

Bachelorprüfung

Fach: Empirische Wirtschaftsforschung II

Prüfer: Prof. Regina T. Riphahn, Ph.D.

Name, Vorname	
Matrikelnr.	
Studiengang	
Semester	
Email	
Datum	
Hausaufgabe	Haben Sie im WS10/11 eine Hausaufgabe abgegeben? <input type="checkbox"/> Ja <input type="checkbox"/> Nein
Unterschrift	

Vorbemerkungen:

- Anzahl der Aufgaben:** Die Klausur besteht aus 7 Aufgaben, die alle bearbeitet werden müssen.
- Bewertung:** Es können maximal 90 Punkte erworben werden. Die Punktzahl ist für jede Aufgabe in Klammern angegeben. Sie entspricht der für die Aufgabe empfohlenen Bearbeitungszeit in Minuten.
- Erlaubte Hilfsmittel:** Tabellen der statistischen Verteilung und Liste der Annahmen (sind der Klausur beigelegt)
1 DIN-A4-Seite mit Notizen
Taschenrechner
Fremdwörterbuch
- Wichtige Hinweise:** Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den exakten Wert der Freiheitsgrade nicht ausweisen, machen Sie dies kenntlich und verwenden Sie den nächstgelegenen Wert.
Sollte es vorkommen, dass bei einer Berechnung eine erforderliche Information fehlt, machen Sie dies kenntlich und treffen Sie für den fehlenden Wert eine plausible Annahme.

Aufgabe 1**[7 Punkte]**

- a) Erläutern Sie das Konzept der Kovarianzstationarität. (2 Punkte)
- b) Angenommen $\{y_t: t = 1, 2, \dots\}$ folgt $y_t = \beta_0 + \beta_1 t + e_t$ mit $\beta_1 \neq 0$. $\{e_t: t = 1, 2, \dots\}$ ist eine i.i.d. Zeitreihe mit Erwartungswert Null und konstanter Varianz σ^2 .
- Ist $\{y_t\}$ kovarianzstationär? Begründen Sie. (1 Punkt)
 - Ist $y_t - E(y_t)$ kovarianzstationär? Begründen Sie. (1 Punkt)
- c) Definieren Sie die Eigenschaft der schwachen Abhängigkeit einer Zeitreihe. Ist der Prozess $u_t = \rho u_{t-1} + e_t$ (mit e_t als i.i.d. Sequenz mit Erwartungswert Null und konstanter Varianz) bei $\rho = 1$ schwach abhängig? Begründen Sie kurz. (3 Punkte)

Aufgabe 2**[18 Punkte]**

- a) Welche Konsequenzen hat das Vorliegen eines AR(1) Prozesses im Störterm für eine KQ Schätzung der Modells $y_t = \beta_0 + \beta_1 x_t + e_t$? Nennen Sie zwei. (2 Punkte)
- b) Beschreiben Sie im Detail Ziel und Vorgehen des Cochrane-Orcutt Verfahrens für den Fall, dass für den Störterm e_t gilt: $e_t = \rho e_{t-1} + v_t$, v_t den Gauss Markov Annahmen entspricht und der Wert von $|\rho| < 1$ bekannt ist? Ist der Schätzer BLUE? Begründen Sie. (5 Punkte)
- c) Welche über die in Teilaufgabe a) hinausgehende Konsequenz hat das Vorliegen eines stabilen AR(1) Prozesses im Störterm für eine KQ Schätzung des Modells $y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 x_t + e_t$? Begründen Sie! (2 Punkte)
- d) Zeigen Sie, dass die Aufnahme von y_{t-2} in das Modell $y_t = \beta_0 + \beta_1 y_{t-1} + e_t$ einen stabilen AR(1) Prozess im Störterm e_t eliminiert. Erläutern Sie im Rahmen Ihrer Antwort auch kurz das Konzept der dynamischen Vollständigkeit eines Zeitreihenmodells. (5 Punkte)
- e) Beschreiben Sie ein Verfahren mit dem Sie im Modell $y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 x_t + e_t$ das Vorliegen eines AR(1) Prozesses testen können. Nennen Sie Null- und Alternativhypothese, Teststatistik und Schlusslogik. (4 Punkte)

Aufgabe 3**[3 Punkte]**

Sie interessieren sich für den Zusammenhang zwischen Löhnen und Betriebsgröße und schätzen folgendes Modell:

$$\ln \text{lohn}_i = \beta_0 + \beta_1 \text{betriebsgröße}_i + e_i$$

wobei i die Beobachtungseinheit ist (Personen).

$\ln \text{lohn}$ = Stundenlohn (logarithmiert).

betriebsgröße = Anzahl Mitarbeiter des Betriebes, in dem i arbeitet.

- a) Nehmen Sie an, dass die Streuung der Löhne mit der Betriebsgröße zunimmt. Welche Konsequenzen hat dies für den Störterm und $\hat{\beta}_1$, wenn das Modell mit KQ geschätzt wird? (2 Punkte)

- b) Nennen Sie kurz eine Methode, mit deren Hilfe man das in a) angesprochene Problem beheben kann. (1 Punkt)

Aufgabe 4

[9 Punkte]

Das Management eines Bekleidungsunternehmens sucht nach Wegen, die Produktivität seiner Arbeiter zu steigern. Bisher hat jeder Arbeiter für sich allein gearbeitet, im Jahr 1990 dürfen sich Arbeiter jedoch freiwillig zu Teams zusammenschließen. Ein Teil der Arbeiter macht von dieser Möglichkeit Gebrauch.

Mit Hilfe eines Datensatzes aus dem Jahr 1990 möchten Sie untersuchen, ob Teamarbeit einen Einfluss auf die Produktivität hat. Sie schätzen hierfür folgendes Modell mit KQ:

$$\ln \text{Anzahl}_{it} = \beta_0 + \beta_1 \text{Team}_{it} + \beta_2 \text{Frau}_i + \beta_3 \text{Erf}_i + e_{it}$$

i ist der Personenindikator und $t \in \{1, \dots, 52\}$ der Kalenderwochenindikator.

$\ln \text{Anzahl}$ = logarithmierte Anzahl der Kleidungsstücke, die Arbeiter i in Woche t hergestellt hat.

Team = 1, falls i zum Zeitpunkt t im Team gearbeitet hat, sonst = 0.

Frau = 1, falls i eine Frau ist, sonst = 0.

Erf = Berufserfahrung zum Zeitpunkt der Einstellung (in Jahren).

- Sie schätzen $\hat{\beta}_1 = 0,18$ mit p-Wert 0,007. Interpretieren Sie $\hat{\beta}_1$ inhaltlich und statistisch. (2 Punkte)
- Die Arbeiter durften sich freiwillig zu Teams zusammenschließen. Warum könnte dies zu einer verzerrten Schätzung von β_1 führen? Zeigen Sie formal und erläutern Sie. (3 Punkte)
- Schreiben Sie für das gegebene Beispiel die Gleichung für den Within-Schätzer auf. Erläutern Sie, inwieweit mit diesem Schätzer das in Teilaufgabe b) angesprochene Problem gelöst werden kann. Unter welcher Bedingung ist der neue Schätzer konsistent? (4 Punkte)

Aufgabe 5

[13 Punkte]

Sie möchten untersuchen, ob zwischen dem Arbeitsangebot von Frauen und der Kinderzahl ein Zusammenhang besteht. Hierfür steht Ihnen ein Datensatz von Müttern zur Verfügung, die bereits zwei oder mehr Kinder haben und zwischen 21 und 35 Jahre alt sind. Sie verwenden folgendes Regressionsmodell:

$$\text{stunden}_i = \beta_0 + \beta_1 \text{vieleKinder}_i + \beta_2 \text{alter}_i + \beta_3 \text{bildung}_i + e_i$$

wobei i der Personenindikator ist.

stunden = Wochenarbeitszeit in Stunden.

vieleKinder = 1, falls Anzahl Kinder ≥ 3 ist. = 0, falls Anzahl Kinder = 2.

alter = Alter der Mutter in Jahren.

bildung = Anzahl Schuljahre der Mutter.

- a) Sie schätzen das Modell mit KQ und erhalten $\hat{\beta}_1 = -3,76$. Interpretieren Sie $\hat{\beta}_1$ inhaltlich (Gehen Sie davon aus, dass $\hat{\beta}_1$ auf dem 1% Niveau signifikant ist). (2 Punkte)
- b) Welche Annahme des KQ Modells ist verletzt, falls *vieleKinder* endogen ist? (1 Punkt)
- c) Erläutern Sie inhaltlich, warum *vieleKinder* endogen sein könnte. Geben Sie hierfür ein Beispiel an und erklären Sie, in welche Richtung $\hat{\beta}_1$ in diesem Fall verzerrt ist. (4 Punkte)
- d) Eine Freundin von Ihnen schlägt vor, *vieleKinder* mit der Variable *glGeschl* zu instrumentieren. *glGeschl* = 1 falls das erste und zweite Kind das gleiche Geschlecht haben, ansonsten ist *glGeschl* = 0. Sie befolgen den Vorschlag und erhalten nun $\hat{\beta}_1^{IV} = -0,24$ mit Standardfehler 0,06.
- Warum könnte *glGeschl* ein geeignetes Instrument sein? Begründen Sie formal und inhaltlich. (4 Punkte)
 - Wie lässt sich der große Unterschied zwischen $\hat{\beta}_1$ aus Teilaufgabe a) und $\hat{\beta}_1^{IV}$ erklären? (2 Punkte)

Aufgabe 6

[25 Punkte]

Wahr oder falsch? Tragen Sie für jede der folgenden Aussagen ein „w“ für „wahr“ oder ein „f“ für „falsch“ ein. Für jede richtige Antwort gibt es 0,5 Punkte, für jede falsche Antwort werden 0,5 Punkte abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

	Bei Heteroskedastie ist der KQ-Schätzer dem White-Schätzer zur Bestimmung der Standardfehler vorzuziehen.
	Finite Distributed Lag Modelle können mehr als 3 verzögerte Werte von x berücksichtigen.
	Sind die Regressoren eines Zeitreihenmodells nicht kontemporär exogen, so sind die KQ Schätzer inkonsistent.
	Die Relevanz eines Instruments lässt sich empirisch testen.
	Das Spurious Regression Problem bei Zeitreihenregression tritt nur dann auf, wenn abhängige Variable und Störterm einem gleichlaufenden Trend folgen.
	Für einen Moving Average Prozess x_t der Ordnung a gilt $Cov[x_t, x_{t+h}] = 0$ sobald $h > a$.
	Einheitswurzelprozesse können durch Bildung der ersten Differenzen in schwach abhängige Prozesse überführt werden.
	Das LAD Verfahren minimiert die Summe der Absolutwerte der Residuen.
	In dynamisch vollständigen Modellen folgen die Störterme einem Random Walk.
	Beim Schätzen mit Paneldaten in ersten Differenzen werden zeitvariable, unbeobachtete Effekte kontrolliert.
	Im Fixed Effects Modell können keine Interaktionen der erklärenden Variablen mit einem Zeittrend aufgenommen werden, da Zeittrends im Fixed Effects Modell durch die Within Transformation heraus gekürzt werden.
	KQ-Schätzer sind inkonsistent, wenn die abhängige Variable stationär ist.
	Eine Trendbereinigung kann auch ohne zusätzliche Bereinigung um Saisoneffekte Sinn machen.
	Heteroskedastierobuste t-Werte sind stets größer als die herkömmlichen.
	Systematische Stichprobenselektion anhand der Ausprägungen der abhängigen Variable kann zu verzerrten Parameterschätzern führen.
	Eine Dummy-Variable kann aufgrund der sich ergebenden Heteroskedastieproblematik nicht als Proxy-Variable verwendet werden.

	Werden relevante erklärende Variablen nicht berücksichtigt, so liefert die KQ Schätzung stets verzerrte Schätzer für die Koeffizienten der anderen Regressoren.
	Kann die H_0 im White-Test verworfen werden, schließen wir auf Heteroskedastie.
	Treten in den Daten fehlende Werte nicht zufällig auf, kann das Weglassen der Beobachtungen mit fehlenden Werten zu verzerrten KQ Schätzern führen.
	In Finite Distributed Lag Modellen führt starke Multikollinearität zwischen den verwendeten Lags der exogenen Variablen häufig zu verzerrten Schätzergebnissen.
	Wenn nicht für alle Beobachtungseinheiten i gleich viele Beobachtungsperioden t vorliegen, spricht man von „Unbalanced Panels“.
	Die Verwendung von verzögerten endogenen Variablen auf der rechten Seite eines Zeitreihenmodells führt zwangsläufig zu inkonsistenten Schätzergebnissen.
	Ausreißer werden beim LAD Verfahren geringer gewichtet als beim KQ Verfahren.
	Gepoolte Querschnitte betrachten dieselben Beobachtungseinheiten zu unterschiedlichen Zeitpunkten.
	Um einen Difference-in-Difference-Schätzer zu ermitteln, werden erste Differenzen in der abhängigen Variable von der ersten Differenz der erklärenden Variablen abgezogen.
	Eine Schätzung in ersten Differenzen ist nicht möglich, wenn die betrachteten Variablen über die Zeit hinweg konstant sind.
	Wenn der Between Schätzer verzerrt ist, ist auch der Within Schätzer verzerrt.
	Wenn $T = 2$, liefern der Within Schätzer und der Schätzer in ersten Differenzen identische Koeffizienten und Standardfehler.
	Im Random Effects Modell werden den Beobachtungseinheiten zufällige Effekte der erklärenden Variablen auf die abhängige Variable zugewiesen.
	Der Within Schätzer kann den Effekt zeitkonstanter erklärender Variablen nicht bestimmen.
	Das Fixed Effects Modell kann auch für gepoolte Querschnittsdaten angewendet werden.
	Im Trendmodell $\log(y_t) = \beta_0 + \beta_1 t + e_t$ kann β_1 approximativ als mittlere periodische Wachstumsrate von y interpretiert werden.
	Instrumentvariablen müssen mit den endogenen erklärenden Variablen unkorreliert sein.
	Finite Distributed Lag Modelle weisen mindestens einen verzögerten Wert von x auf.
	Sind die Regressoren eines Zeitreihenmodells nicht strikt exogen, so sind die KQ Schätzer stets inkonsistent.
	Eine wichtige Annahme beim Difference-in-Differences Verfahren ist, dass sich die Treatmentgruppe zu beiden Beobachtungszeitpunkten identisch zusammensetzt.
	Das White-Schätzverfahren dient zur Korrektur der Standardfehler bei Vorliegen von Heteroskedastie.
	Newey-West t-Werte sind stets kleiner als die bei KQ ausgewiesenen t-Werte.
	Eine stetige Variable kann als Proxy-Variable verwendet werden.
	Im Fall einer binären abhängigen Variable liegt bei einer KQ Schätzung perfekte Multikollinearität im Störterm vor.
	Kann die H_0 im Durbin-Watson Test verworfen werden, schließen wir, dass Heteroskedastie vorliegt.
	Der Durbin-Watson Test unterstellt strikt exogene Regressoren.
	Interne Validität ist gegeben, wenn die Ergebnisse einer Studie auch auf andere Grundgesamtheiten übertragen werden können.
	Es ist möglich, mittels FGLS Schätzung Heteroskedastie und Autokorrelation gleichzeitig zu berücksichtigen.
	Ob eine Zeitreihe schwach abhängig ist, lässt sich nicht überprüfen.

	Ein Moving Average Prozess kann nicht schwach abhängig sein.
	Beobachtungen mit Ausprägungen der erklärenden Variablen, die weit entfernt vom Stichprobenmittelwert liegen, beeinflussen die Punktschätzer für die Steigungsparameter nicht notwendigerweise.
	In tagesgenauen Daten lassen sich keine Jahreseffekte kontrollieren.
	Bei Vorliegen von Ausreißerbeobachtungen ist das LAD Verfahren BLUE.
	Mit dem Durbin-Watson Test lässt sich auch negative Autokorrelation erster Ordnung testen.

Aufgabe 7

[15 Punkte]

Welche Antwort ist richtig? Kreuzen Sie nur **eine Antwort** pro Aufgabe an. Falls mehrere Aussagen korrekt sind, kreuzen Sie **nur** die entsprechende **Antwortkombination** an. Für jede richtige Antwort gibt es 1 Punkt. Für falsche Antworten werden keine Punkte abgezogen.

1.	Heteroskedastie	
a	<input type="checkbox"/>	Kann gleichzeitig mit Autokorrelation vorkommen.
b	<input type="checkbox"/>	Führt in Finite Distributed Lag Modellen zu inkonsistenten KQ-Schätzern.
c	<input type="checkbox"/>	Führt in dynamischen Modellen zu inkonsistenten KQ-Schätzern.
d	<input type="checkbox"/>	ist bei i.i.d Prozessen oft anzutreffen.
e	<input type="checkbox"/>	a und c
f	<input type="checkbox"/>	Keine der Antworten.

2.	Das LAD Verfahren	
a	<input type="checkbox"/>	ist bei Gültigkeit der Gauss-Markov Bedingungen effizienter als KQ.
b	<input type="checkbox"/>	unterschätzt $E(y x)$.
c	<input type="checkbox"/>	kann von Heteroskedastie betroffen sein.
d	<input type="checkbox"/>	kann von Autokorrelation betroffen sein.
e	<input type="checkbox"/>	a, b und d
f	<input type="checkbox"/>	c und d

3.	Ein RESET-Test	
a	<input type="checkbox"/>	testet auf Fehlspezifikation der funktionalen Form.
b	<input type="checkbox"/>	fügt Polynome der vorhergesagten abhängigen Variable als Regressoren in das Modell ein.
c	<input type="checkbox"/>	liefert Hinweise auf endogene Stichprobenselektion.
d	<input type="checkbox"/>	zeigt auf, welche erklärende Variable fehlspezifiziert ist.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	a und d.

4.	Der First Difference Schätzer	
a	<input type="checkbox"/>	ist heteroskedastisch.
b	<input type="checkbox"/>	ist bei Gültigkeit der Gauss-Markov-Annahmen BLUE.
c	<input type="checkbox"/>	erfordert serielle Korrelation in den zeitkonstanten Variablen.
d	<input type="checkbox"/>	ist zur Null hin verzerrt.
e	<input type="checkbox"/>	c und d .
f	<input type="checkbox"/>	keine der Antworten

5.	Ein Two-Stage-Least-Squares-Schätzer	
a	<input type="checkbox"/>	kann zur konsistenten Schätzung bei Vorliegen von Endogenität genutzt werden.
b	<input type="checkbox"/>	ist umso präziser, je stärker die Instrumente sind.
c	<input type="checkbox"/>	kann höchstens ein exogenes Instrument pro endogener Variable berücksichtigen.
d	<input type="checkbox"/>	ist für zweistufige Variablen nicht anwendbar.
e	<input type="checkbox"/>	a und b
f	<input type="checkbox"/>	a, b und c

6.	Ist bei Paneldaten die zeitkonstante, unbeobachtete Heterogenität a_i unkorreliert mit den erklärenden Variablen x_{it} [d.h., $\text{Cov}(a_i, x_{it}) = 0$], dann	
a	<input type="checkbox"/>	ist der KQ-Schätzer BLUE
b	<input type="checkbox"/>	ist der KQ Schätzer inkonsistent.
c	<input type="checkbox"/>	ist der Random Effects Schätzer effizienter als der Fixed Effects Schätzer.
d	<input type="checkbox"/>	ist der Random Effects Schätzer inkonsistent.
e	<input type="checkbox"/>	a und c.
f	<input type="checkbox"/>	c und d.

7.	Endogene Stichprobenselektion	
a	<input type="checkbox"/>	bedeutet systematische Stichprobenselektion auf Basis exogener erklärender Variablen.
b	<input type="checkbox"/>	erhöht die Stichprobengröße und senkt daher die Varianz der geschätzten Parameter.
c	<input type="checkbox"/>	entsteht bei Verwendung von Instrumentvariablenverfahren.
d	<input type="checkbox"/>	führt tendenziell zu insignifikanten Parameterschätzern.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	Keine der Antworten.

8.	Die Annahme TS.5', keine Autokorrelation der Störterme,	
a	<input type="checkbox"/>	ist im Querschnittsdatenfall erforderlich wenn die Annahme der Zufallsstichprobe gilt.
b	<input type="checkbox"/>	bedeutet, dass die unbeobachteten Faktoren verschiedener Zeitpunkte miteinander korreliert sein dürfen.
c	<input type="checkbox"/>	ist notwendig für die Gültigkeit von Inferenzverfahren nach einer KQ Schätzung.
d	<input type="checkbox"/>	ist für die Konsistenz von KQ im Zeitreihenfall unbedingt erforderlich.
e	<input type="checkbox"/>	a und c
f	<input type="checkbox"/>	Keine der genannten Antworten.

9.	Eine Proxy-Variable	
a	<input type="checkbox"/>	sollte mit der ausgelassenen Variable möglichst hoch korreliert sein.
b	<input type="checkbox"/>	sollte über die Korrelation mit der ausgelassenen Variable hinaus auch einen eigenständigen Effekt auf die abhängige Variable haben.
c	<input type="checkbox"/>	sollte möglichst hoch mit den Koeffizienten der übrigen erklärenden Variablen des Modells korreliert sein.
d	<input type="checkbox"/>	sollte mit allen erklärenden Variablen des ursprünglichen Modells möglichst hoch korreliert sein.
e	<input type="checkbox"/>	a und d.
f	<input type="checkbox"/>	b und c.

10.	Attenuation bias	
a	<input type="checkbox"/>	tritt auf bei Messfehlern in der abhängigen Variable.
b	<input type="checkbox"/>	ist eine Folge von Heteroskedastie.
c	<input type="checkbox"/>	ist eine Folge von Autokorrelation.
d	<input type="checkbox"/>	ist eine Folge von Multikollinearität.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	Keine der genannten Antworten.

11.	Im linearen Wahrscheinlichkeitsmodell	
a	<input type="checkbox"/>	ist die erklärende Variable eine Dummy Variable.
b	<input type="checkbox"/>	kann Heteroskedastie im Störterm mit dem Prais-Winston Verfahren korrigiert werden.
c	<input type="checkbox"/>	tritt zwangsläufig Heteroskedastie auf.
d	<input type="checkbox"/>	können bei einer FGLS Schätzung zur Behebung von Heteroskedastie nie alle Beobachtungen verwendet werden.
e	<input type="checkbox"/>	b und c.
f	<input type="checkbox"/>	Keine der genannten Antworten.

12.	Bei Fehlspezifikation der funktionalen Form	
a	<input type="checkbox"/>	bleiben die KQ Schätzer konsistent.
b	<input type="checkbox"/>	ist der Störterm mit den Koeffizienten der erklärenden Variablen korreliert.
c	<input type="checkbox"/>	sollte der RESET Test die H0 verwerfen.
d	<input type="checkbox"/>	kann es vorteilhaft sein, weitere Polynome der erklärenden Variablen mit aufzunehmen.
e	<input type="checkbox"/>	a, b und d.
f	<input type="checkbox"/>	c und d.

13.	Der Between Schätzer für Paneldaten mit $e_{it} = a_i + u_{it}$	
a	<input type="checkbox"/>	ist konsistent, wenn die erklärenden Variablen und a_i korreliert sind.
b	<input type="checkbox"/>	ist inkonsistent, wenn die abhängige Variable mit u_{it} korreliert ist.
c	<input type="checkbox"/>	nutzt weniger Information als der KQ Schätzer.
d	<input type="checkbox"/>	ist inkonsistent bei Heteroskedastie.
e	<input type="checkbox"/>	b und c.
f	<input type="checkbox"/>	Keine der genannten Antworten.

14.	Der LSDV Schätzer für Paneldaten mit $e_{it} = a_i + u_{it}$	
a	<input type="checkbox"/>	ist konsistent, wenn die erklärenden Variablen und a_i korreliert sind.
b	<input type="checkbox"/>	ist inkonsistent, wenn die erklärenden Variablen und u_{it} korreliert sind.
c	<input type="checkbox"/>	und der Within Schätzer für Paneldaten liefern stets identische Schätzergebnisse bzgl. der Steigungsparameter.
d	<input type="checkbox"/>	Berechnet für jede Beobachtungseinheit einen eigenen Achsenabschnitt.
e	<input type="checkbox"/>	Alle der genannten Antworten.
f	<input type="checkbox"/>	Keine der genannten Antworten.

15.	Gepoolte Querschnittsdaten	
a	<input type="checkbox"/>	liegen nur vor, wenn Querschnittsdaten aus direkt aufeinanderfolgenden Jahren vorhanden sind.
b	<input type="checkbox"/>	erlauben die Anwendung des Difference-in-Differences-Schätzers.
c	<input type="checkbox"/>	erlauben die Anwendung des Random Effects Schätzers.
d	<input type="checkbox"/>	führen bei der Anwendung von KQ stets zu autokorrelierten Störtermen.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	a, b und c.

Annahmen im linearen Regressionsmodell

Einfaches Modell

SLR.1 Das Bevölkerungsmodell ist linear in den Parametern

SLR.2 Die Stichprobe ist zufällig

SLR.3 Die Realisationen von x_i in der Stichprobe sind nicht alle identisch

SLR.4 $E(u|x) = 0$

SLR.5 $\text{Var}(u|x) = \sigma^2$

Multiples Modell

MLR.1 Das Bevölkerungsmodell ist linear in den Parametern

MLR.2 Die Stichprobe ist zufällig

MLR.3 Keine perfekte Kollinearität

MLR.4 $E(u|x_1, \dots, x_k) = E(u) = 0$

MLR.5 $\text{Var}(u|x_1, \dots, x_k) = \text{Var}(u) = \sigma^2$

MLR.6 $u \sim \text{Normal}(0, \sigma^2)$

Zeitreihenmodelle

TS.1 Der stochastische Prozess ist linear in den Parametern

TS.2 Keine perfekte Kollinearität

TS.3 $E(u_t | X) = 0$, $t=1, 2, \dots, n$

TS.4 $\text{Var}(u_t | X) = \text{Var}(u_t) = \sigma^2$, $t = 1, 2, \dots, n$

TS.5 $\text{corr}(u_t, u_s | X) = 0$, $t \neq s$

TS.6 $u_t \sim N(0, \sigma^2)$

TS.1' Der stochastische Prozess ist stationär, schwach abhängig und linear in den Parametern

TS.2' Keine perfekte Kollinearität

TS.3' $E(u_t | x_t) = 0$

TS.4' $\text{Var}(u_t | x_t) = \sigma^2$

TS.5' $E(u_s u_t | x_t, x_s) = 0$ für alle $t \neq s$.