

## Bachelorprüfung

**Fach:** Empirische Wirtschaftsforschung II

**Prüfer:** PD Dr. Richard Reichel

<b>Name, Vorname</b>	
<b>Matrikelnr.</b>	
<b>Studiengang</b>	
<b>Semester</b>	
<b>Email</b>	
<b>Datum</b>	
<b>Hausaufgabe</b>	Haben Sie im WS11/12 oder SS 11 eine Hausaufgabe abgegeben? <input type="checkbox"/> Ja <input type="checkbox"/> Nein
<b>Unterschrift</b>	

### Vorbemerkungen:

Anzahl der Aufgaben: Die Klausur besteht aus 6 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:****[10 Punkte]**

Sie möchten den Zusammenhang zwischen der Geburtenrate und der Population von Störchen in Deutschland untersuchen. Ein Zeitreihenmodell  $\log(fertil_t) = \beta_0 + \beta_1 \log(stork_t) + u_t$  mit Jahresdaten ( $n = 30$ ) liefert folgendes Ergebnis einer KQ-Schätzung:

$$\beta_0 = 0.25 \quad SE(\beta_0) = 0.07$$

$$\beta_1 = 0.992 \quad SE(\beta_1) = 0.034$$

$$R^2 = 0.92$$

$$DW = 0.36$$

- a) Interpretieren Sie das Ergebnis und zeigen Sie mögliche Probleme des Ansatzes auf. (5 Punkte)
- b) Wie könnte man obiges Modell transformieren um es mit KQ konsistent schätzen zu können? Welches Ergebnis würden Sie erwarten, wenn es sich
- b1) um einen ökonomisch begründbaren Zusammenhang zwischen beiden Variablen handeln würde, bzw. wenn es sich (2,5 Punkte)
- b2) um den Fall einer „spurious regression“ handeln würde? (2,5 Punkte)

**Aufgabe 2:****[20 Punkte]**

Regressionschätzungen können durch Heteroskedastie des Störterms gekennzeichnet sein.

- a) Welche Auswirkungen hat Heteroskedastie auf die KQ-Schätzung? (3 Punkte)
- b) Erläutern Sie den Ansatz, Heteroskedastie durch Logarithmierung zu beseitigen. (2 Punkte)
- c) Wie kann man auf Heteroskedastie testen? Vergleichen Sie den Breusch-Pagan-Test mit dem White-Test. (5 Punkte)
- d) Beschreiben Sie das WLS-Verfahren zur Schätzung bei Heteroskedastie. Worin bestehen Unterschiede zur „üblichen“ KQ-Schätzung, wo Gemeinsamkeiten? (5 Punkte)
- e) Inwiefern weicht der FGLS-Schätzer vom Verfahren unter d) ab und welche Konsequenzen hat dies? (5 Punkte)

**Aufgabe 3:****[12 Punkte]**

- a) Wann ist es sinnvoll, Instrumentvariablenschätzer einzusetzen und wodurch unterscheidet sich dieser Ansatz von einer „normalen“ KQ-Schätzung? (6 Punkte)
- b) Welche Annahmen müssen für „gute“ Instrumente getroffen werden? (3 Punkte)
- c) Welche Annahmen sind demgegenüber bei Proxy-Variablen erforderlich? (3 Punkte)

**Aufgabe 4:****[16 Punkte]**

Wenn gleichzeitig Querschnitts- und Zeitreihendaten verfügbar sind, sind Panelschätzungen eine Möglichkeit, den Informationsgehalt der Daten bestmöglich auszunutzen.

- a) Erläutern Sie den „difference-in-difference“-Schätzer und diskutieren Sie, wie man diesen Ansatz zur Abschätzung der Wirksamkeit wirtschaftspolitischer Maßnahmen verwenden könnte. (10 Punkte)
- b) Wie wird demgegenüber beim LSDV-Verfahren vorgegangen und warum ergibt sich hier meist ein hohes  $R^2$ ? (6 Punkte)

**Aufgabe 5:****[22 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.

	Damit Kovarianzstationarität vorliegt, muss die Varianz der Zufallsvariable konstant sein.
	Finite Distributed Lag Modelle weisen mindestens einen verzögerten Wert von $y$ auf.
	Sind die Regressoren eines Zeitreihenmodells nicht strikt exogen, so sind die KQ Schätzer stets inkonsistent.
	Stichprobenselektion anhand der Ausprägungen der erklärenden Variablen führt stets zu verzerrten Schätzungen.
	Beim Schätzen mit Paneldaten in ersten Differenzen werden zeitkonstante, unbeobachtete Effekte herausgerechnet.
	Sind die Störterme autokorreliert, ist der KQ-Schätzer effizient.
	Ist bei Paneldaten der Fixed Effects Schätzer inkonsistent, so gilt das stets auch für den KQ Schätzer.
	Ein Random Walk Prozess ist kovarianzstationär.
	Die erste Differenz eines Random Walk Prozesses ist schwach abhängig.
	Einheitswurzelprozesse können durch Wurzelziehen in schwach abhängige Prozesse überführt werden.
	Tritt in einem Modell $y_t = \beta_0 + \beta_1 y_{t-1} + u_t$ Autokorrelation erster Ordnung in $u_t$ auf, liefert die KQ Schätzung inkonsistente Koeffizienten.
	Das LAD Verfahren maximiert die Summe der Absolutwerte der Residuen.
	In dynamisch vollständigen Modellen sind die Störterme nicht autokorreliert.
	Liefern White-Schätzer und Prais-Winsten Schätzer stark unterschiedliche Koeffizienten, sollte das Schätzverfahren mit der niedrigeren Fehlerquadratsumme verwendet werden.
	Die FGLS Schätzungen des Cochrane-Orcutt und Prais-Winsten Verfahrens sind unverzerrt.
	Bei positiver Autokorrelation sind die von KQ ausgewiesenen Standardfehler falsch.
	Der difference-in-difference Schätzer eignet sich nur bei balanced panels.
	Eine Trendbereinigung kann auch dann Sinn machen, wenn Saisoneffekte nicht bereinigt werden.
	Das White-Schätzverfahren dient zur Korrektur der Standardfehler bei Vorliegen von Heteroskedastie.
	Newey-West t-Werte sind stets größer als die bei KQ ausgewiesenen.
	Eine binäre Variable kann als Proxy-Variable verwendet werden.
	Im Fall einer binären abhängigen Variable liegt bei einer KQ Schätzung perfekte Multikollinearität vor.
	Kann die $H_0$ im Breusch-Godfrey Test verworfen werden, schließen wir, dass Heteroskedastie

	vorliegt.
	Treten in den Daten fehlende Werte zufällig auf, führt das Weglassen der Beobachtungen mit fehlenden Werten nicht zu verzerrten KQ Schätzern.
	Die Verwendung von verzögerten endogenen Variablen auf der rechten Seite eines Zeitreihenmodells führt nicht zwangsläufig zu inkonsistenten Schätzergebnissen.
	Ausreißer werden beim LAD Verfahren stärker gewichtet als beim KQ Verfahren.
	Der Durbin-Watson Test testet auf Autokorrelation 1. Ordnung.
	Eine Dummy-Variable kann aufgrund der sich ergebenden Heteroskedastieproblematik nicht als Proxy-Variable verwendet werden.
	Externe Validität ist gegeben, wenn die Ergebnisse einer Studie auch auf andere Grundgesamtheiten übertragen werden können.
	Interne Validität setzt die unverzerrte Schätzung der Koeffizienten voraus.
	Man spricht von einem unbalanced Panel, sobald die Anzahl der Querschnittsbeobachtungen nicht der Anzahl der Wiederholungsbeobachtungen entspricht.
	Der Breusch-Pagan Test kann auch als F-Test durchgeführt werden.
	Wenn es keine zeitkonstante unbeobachtete Heterogenität gibt (alle $a_i = 0$ ), ist das First Difference Verfahren effizienter als eine gepoolte KQ-Schätzung.
	Wenn $T=2$ , liefern der Within Schätzer und der Schätzer in ersten Differenzen identische Koeffizienten und Standardfehler.
	Damit Kovarianzstationarität vorliegt, muss der Erwartungswert der Zufallsvariable konstant sein.
	Für den Prozess $u_t = 0,7u_{t-1} + e_t$ mit $e_t = \text{iid}$ . sind $u_t$ und $u_{t+h}$ ab $h=3$ unkorreliert.
	Der Durbin Watson Test lässt sich auf AR(n) Prozesse mit $n>1$ anwenden.
	Im Random Effects Modell wird unterstellt, dass die unbeobachteten zeitkonstanten Effekte mit den erklärenden Variablen unkorreliert sind.
	Der Between Schätzer kann den Effekt zeitvariabler erklärender Variablen nicht bestimmen.
	Das Fixed Effects Modell kann auch auf gepoolte Querschnittsdaten angewendet werden.
	Das Spurious Regression Problem bei Zeitreihenregressionen tritt nur dann auf, wenn abhängige und erklärende Variablen einem gleichlaufenden Trend folgen.
	Sind die Regressoren eines Zeitreihenmodells nicht kontemporär exogen, so sind die KQ Schätzer inkonsistent.
	Ist der Heteroskedastie verursachende Prozess bei einer WLS Schätzung falsch spezifiziert, ist das Verwenden von robusten Standardfehlern nach der WLS Schätzung sinnvoll.
	Im Trendmodell $\log(y_t) = \beta_0 + \beta_1 t + e_t$ kann $\beta_1$ approximativ als mittlere Wachstumsrate von $y$ interpretiert werden.

### Aufgabe 6:

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

2.	Der RESET-Test	
a	<input type="checkbox"/>	ermittelt die Höhe der Verzerrung der Koeffizienten bei Heteroskedastie.
b	<input type="checkbox"/>	Basiert auf einer Regression der abhängigen Variable auf ein Polynom der vorhergesagten abhängigen Variable.
c	<input type="checkbox"/>	testet auf Fehlspezifikation der funktionalen Form.
d	<input type="checkbox"/>	gibt Hinweise auf die Ursache einer Fehlspezifikation.
e	<input type="checkbox"/>	b und c.
f	<input type="checkbox"/>	Keine der genannten Antworten.

3.	Bei Fehlspezifikation der funktionalen Form	
a	<input type="checkbox"/>	müssen Panelschätzverfahren eingesetzt werden.
b	<input type="checkbox"/>	Ist der Störterm mit den erklärenden Variablen korreliert.
c	<input type="checkbox"/>	Sollte der RESET Test die H1 verwerfen.
d	<input type="checkbox"/>	Sollte der Breusch-Pagan Test die H0 verwerfen.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	Alle genannten Antworten.

4.	Der Between Schätzer für Paneldaten mit $e_{it} = a_i + u_{it}$	
a	<input type="checkbox"/>	ist inkonsistent, 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"/>	schätzt KQ für ein Modell querschnittsspezifischer Mittelwerte.
d	<input type="checkbox"/>	ist inkonsistent bei Heteroskedastie.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	a, b und c.

5.	Periodendummies in Fixed Effects Schätzungen	
a	<input type="checkbox"/>	erlauben die Schätzung periodenspezifischer Achsenabschnitte.
b	<input type="checkbox"/>	haben keine Within-Variation.
c	<input type="checkbox"/>	führen nur bei einem Balanced Panel zu Autokorrelation.
d	<input type="checkbox"/>	erlauben die Schätzung individuenspezifischer Achsenabschnitte.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	Keine der genannten Antworten.

6.	Folgt der Störterm einem Moving Average Prozess der Art $u_t = e_t + 0,5e_{t-1}$ und ist $\text{Var}(e_{t-1}) > 0$ ,	
a	<input type="checkbox"/>	kann der Prozess nicht mehr schwach abhängig sein.
b	<input type="checkbox"/>	sind $u_t$ und $u_{t-1}$ positiv korreliert.
c	<input type="checkbox"/>	beträgt die Durbin-Watson Teststatistik ungefähr zwei.
d	<input type="checkbox"/>	ist KQ inkonsistent.
e	<input type="checkbox"/>	a, b und d.
f	<input type="checkbox"/>	a und d.

7.	Schwache Abhängigkeit	
a	<input type="checkbox"/>	gilt für $u_t = 0,7u_{t-1} + e_t$ wenn $e_t$ iid. ist.
b	<input type="checkbox"/>	bedeutet, dass die Kovarianz zwischen den Ausprägungen einer Variable zu zwei Zeitpunkten $t$ und $t+h$ mit wachsendem $h$ gegen Null geht.
c	<input type="checkbox"/>	impliziert Kovarianzstationarität.
d	<input type="checkbox"/>	impliziert Stationarität.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	a, b und c.

8.	Bei Messfehlern	
a	<input type="checkbox"/>	in der erklärenden Variable kann der Koeffizient gegen 0 verzerrt geschätzt sein.
b	<input type="checkbox"/>	in der abhängigen Variable werden die Koeffizienten inkonsistent geschätzt.
c	<input type="checkbox"/>	in der abhängigen Variable liegt exogene Stichprobenselektion vor.
d	<input type="checkbox"/>	in der erklärenden Variable werden die $R^2$ Werte zu hoch ausgewiesen.
e	<input type="checkbox"/>	a und b.
f	<input type="checkbox"/>	Keine der genannten Antworten.

9.	Bei der Schätzung von FDL-Modellen	
a	<input type="checkbox"/>	entspricht der impact multiplier immer dem long-run multiplier.
b	<input type="checkbox"/>	tritt oftmals das Problem der Multikollinearität auf.
c	<input type="checkbox"/>	ist serielle Korrelation der Störterme immer beseitigt.
d	<input type="checkbox"/>	ist der dynamische Effekt des Regressors $x$ auf die abhängige Variable vorgegeben.
e	<input type="checkbox"/>	a und d.
f	<input type="checkbox"/>	Keine der genannten Antworten.

10.	Autokorrelation	
a	<input type="checkbox"/>	kann nicht gleichzeitig mit Heteroskedastie vorkommen.
b	<input type="checkbox"/>	kennzeichnet die Störterme von dynamisch vollständigen Modellen.
c	<input type="checkbox"/>	Führt in statischen Modellen zu inkonsistenten KQ-Schätzern.
d	<input type="checkbox"/>	ist bei random walk Prozessen gleich 0.
e	<input type="checkbox"/>	a und c.
f	<input type="checkbox"/>	Keine der Antworten.

## 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  $t \neq s$ .