falls die Annahme von Zufallsstichproben nicht adäquat ist, z.B. wenn die untersuchten Beobachtungen einen großen Anteil der Grundgesamtheit ausmachen)

Damit ergibt sich auch unter den Annahmen B1 bis B5 für die bedingte Varianz der mit OLS geschätzten Steigungsparameter:

$$(5.6)\quad \mathrm{Var}(\hat{\beta}_j|X) = \frac{\sigma^2}{(1-R_j^2)\sum_{t=1}^{n}(x_{tj}-\bar{x}_j)^2} = \frac{\sigma^2}{(1-R_j^2)SST_j} \quad \text{für } j = 1,\ldots, k$$

Dabei stellt R_j^2 das Bestimmtheitsmaß bei einer Regression von x_j auf alle anderen erklärenden Variablen (einschließlich einer Konstante) dar.

Des Weiteren gilt:

- Auch unter den Annahmen B1 bis B5 ergibt sich folgender erwartungstreuer Schätzer für die Varianz σ^2 des Fehlerterms u_t:

  $$\hat{\sigma}^2 = \frac{1}{n-k-1}\sum_{t=1}^{n}\hat{u}_t^2 = \frac{SSR}{n-k-1}$$

- Auch unter den Annahmen B1 bis B5 sind die OLS-Schätzer die besten linearen unverzerrten Schätzer unter der Bedingung von X

Somit liegen bei den Annahmen B1 bis B5 in Regressionsanalysen mit Zeitreihendaten dieselben wünschenswerten Eigenschaften bei endlichen Stichproben vor wie bei den Annahmen A1 bis A5 in Querschnittsanalysen.

→ Für die Ableitung der geschätzten Varianzen der mit OLS geschätzten Steigungsparameter sowie von t- und F-Statistiken muss aber zunächst erneut eine zusätzliche Annahme getroffen werden.

Annahmen:

- Annahmen B1 bis B5 (die bei der Betrachtung der Varianz von OLS-Schätzern getroffen werden)
- Annahme B6: Normalverteilung
  Die Störterme u_t sind unabhängig von X und unabhängig und identisch normalverteilt mit einem Erwartungswert null und einer Varianz σ^2, d.h. es gilt: $u_t \sim N(0; \sigma^2)$. Dabei impliziert B6 die Annahmen B3 bis B5, jedoch ist diese Annahme aufgrund der Unabhängigkeit und Normalverteilung stärker.

Auch unter den klassischen linearen Modellannahmen B1 bis B6 in Zeitreihenanalysen ergibt sich:

Die OLS-Schätzer sind unter der Bedingung von X normalverteilt, die konventionelle Konstruktion von Konfidenzintervallen ist gültig und unter den entsprechenden Nullhypothesen folgen die t- und F-Statistiken der t- und F-Verteilung

Damit ergeben sich mit B1 bis B6 dieselben Folgerungen wie mit den Annahmen A1 bis A6. Jedoch sind B1 bis B6 für Regressionsanalysen mit Zeitreihendaten restriktiver (insbesondere die strikte Exogenität sowie keine Autokorrelation sind häufig unrealistisch) als A1 bis A6 für Querschnittsanalysen.

→ Aufgrund der zuvor betrachteten bei Regressionsanalysen mit Zeitreihendaten sehr restriktiven Eigenschaften von OLS-Schätzern bei endlichen Stichprobenumfängen sind asymptotische Eigenschaften unter weniger restriktiven Annahmen bei Zeitreihenanalysen noch wichtiger als bei Querschnittsanalysen

Annahmen zur Betrachtung der Konsistenz von OLS-Schätzern:

- Annahme B1‘: Linearität und schwache Abhängigkeit
  Es gilt Annahme B1, d.h. der Zeitreihenprozess $\{(x_t, y_t): t = 1,\ldots, n\}$ folgt dem linearen Modell $y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t$. Zusätzlich ist der Zeitreihenprozess aber stationär und schwach abhängig.
- Annahme B2‘: Keine perfekte Kollinearität (so dass Annahme B2 gilt)
- Annahme B3‘: Bedingter Erwartungswert von u_t ist null
  Im Gegensatz zu Annahme B3 wird jetzt nicht mehr die strikte Exogenität der erklärenden Variablen, sondern lediglich die kontemporäre Exogenität entsprechend (5.3) betrachtet, d.h. $E(u_t|x_t) = 0$. Häufig wird für die Konsistenzeigenschaft auch lediglich folgendes vorausgesetzt:

$$(5.7) \quad E(u_t) = 0, \; Cov(x_{tj}, u_t) = 0 \quad j = 1,..., k$$

Unter diesen drei Annahmen sind die OLS-Schätzer $\hat{\beta}_j$ konsistent (wenngleich nicht unbedingt erwartungstreu), d.h. es gilt $\text{plim}(\hat{\beta}_j) = \beta_j$ für $j = 0,1,\ldots, k$.

Die Annahmen zur Ableitung der asymptotischen Normalverteilung von Funktionen von OLS-Schätzern in Zeitreihenanalysen und damit zur Durchführung von Testverfahren sind etwas weniger restriktiv als die klassischen linearen Modellannahmen B1 bis B6:

- Annahmen B1‘ bis B3‘ (die bei der Betrachtung der Konsistenz von OLS-Schätzern getroffen werden)
- Annahme B4‘: Kontemporäre Homoskedastizität
  Die bedingte Varianz des Fehlerterms u_t ist konstant und bezieht sich nicht mehr auf die erklärenden Variablen in allen Zeitperioden t = 1,…, n, sondern nur noch in Zeitperiode t: $Var(u_t|x_t) = Var(u_t) = \sigma^2$.
- Annahme B5‘: Keine Autokorrelation
  Für alle $t \neq s$ gilt $E(u_t u_s|x_t, x_s) = 0$, d.h. es wird nur noch auf die erklärenden Variablen in den Zeitperioden von u_t und u_s bedingt. Auch bei dieser Annahme wird zur einfacheren Interpretation oft von der Bedingung abstrahiert und lediglich die Unkorreliertheit von u_t und u_s betrachtet.

Unter den Annahmen B1‘ bis B5‘ ergibt sich, dass Funktionen der OLS-Schätzer asymptotisch normalverteilt und die OLS-Schätzer asymptotisch effizient sind sowie dass die t- und F-Statistiken asymptotisch t- und F-verteilt sind. Dadurch können dann die konventionellen t- und F-Tests durchgeführt und Konfidenzintervalle konstruiert werden.

Beispiel: Erklärung von Zinssätzen

Mit Hilfe eines linearen Regressionsmodells soll für die Jahre 1948 bis 2003 der Effekt der mit dem Preisindex für die Lebenshaltung gemessenen prozentualen jährlichen Inflationsrate (inf) und des staatlichen Haushaltsdefizits in % vom BIP (def) auf den Zinssatz für Schatzanweisungen (d.h. den kurzfristigen Geldmarktzins, „three-month T-bill rate“) (i3) in % untersucht werden. Dabei haben sich mit STATA folgende OLS-Schätzergebnisse gezeigt (n = 56, R^2 = 0,602):

```
reg i3 inf def

------------------------------------------------------------------------------
          i3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inf |   .6058659   .0821348     7.38   0.000     .4411243    .7706075
         def |   .5130578   .1183841     4.33   0.000     .2756095    .7505062
       _cons |   1.733266    .431967     4.01   0.000     .8668497    2.599682
-----------------------------------------------------------------------------
```

Interpretation:

- Sowohl die jährliche Inflationsrate als auch das staatliche Haushaltsdefizit haben einen signifikant positiven Effekt auf die kurzfristigen Zinssätze
- Zum Beispiel führt eine Erhöhung der jährlichen Inflationsrate um einen Prozentpunkt ceteris paribus zu einem geschätzten Anstieg des kurzfristigen Geldmarktzinses um 0,606 Prozentpunkte

5.3 Dummy-Variablen, Zeittrends und Saisonalität

→ Auch in Regressionsanalysen mit Zeitreihendaten werden häufig Dummy-Variablen, Interaktionsterme sowie logarithmierte abhängige und erklärende Variablen verwendet

Besonderheiten der Einbeziehung von Dummy-Variablen in Zeitreihenanalysen:

Da eine Beobachtungseinheit eine Zeitperiode darstellt, repräsentieren Dummy-Variablen, ob sich ein spezifisches Ereignis in einzelnen oder mehreren Perioden ergeben hat

Ausgehend von (5.1) werden nun zusätzlich qualitative erklärende Variablen mit q verschiedenen Ausprägungen betrachtet. Für diesen Fall können (maximal) q-1 Dummy-Variablen $x_{t01}, x_{t02},\ldots, x_{t,0,q-1}$ (neben den jetzt k-q+1 quantitativen erklärenden Variablen $x_{t1}, x_{t2},\ldots, x_{t,k-q+1}$) einbezogen werden:

$$(5.8)\quad y_t = \beta_0 + \beta_1 x_{t01} + \beta_2 x_{t02} + \cdots + \beta_{q-1} x_{t,0,q-1} + \beta_q x_{t1} + \beta_{q+1} x_{t2} + \cdots + \beta_k x_{t,k-q+1} + u_t$$

Die q-te Ausprägung der qualitativen Variablen (d.h. die Dummy-Variable x_{t0q}) dient dabei als Basisgruppe.

→ In Regressionsanalysen mit Zeitreihendaten werden Dummy-Variablen häufig zur Isolierung einzelner Zeitperioden, die sich systematisch unterscheiden, verwendet

---

Beispiel: Erklärung von Fertilitätsraten

Mit Hilfe eines linearen Regressionsmodells soll für die Jahre von 1913 bis 1984 der Effekt des durchschnittlichen Steuerfreibetrags (pe) sowie der Zeitperioden des Zweiten Weltkrieges von 1941 bis 1945 (ww2) und seit Einführung der Antibabypille ab 1963 (pill) auf die Anzahl der Geburten auf 1000 Frauen im gebärfähigen Alter (gfr) in den USA untersucht werden. Dabei haben sich mit STATA folgende OLS-Schätzergebnisse gezeigt (n = 72, R^2 = 0,473):

```
reg gfr pe ww2 pill
```

gfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pe	.08254	.0296462	2.78	0.007	.0233819	.1416981
ww2	-24.2384	7.458253	-3.25	0.002	-39.12111	-9.355686
pill	-31.59403	4.081068	-7.74	0.000	-39.73768	-23.45039
_cons	98.68176	3.208129	30.76	0.000	92.28003	105.0835

Die geschätzten Regressionsparameter implizieren, dass die Anzahl der Geburten auf 1000 Frauen im gebärfähigen Alter (bei gleichem durchschnittlichen Steuerfreibetrag) während des Zweiten Weltkrieges durchschnittlich um mehr als 24 und seit Einführung der Antibabypille um mehr als 31 geringer war als in den anderen Zeitperioden.

→ Eine besondere Bedeutung hat die Einbeziehung von Dummy-Variablen bei (finanzökonomischen) Event-Studien. Dabei wird untersucht, ob ein spezifisches Ereignis einen Einfluss auf Aktienrenditen hat.

Versionen von linearen Regressionsmodellen in Event-Studien:

$$R_t = \beta_0 + \beta_1 R_t^m + \beta_2 d_t + u_t$$

$$R_t = \beta_0 + \beta_1 R_t^m + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 d_t + u_t$$

Dabei bezeichnen R_t die Aktienrendite eines Unternehmens (oder Portfolios), R^m_t die Marktrendite sowie SMB_t und HML_t zwei weitere Kapitalmarktfaktoren (nach Fama und French) zur Erklärung von Aktienrenditen. Die Dummy-Variable d_t symbolisiert schließlich die Zeitperiode(n) für das spezifische Ereignis.

Beispiele für (neue überraschende zuvor nicht bekannte) Ereignisse :

- Zusammenschluss oder Übernahme von Unternehmen
- Einbeziehung in Aktienindizes (z.B. Nachhaltigkeitsindizes)
- Stör- oder Unfälle eines Unternehmens oder einer Branche (z.B. Unfall auf Ölplattform Deepwater Horizon, Atomkatastrophe von Fukushima)
- Phasen von Regulierungen wie z.B. Umweltgesetzgebungen
- Information über umwelt- oder sozialrelevante unternehmerische Maßnahmen (“Corporate Social Responsibility“)

Meist werden dabei jedoch alternative zweistufige Ansätze verwendet. Dabei werden zunächst vor dem Ereignis für ein Unternehmen z.B. folgende lineare Regressionsmodelle betrachtet:

$$R_t = \beta_0 + \beta_1 R_t^m + u_t$$

$$R_t = \beta_0 + \beta_1 R_t^m + \beta_2 SMB_t + \beta_3 HML_t + u_t$$

Die mit OLS geschätzten Regressionsparameter $\hat{\beta}_0$ und $\hat{\beta}_1$ im ersten Ansatz oder $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$ und $\hat{\beta}_3$ im zweiten Ansatz werden dann zur Schätzung der abnormalen Renditen AR_t in einer Zeitperiode t im Event-Fenster („event window“), d.h. während der Zeit des Einflusses des Ereignisses, verwendet:

$$A\hat{R}_t = R_t - \hat{\beta}_0 - \hat{\beta}_1 R_t^m$$

$$A\hat{R}_t = R_t - \hat{\beta}_0 - \hat{\beta}_1 R_t^m - \hat{\beta}_2 SMB_t - \hat{\beta}_3 HML_t$$

Diese geschätzten abnormalen Renditen können dann über verschiedene Zeitperioden im Event-Fenster und/oder über verschiedene Unternehmen aggregiert werden, so dass man geschätzte durchschnittliche, geschätzte kumulierte und geschätzte durchschnittliche kumulierte abnormale Renditen erhält. Mit Hilfe verschiedener statistischer Testverfahren wird dann überprüft, ob diese geschätzten abnormalen Renditen signifikant von null verschieden sind.

Zeittrends:

Zeitreihendaten können eine sinkende und vor allem wachsende Tendenz über die Zeit haben (z.B. Arbeitsproduktivität, nominale Importe)

Einfachstes Modell für lineare Zeittrends eines stochastischen Prozesses $\{y_t\}$:

$$(5.9) \quad y_t = \alpha_0 + \alpha_1 t + e_t \quad \text{für } t = 1, 2, ...$$

Im einfachen Fall stellt $\{e_t\}$ eine unabhängige identisch verteilte zufällige Sequenz mit $E(e_t) = 0$ und $Var(e_t) = \sigma_e^2$ dar. Falls alle Faktoren in e_t konstant sind mit $\Delta e_t = 0$, bildet α_1 die Veränderung von y von einer Zeitperiode zur nächsten Periode ab, d.h. $\Delta y_t = y_t - y_{t-1} = \alpha_1$. Zudem ergibt sich:

$$(5.10) \quad E(y_t) = \alpha_0 + \alpha_1 t$$

Bei $\alpha_1 > 0$ liegt im Durchschnitt ein wachsender Trend und bei $\alpha_1 < 0$ ein sinkender Trend vor. Im Gegensatz zum Erwartungswert ist die Varianz von y_t konstant über die Zeit.

→ Falls bei abhängigen und/oder erklärenden Variablen in Regressionsanalysen mit Zeitreihendaten Trends vorliegen, sollten diese einbezogen werden, da ansonsten scheinbare Zusammenhänge („spurious regression“) entstehen. Ohne die Einbeziehung von Trendvariablen könnten sich durch die Vernachlässigung relevanter erklärender Variablen verzerrte Schätzungen der Regressionsparameter ergeben („omitted variable bias“).

Zusätzliche Einbeziehung eines linearen Zeittrends in (5.1) (neben jetzt k-1 sonstigen erklärenden Variablen $x_{t1}, x_{t2}, \ldots, x_{t,k-1}$):

$$(5.11) \quad y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_{k-1} x_{t,k-1} + \beta_k t + u_t$$

Anmerkungen:

- Neben linearen Zeittrends können auch quadratische Zeittrends sowie weitere Polynome von t einbezogen werden, wenngleich die Trends möglichst einfach gehalten werden sollten
- Falls der Zeittrend signifikant von null verschieden ist und sich die Schätzergebnisse stark verändern, sollten die geschätzten Regressionsparameter ohne Einbeziehung der Trendvariablen sehr vorsichtig interpretiert werden
- Die Einbeziehung eines Zeittrends als erklärende Variable führt zu einer Trendbereinigung, d.h. die geschätzten Steigungsparameter der erklärenden Variablen können als geschätzte Effekte ohne den Zeittrend interpretiert werden
- Konventionelle und angepasste Bestimmtheitsmaße in Regressionsanalysen mit Zeitreihendaten können artifiziell sehr hohe Werte aufweisen, falls die abhängige Variable einen Trend aufweist. Dies bedeutet aber nicht unbedingt, dass ein sehr hoher Anteil der Variation der abhängigen Variablen durch die Variation der erklärenden Variablen erklärt wird, da die Varianz der abhängigen Variablen nicht mehr unverzerrt oder konsistent geschätzt wird.

Beispiel 1: Erklärung von Wohnungsbauinvestitionen

Mit Hilfe eines linearen Regressionsmodells soll für die USA für die Jahre 1947 bis 1988 der Effekt des Logarithmus des Wohnungspreisindexes (der in 1982 den Wert eins annimmt) (logprice) auf den Logarithmus der Pro-Kopf-Wohnungsbauinvestitionen in 1000 Dollar (loginvpc) untersucht werden. Dabei haben sich mit STATA folgende OLS-Schätzergebnisse gezeigt:

```
reg loginvpc logprice
      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  1,    40) =   10.53
       Model |  .254364557     1  .254364557           Prob > F      =  0.0024
    Residual |  .966255395    40  .024156385           R-squared     =  0.2084
-------------+------------------------------           Adj R-squared =  0.1886
       Total |  1.22061995    41  .029771218           Root MSE      =  .15542

------------------------------------------------------------------------------
    loginvpc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    logprice |   1.240943   .3824192     3.24   0.002     .4680455    2.013841
       _cons |  -.5502345   .0430266   -12.79   0.000    -.6371945   -.4632745
------------------------------------------------------------------------------
```

Beispiel 1: Erklärung von Wohnungsbauinvestitionen (Fortsetzung)

Allerdings muss bei dem Ergebnis einer signifikant positiven Preiselastizität beachtet werden, dass sowohl invpc als auch price einen wachsenden Trend aufweisen. Diese (lineare) Trendvariable muss einbezogen werden, um die Schätzung scheinbarer Zusammenhänge zu vermeiden. Dabei haben sich mit STATA folgende OLS-Schätzergebnisse gezeigt:

```
reg loginvpc logprice t
      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  2,    39) =   10.08
       Model |  .415945125     2  .207972562           Prob > F      =  0.0003
    Residual |  .804674828    39  .020632688           R-squared     =  0.3408
-------------+------------------------------           Adj R-squared =  0.3070
       Total |  1.22061995    41  .029771218           Root MSE      =  .14364

------------------------------------------------------------------------------
    loginvpc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    logprice |  -.3809609   .6788352    -0.56   0.578    -1.754035     .992113
           t |   .0098287   .0035122     2.80   0.008     .0027246    .0169328
       _cons |  -.9130595   .1356134    -6.73   0.000    -1.187363   -.6387556
------------------------------------------------------------------------------
```

Die geschätzte Preiselastizität ist nun negativ und nicht signifikant von null verschieden. Der Zeittrend ist signifikant positiv. Der geschätzte Regressionsparameter impliziert einen approximativen Anstieg von invpc um durchschnittlich 1% pro Jahr.

Beispiel 2: Erklärung von Fertilitätsraten

Wie zuvor soll erneut mit Hilfe eines linearen Regressionsmodells für die Jahre von 1913 bis 1984 der Effekt des durchschnittlichen Steuerfreibetrags (pe) sowie der Zeitperioden des Zweiten Weltkrieges von 1941 bis 1945 (ww2) und seit Einführung der Antibabypille ab 1963 (pill) auf die Fertilitätsrate (gfr) in den USA untersucht werden. Jetzt wird aber durch die Einbeziehung einer linearen Trendvariable eine Trendbereinigung durchgeführt. Dabei haben sich mit STATA folgende OLS-Schätzergebnisse gezeigt (n = 72, R^2 = 0,662):

```
reg gfr pe ww2 pill t
```

gfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pe	.2788778	.0400199	6.97	0.000	.1989978	.3587578
ww2	-35.59228	6.297377	-5.65	0.000	-48.1619	-23.02267
pill	.997447	6.26163	0.16	0.874	-11.50082	13.49571
t	-1.149872	.1879038	-6.12	0.000	-1.524929	-.7748145
_cons	111.7694	3.357765	33.29	0.000	105.0673	118.4716

Der geschätzte sinkende lineare Zeittrend ist hochsignifikant von null verschieden, so dass nun der geschätzte Regressionsparameter für pe mehr als dreimal so groß wie vorher und deutlicher signifikant von null verschieden ist .Vor allem aber hat dadurch nun die Einführung der Antibabypille ab 1963 keinen signifikanten Effekt mehr jenseits des Zeittrends.

Saisonalität:

Falls sich Zeitreihendaten auf Monate oder Quartale (bzw. auch Wochen oder Tage) beziehen, kann Saisonalität vorliegen, z.B. Wettereinflüsse bei makroökonomischen Variablen wie Beschäftigungszahlen. Andere Variablen wie z.B. Zins- oder Inflationsraten weisen dagegen selten Saisonalität auf. Daten für einige Variablen mit Saisonalität können bereits im Vorfeld saisonbereinigt sein.

Falls Saisonalität bei abhängigen und/oder erklärenden Variablen in Regressionsanalysen mit Zeitreihendaten vorliegt und die Daten noch nicht saisonbereinigt wurden, sollten entsprechende saisonale Dummy-Variablen betrachtet werden. Bei Monatsdaten und der zusätzlichen Einbeziehung von elf Dummy-Variablen für die Monate Februar (feb) bis Dezember (dec) in (5.1) (neben jetzt k-11 sonstigen erklärenden Variablen $x_{t1}, x_{t2},\dots, x_{t,k-11}$) ergibt sich:

$$(5.12) \quad y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_{k-11} x_{t,k-11} + \beta_{k-10} feb_t + \cdots + \beta_k dec_t + u_t$$

Anmerkungen:

- Bei Quartalsdaten können Dummy-Variablen für maximal drei Quartale einbezogen werden
- Die Einbeziehung von saisonalen Dummy-Variablen führt zu einer Saisonbereinigung, d.h. die geschätzten Steigungsparameter der erklärenden Variablen können als geschätzte Effekte ohne Saisonalität interpretiert werden
- Saisonale Dummy-Variablen können neben Trendvariablen einbezogen werden, so dass eine Saison- und Trendbereinigung durchgeführt wird

5.4 Autokorrelierte Fehlerterme

Eigenschaften von OLS-Schätzern bei autokorrelierten Fehlertermen:

- Unter den Annahmen B1 bis B3 sind die OLS-Schätzer $\hat{\beta}_j$ unabhängig von der Stärke der Autokorrelation der Störterme erwartungstreu, d.h. der Erwartungswert der Schätzer ist gleich den unbekannten Regressionsparametern
- Unter den Annahmen B1‘ bis B3‘ sind die OLS-Schätzer $\hat{\beta}_j$ unabhängig von der Stärke der Autokorrelation der Störterme konsistent, d.h. es gilt somit $\text{plim}(\hat{\beta}_j) = \beta_j$ für $j = 0,1,\ldots, k$
- Allerdings werden bei autokorrelierten Fehlertermen nicht mehr die Gauss-Markov-Annahmen erfüllt, so dass die OLS-Schätzer in diesem Fall nicht mehr die BLUE-Eigenschaft (bzw. Effizienz) aufweisen
- Vor allem aber sind bei autokorrelierten Fehlertermen die konventionellen Schätzungen der Varianzen der mit OLS geschätzten Steigungsparameter verzerrt und somit auch die t- und F-Statistiken nicht einmal mehr asymptotisch t- und F-verteilt
- Die wichtigste Form autokorrelierter Fehlerterme ergibt sich durch einen entsprechenden AR(1) Prozess, also durch folgende AR(1) Autokorrelation:

$$(5.13) \quad u_t = \rho u_{t-1} + e_t \quad \text{für } t = 1, 2, \ldots, n$$

Dabei gilt die Stabilitätsbedingung $|\rho| < 1$ und die e_t sind unkorreliert mit Erwartungswert null und Varianz σ_e^2.

Ausgangspunkt der folgenden Tests auf autokorrelierte Fehlerterme ist das multiple lineare Regressionsmodell nach (5.1):

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + u_t$$

t-Test auf AR(1) Autokorrelation der Fehlerterme mit $u_t = \rho u_{t-1} + e_t$ (t = 2, 3, …) bei strikter Exogenität der erklärenden Variablen:

- Angenommen wird dabei, dass $\{e_t\}$ ein i.i.d. Zeitreihenprozess ist, so dass auch Homoskedastizität vorliegt:

  (5.14) $E(e_t|u_{t-1}, u_{t-2},...) = 0$

  (5.15) $Var(e_t|u_{t-1}) = Var(e_t) = \sigma_e^2$

- Die Nullhypothese lautet:

  (5.16) $H_0: \rho = 0$

- Die Nullhypothese könnte dadurch überprüft werden, dass ein gewöhnlicher t-Test zu ρ bei der Regression von u_t auf u_{t-1} angewendet wird. Allerdings sind die u_t unbekannt und werden deshalb durch die entsprechenden OLS-Residuen $\hat{u}_t$ ersetzt. Durch die Annahme der strikten Exogenität der erklärenden Variablen wird die asymptotische t-Verteilung der Prüfgröße gewährleistet.

- Somit ergibt sich folgendes Vorgehen bei diesem t-Test:

(1) Mit Hilfe der OLS-Regressionswerte bei der Regression von y_t auf die erklärenden Variablen $x_{t1},\ldots, x_{tk}$ werden zunächst die Residuen $\hat{u}_t$ für alle $t = 1,2,\ldots, n$ ermittelt

(2) Danach wird in folgendem Hilfsregressionsmodell der OLS-Schätzer $\hat{\rho}$ ermittelt (dabei kann eine Konstante einbezogen werden, da sich auch dann eine asymptotische t-Verteilung der Teststatistik ergibt):

$$(5.17)\quad \hat{u}_t = \rho\hat{u}_{t-1} + e_t \quad \text{für } t = 2, 3,\ldots, n$$

(3) Die t-Statistik $t_{\hat{\rho}}$ wird dann zur Überprüfung von H_0: $\rho = 0$ verwendet

- Obwohl dieser Test zur Überprüfung von AR(1) Autokorrelationen in den Fehlertermen entwickelt wurde, können damit auch andere Formen von (angrenzenden) Autokorrelationen aufgedeckt werden
- Falls keine Homoskedastizität nach (5.15) vorliegt, können heteroskedastizitäts-robuste t-Statistiken angewendet werden

→ Es können auch t-Tests auf AR(1) Autokorrelation der Fehlerterme ohne strikte Exogenität der erklärenden Variablen durchgeführt werden. Dabei werden die $\hat{u}_t$ zur Überprüfung von H_0: $\rho = 0$ nicht nur auf $\hat{u}_{t-1}$, sondern auch auf alle erklärenden Variablen und einer Konstante regressiert. Dadurch ergibt sich die approximative t-Verteilung der Prüfgröße, selbst wenn die erklärenden Variablen nicht strikt exogen sind.

Beispiel: Philips-Kurve

Ein sehr einfacher Ansatz zum Zusammenhang zwischen Inflationsrate (inf) und Arbeitslosenrate (unem) basiert auf der statischen Philips-Kurve. Für die USA wurde dabei für die Jahre von 1948 bis 1996 folgende OLS-Regressionsfunktion geschätzt ($n = 49$, $R^2 = 0{,}053$):

$$\widehat{\text{inf}}_t = \underset{(1{,}72)}{1{,}42} + \underset{(0{,}289)}{0{,}468}\,\text{unem}_t$$

Auf Basis dieser OLS-Schätzung wurden die $\hat{u}_t$ auf die $\hat{u}_{t-1}$ ($n = 48$) regressiert:

$$\hat{\hat{u}}_t = \underset{(0{,}115)}{0{,}573}\,\hat{u}_{t-1}$$

Damit ergibt sich $t_{\hat{\rho}} = 4{,}98$ und damit eine sehr starke Evidenz für eine AR(1) Autokorrelation der Störterme. Daraus folgt, dass die obige t-Statistik zur Überprüfung des Effektes der Arbeitslosigkeit auf die Inflation nicht zuverlässig ist.

→ Bei der Untersuchung einer flexibleren Form der Philips-Kurve („expectations augmented Philips curve“), kann dagegen keine Autokorrelation erster Ordnung der Störterme nachgewiesen werden

---

Beispiel: Philips-Kurve (STATA-Output)

Mit STATA haben sich folgende OLS-Schätzergebnisse gezeigt:

```
tsset year
reg inf unem
      Source |       SS       df       MS              Number of obs =      49
-------------+------------------------------           F(  1,    47) =    2.62
       Model |  25.6369586     1  25.6369586           Prob > F      =  0.1125
    Residual |  460.619776    47  9.80042077           R-squared     =  0.0527
-------------+------------------------------           Adj R-squared =  0.0326
       Total |  486.256735    48  10.1303486           Root MSE      =  3.1306

------------------------------------------------------------------------------
         inf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        unem |   .4676257   .2891262     1.62   0.112    -.1140212    1.049273
       _cons |    1.42361   1.719015     0.83   0.412    -2.034602    4.881822
------------------------------------------------------------------------------
predict u, resid
reg u l.u, noconstant
      Source |       SS       df       MS              Number of obs =      48
-------------+------------------------------           F(  1,    47) =   24.80
       Model |  150.799931     1  150.799931           Prob > F      =  0.0000
    Residual |  285.815602    47  6.08118302           R-squared     =  0.3454
-------------+------------------------------           Adj R-squared =  0.3315
       Total |  436.615533    48  9.09615694           Root MSE      =   2.466

------------------------------------------------------------------------------
           u |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           u |
         L1. |   .5727355   .1150132     4.98   0.000     .3413588    .8041121
------------------------------------------------------------------------------
```

Durbin-Watson-Test auf AR(1) Autokorrelation der Störterme:

- Die Durbin-Watson-Teststatistik basiert ebenso wie der vorherige t-Test auf den OLS-Residuen:

$$(5.18) \quad DW = \frac{\sum_{t=2}^{n} \left(\hat{u}_t - \hat{u}_{t-1}\right)^2}{\sum_{t=1}^{n} \hat{u}_t^2}$$

- Dabei sind die Durbin-Watson-Teststatistik und der vorher diskutierte OLS-Schätzer $\hat{\rho}$ eng miteinander verknüpft:

$$(5.19) \quad DW \approx 2(1-\hat{\rho})$$

- Der Zusammenhang ist nicht exakt, da der Nenner bei der OLS-Schätzung von $\hat{\rho}$ im Gegensatz zur Durbin-Watson-Teststatistik die Summe $\sum_{t=2}^{n} \hat{u}_{t-1}{}^2$ enthält. Allerdings ist die Approximation meistens sehr gut. Dies weist auf die konzeptuelle Gleichheit des Durbin-Watson-Tests und des zuvor betrachteten t-Tests hin.
- Durbin und Watson haben die Verteilung von DW abgeleitet. Diese basiert allerdings zum einen auf allen klassischen linearen Modellannahmen B1 bis B6 und damit insbesondere auch auf der Normalverteilung der Störterme. Darüber hinaus hängt die Verteilung von der Anzahl und den Werten der erklärenden Variablen, der Anzahl der betrachteten Zeitperioden sowie davon ab, ob eine Konstante in das Regressionsmodell einbezogen wurde.

→ Falls autokorrelierte Fehlerterme vorliegen, sollte hierauf bei der Schätzung der Regressionsparameter und/oder bei darauf aufbauenden Testverfahren in linearen Regressionsmodellen reagiert werden.

Mögliche Ansätze:

- Ein Ansatz ist die Transformation der Variablen durch die Einbeziehung von (z.B. ersten) Differenzen der abhängigen und erklärenden Variablen. Durch diese Transformation kann häufig eine Autokorrelation der Störterme ganz vermieden werden.
- Ein alternativer Ansatz ist die Anwendung einer zu OLS alternativen Schätzmethode. Beispiele hierfür sind verschiedene verallgemeinerte Methoden der kleinsten Quadrate (GLS) wie z.B. die Cochrane-Orcutt- oder die Prais-Winston-Methode. Damit können einzelne Formen autokorrelierter Störterme bei der Parameterschätzung berücksichtigt werden.

→ Allerdings hat die Anwendung von GLS-Methoden eine Reihe von strengen Anforderungen. So müssen z.B. die erklärenden Variablen strikt exogen sein, da GLS-Schätzer ansonsten nicht einmal konsistent sind. Zudem wird häufig die spezifische AR(1) Autokorrelation der Fehlerterme angenommen.

→ Aus diesem Grund werden viel häufiger die Regressionsparameter (ineffizient) mit OLS geschätzt, die Schätzung der Varianzen der geschätzten Regressionsparameter aber korrigiert. Damit können in Analogie zur Betrachtung von heteroskedastizitäts-robusten t-Statistiken autokorrelations-robuste Konfidenzintervalle und vor allem t-Statistiken abgeleitet werden.

Ausgangspunkt eines Ansatzes zur Ableitung von autokorrelations-robusten Schätzungen der Varianz der geschätzten Regressionsparameter ist das multiple lineare Regressionsmodell nach (5.1):

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + u_t$$

Dabei soll zunächst eine autokorrelations-robuste Schätzung der Standardabweichung des OLS-Schätzers $\hat{\beta}_1$ betrachtet werden. Hierzu wird folgendes Hilfsregressionsmodell untersucht, das die erste erklärende Variable x_{t1} als abhängige Variable und alle anderen erklärenden Variablen auch hier als erklärende Variablen beinhaltet:

$$(5.20) \quad x_{t1} = \delta_0 + \delta_1 x_{t2} + \cdots + \delta_{k-1} x_{tk} + r_{t1} \quad \text{für } t = 1, ..., n$$

Dabei hat der Störterm r_{t1} einen Erwartungswert von null und ist unkorreliert mit den erklärenden Variablen $x_{t2}, \ldots, x_{tk}$.

Für den Schätzer der Standardabweichung des OLS-Schätzers $\hat{\beta}_1$ wird im Folgenden weiterhin der Schätzer $\hat{\sigma}$ der Standardabweichung σ des Fehlerterms u_t betrachtet. Mit $\hat{r}_{t1}$ als Residuum der OLS-Schätzung in (5.20) in Zeitperiode t kann mit g > 0 (wobei g kontrolliert, welches Ausmaß an Autokorrelation in die Betrachtung einbezogen wird) folgender Ansatz abgeleitet werden:

$$(5.21) \quad \hat{v}_1 = \sum_{t=1}^{n} \hat{r}_{t1}^2 \hat{u}_t^2 + 2\sum_{h=1}^{g} \left(1 - \frac{h}{g+1}\right)\left(\sum_{t=h+1}^{n} \hat{r}_{t1}\hat{u}_t \hat{r}_{t-h,1}\hat{u}_{t-h}\right)$$

Je größer g ist, desto mehr Terme werden zur Korrektur der Autokorrelation einbezogen. Im einfachsten Fall mit g = 1 ergibt sich:

$$\hat{v}_1 = \sum_{t=1}^{n} \hat{r}_{t1}^2 \hat{u}_t^2 + \sum_{t=2}^{n} \hat{r}_{t1}\hat{u}_t \hat{r}_{t-1,1}\hat{u}_{t-1}$$

Mit der Division des üblichen Schätzers der Standardabweichung von $\hat{\beta}_1$ durch $\hat{\sigma}$ sowie der anschließenden Multiplikation des entsprechenden quadrierten Terms mit der Wurzel von $\hat{v}_1$ in (5.21) ergibt sich nun folgender autokorrelations-robuster Schätzer der Standardabweichung:

$$(5.22) \quad \sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_1)} = \left[\frac{\dfrac{\hat{\sigma}}{\sqrt{(1-R_1^2)SST_1}}}{\hat{\sigma}}\right]^2 \sqrt{\hat{v}_1} = \frac{\sqrt{\hat{v}_1}}{(1-R_1^2)SST_1} = \frac{\sqrt{\hat{v}_1}}{SSR_1}$$

Diese Schätzung kann analog auf beliebige OLS-Schätzer $\hat{\beta}_j$ angewendet werden (mit Störterm r_{tj} sowie $\hat{v}_j$). Dabei sind die Schätzungen nicht nur autokorrelations-robust, sondern auch robust für beliebige Formen der Heteroskedastizität, so dass sie auch als heteroskedastizitäts- und autokorrelations-robuste (HAC) Schätzungen der Standardabweichung der geschätzten Steigungsparameter bezeichnet werden. Diese Robustheit für Heteroskedastizität zeigt sich, wenn für $\hat{\beta}_j$ nur der erste Term von $\hat{v}_j$ in Analogie zu (5.21) einbezogen wird:

$$(5.23)\quad \sqrt{\hat{\operatorname{Var}}(\hat{\beta}_j)} = \frac{\sqrt{\sum_{t=1}^{n} \hat{r}_{tj}^2 \hat{u}_t^2}}{(1-R_j^2)SST_j} = \frac{\sqrt{\sum_{t=1}^{n} \hat{r}_{tj}^2 \hat{u}_t^2}}{SSR_j}$$

Damit ergibt sich die Analogie zur heteroskedastizitäts-robusten Schätzung der Standardabweichung von mit OLS geschätzten Steigungsparametern in (4.16).

Anmerkungen:

- Die Einbeziehung der HAC-Schätzung der Standardabweichung der geschätzten Steigungsparameter in t-Statistiken führt zu heteroskedastizitäts- und autokorrelations-robusten t-Statistiken
- Mit wachsendem n sollte auch die Zahl g wachsen, da bei einer großen Anzahl an Zeitperioden auch das Ausmaß der Autokorrelation der Fehlerterme steigen kann. Faustregeln sind z.B. $g = 4(n/100)^{2/9}$ (nach Newey und West, 1987) oder aber $g = n^{1/4}$.

Beispiel: Erklärung von Beschäftigungsraten

Mit Hilfe eines linearen Regressionsmodells soll für die Jahre von 1950 bis 1987 der Effekt des Logarithmus der Bedeutung des U.S. Mindestlohnes (logmincov), des Logarithmus des U.S. Bruttosozialprodukts (logusgnp) und des Logarithmus des Bruttosozialprodukts in Puerto Rico (logprgnp) unter Einbeziehung einer linearen Trendvariablen auf den Logarithmus der Beschäftigungsrate in Puerto Rico (logprepop) untersucht werden. Bei einem t-Test auf AR(1) Autokorrelation der Fehlerterme ohne strikte Exogenität der erklärenden Variablen hat sich ein klarer Hinweis auf AR(1) Autokorrelation ergeben, so dass autokorrelations-robuste t-Statistiken betrachtet werden.

Ergebnisse:

- Der OLS-Schätzer der Elastizität der Beschäftigungsrate in Bezug auf den Mindestlohn beträgt -0,2123 und der übliche Schätzwert der Standardabweichung des geschätzten Steigungsparameters beträgt 0,0402
- Mit g = 2 ergibt sich ein heteroskedastizitäts- und autokorrelations-robuster Schätzer der Standardabweichung des Steigungsparameters von 0,0457 und ist damit nur geringfügig höher
- Die robuste t-Statistik beträgt -4,64, so dass die geschätzte Elastizität weiterhin hoch signifikant von null verschieden ist

Beispiel: Erklärung von Beschäftigungsraten (STATA-Output)

Mit STATA haben sich folgende OLS-Schätzergebnisse gezeigt:

```
tsset year
reg logprepop logmincov logusgnp logprgnp t
      Source |       SS       df       MS              Number of obs =      38
-------------+------------------------------           F(  4,    33) =   66.23
       Model |  .284429837     4  .071107459           Prob > F      =  0.0000
    Residual |   .03542846    33   .00107359           R-squared     =  0.8892
-------------+------------------------------           Adj R-squared =  0.8758
       Total |  .319858296    37  .008644819           Root MSE      =  .03277

------------------------------------------------------------------------------
   logprepop |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   logmincov |  -.2122612   .0401525    -5.29   0.000    -.2939519   -.1305704
    logusgnp |   .4860432   .2219834     2.19   0.036     .0344145    .9376719
    logprgnp |   .2852395   .0804922     3.54   0.001     .1214768    .4490022
           t |  -.0266633   .0046267    -5.76   0.000    -.0360764   -.0172501
       _cons |  -6.663416   1.257835    -5.30   0.000    -9.222501   -4.104331
------------------------------------------------------------------------------

newey logprepop logmincov logusgnp logprgnp t, lag(2)
Regression with Newey-West standard errors          Number of obs  =        38
maximum lag: 2                                      F(  4,    33)  =     37.84
                                                    Prob > F       =    0.0000

------------------------------------------------------------------------------
             |             Newey-West
   logprepop |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   logmincov |  -.2122611   .0457188    -4.64   0.000    -.3052768   -.1192455
    logusgnp |   .4860416   .2791144     1.74   0.091     -.081821    1.053904
    logprgnp |   .2852399   .0996364     2.86   0.007      .082528    .4879518
           t |  -.0266632   .0057559    -4.63   0.000    -.0383736   -.0149528
       _cons |  -6.663407   1.536445    -4.34   0.000    -9.789328   -3.537485
------------------------------------------------------------------------------
```