

Aufgabe 1:

[32 Punkte]

Sie interessieren sich für Wohnungslosigkeit und ihre Determinanten. Sie möchten eine Studie US-amerikanischer Ökonomen auf Deutschland übertragen und sammeln hierfür in mehreren Gemeinden Daten über die Zahl der Wohnungslosen sowie Merkmale des lokalen Wohnungsmarktes. Sie unterstellen folgendes Modell (A):

$$WL_i = \beta_0 + \beta_1 \cdot Leer_i + \beta_2 \cdot Mietek_i + \beta_3 \cdot JanTemp_i + \beta_4 \cdot ALQ_i + \varepsilon_i$$

WL_i : Anteil der Wohnungslosen an der Gesamtbevölkerung in der Gemeinde i

$Leer_i$: Anteil der leer stehenden Wohnungen an der Gesamtzahl der Wohnungen in der Gemeinde i

$Mietek_i$: Medianmiete im Verhältnis zum Median des Haushaltseinkommens in der Gemeinde i

$JanTemp_i$: Durchschnittstemperatur in der Gemeinde i im Januar (in °C)

ALQ_i : Anteil der Arbeitslosen an der Erwerbsbevölkerung in der Gemeinde i

Die Mittelwerte der Variablen betragen:

	WL_i	$Leer_i$	$Mietek_i$	$JanTemp_i$	ALQ_i
Mittelwert	0.03	0.12	0.29	-1.8	0.12

Die Auswertung der Daten mit R ergibt folgenden Output:

```
Call:
lm(formula = WL ~ Leer + Mietek + JanTemp + ALQ)

Coefficients:
            Estimate      Std. Error  t value Pr(>|t|)
(Intercept)   0.434         0.219    1.982  0.0477 *
Leer          -0.668         0.057   -11.632 6.55e-30 ***
Mietek         0.352          ?     1.751  0.0802 .
JanTemp        0.051         0.016     ?    0.0015 **
ALQ            0.004         0.098     0.041  0.9673
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.46 on 1390 degrees of freedom
Multiple R-Squared:  0.973,    Adjusted R-squared:  0.973
F-statistic: ? on ? and 1390 DF,  p-value: <2.2e-16
```

a) Berechnen Sie unter Angabe des Rechenweg

(5 Punkte)

a1) den t-Wert für b_3

- $t_{b_3} = \frac{b_3}{se(b_3)} = \frac{0.051}{0.016} = 3.188$

a2) den Standardfehler von b_2

- $se(b_2) = \frac{b_2}{t_{b_2}} = \frac{0.352}{1.751} = 0.201$

a3) die geschätzte Fehlertermvarianz

- $\hat{V}(e_i) = s^2 = 0,46^2 = 0.2116$

a4) die Zählerfreiheitsgrade der F-Statistik

- Erklärende ohne Konstante: 4

a5) den Wert der F -Statistik

$$F = \frac{R^2/(K-1)}{(1-R^2)/(N-K)} = \frac{0.973/4}{0.027/1390} = 12522.87$$

b) Interpretieren Sie die Koeffizienten b_1 und b_4 inhaltlich und statistisch. (2 Punkte)

- b_1 : ist auf 1%-Niveau signifikant verschieden von Null. Wenn sich der Anteil leer stehender Wohnungen an allen Wohnungen um einen Prozentpunkt erhöht, sinkt der Anteil der Wohnungslosen an der Gesamtbevölkerung um 0.668 Prozentpunkte.
- b_4 : der p -Wert beträgt 0.96, d.h. die $H_0: \beta_4=0$ könnte nur abgelehnt werden, wenn eine Fehlerwahrscheinlichkeit von 96% akzeptiert wird. Der Koeffizient besagt, dass der Anteil der Wohnungslosen um 0.004 Prozentpunkte steigt, wenn die Arbeitslosigkeit um einen Prozentpunkt zunimmt.

c) In einer weiteren Schätzung wird dem Modell (A) ein Interaktionsterm $Int = Leer \cdot Mietek$ hinzugefügt. Der Output des resultierenden Modells (B) lautet wie folgt: (3 Punkte)

```
Call:
lm(formula = WL ~ Leer + Mietek + JanTemp + ALQ + Int)

Coefficients:
(Intercept)      0.423      0.239      1.769      0.0771
Leer            -0.768      0.069     -11.130     1.27e-27 ***
Mietek           0.288      0.126       2.286      0.0224 *
JanTemp          0.031      0.014       2.214      0.0269 *
ALQ              -0.002      0.096      -0.021      0.9832
Int              0.346      0.149       2.322      0.0204 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.37 on 1389 degrees of freedom
Multiple R-Squared:  0.983,    Adjusted R-squared:  0.983
F-statistic: 1.338e+04 on 5 and 1389 DF,  p-value: <2.2e-16
```

c1) Berechnen Sie den marginalen Effekt einer Veränderung des Leerstandes für Gemeinden, in denen die Medianmiete einen Anteil von 0,4 des Medianeinkommens ausmacht.

$$\frac{\partial WL}{\partial Leer} = b_1 + b_5 \cdot Mietek = -0.768 + 0.346 \cdot 0.4 = -0.6296$$

c2) Interpretieren Sie den Koeffizienten b_1 aus Modell (B).

- $b_{1(B)}$ misst den marginalen Effekt des Leerstandes auf das Ausmaß der Wohnungslosigkeit bei $Mietek = 0$, also wenn Wohnraum kostenlos wäre.

d) Modellspezifikation

d1) Beschreiben Sie die Vorgehensweise des RESET-Tests und stellen Sie eine mögliche Hilfsregression für das Modell (A) auf. Nehmen Sie an, die resultierende Teststatistik habe einen Wert von 2.48. Testen Sie für die gewählte Hilfsregression auf dem 5%-Signifikanzniveau, ob eine Fehlspezifikation vorliegt. (5 Punkte)

- Aus einer ersten Schätzung werden vorhergesagte Werte für y berechnet: \hat{y}_i
- Diese vorhergesagten Werte werden mit 2, 3, 4, ... potenziert und dann in einer neuen Regression als Erklärende eingefügt.
- Testidee: Haben die vorhergesagten Werte noch einen weiteren Erklärungsgehalt, kann noch nicht alles mit X verbundener Erklärungsgehalt ausgeschöpft sein.
- Daher: Sind die Koeffizienten verschieden von Null, ist das ein Hinweis auf Fehlspezifikation.

- Hier Hilfsregression für $Q=3$:

$$WL_i = b_0 + b_1 \cdot Leer_i + b_2 \cdot Mietek_i + b_3 \cdot JanTemp_i + b_4 \cdot ALQ_i + \gamma_1 \hat{Ob}_i^2 + \gamma_2 \hat{Ob}_i^3 + e_i$$

- $H_0: \gamma_1 = \gamma_2 = 0$
- kritischer Wert: $F_{1389}^2 \approx F_{\infty}^2 = 3.00$
- Teststatistik: $2.48 < 3 =$ kritischer Wert
- H_0 kann nicht verworfen werden, es besteht daher kein Hinweis auf Fehlspezifikation.

d2) Nehmen Sie an, Modell(A): $WL_i = x_i' \beta + \varepsilon_i$ werde geschätzt, obwohl Modell (B)

$WL_i = x_i' \beta + \beta_5 \cdot Int_i + \varepsilon_i$ korrekt sei, wobei $x_i' = (1 \text{ Leer Mietek JanTemp ALQ})$. Zeigen Sie, dass der KQ-Schätzer $b_{(A)}$ verzerrt ist, falls $E\{x_i Int_i\} \neq 0$. (4 Punkte)

- $y_i = x_i' \beta + Int_i \beta_5 + \varepsilon_i$

- $$b_A = \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i y_i = \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i (x_i' \beta + Int_i \beta_5 + \varepsilon_i)$$

$$= \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i x_i' \beta$$

$$+ \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i Int_i \beta_5$$

$$+ \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i \varepsilon_i$$

$$= \beta + \beta_5 \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i Int_i + \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i \varepsilon_i$$

- $$E(b_A) = E \left(\beta + \beta_5 \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i Int_i + \left(\sum_i x_i x_i' \right)^{-1} \sum_i x_i \varepsilon_i \right)$$

$$= \beta + \beta_5 \left(\sum_i x_i x_i' \right)^{-1} \cdot \sum_i E(x_i Int_i) + 0$$

- Nur wenn $E\{x_i Int_i\} = 0$ ist $E(b_A) = \beta$.

d3) Sie schätzen zusätzlich ein Modell, in dem Sie abhängige und erklärende Variablen zuvor logarithmiert haben. Mit welchem Test können Sie überprüfen, welches Modell geeigneter ist? Stellen Sie eine der beiden für diesen Test möglichen Schätzgleichungen auf und nennen Sie die Nullhypothese. Erläutern Sie *kurz* das zentrale Element der Schätzgleichung. (3 Punkte)

- $WL_i = x_i' \beta + \delta_{LN} (\log(\hat{WL}_i) - \log \tilde{WL}_i) + u_i$

- $\log(\hat{WL}_i)$ ist der Logarithmus des aus einem linearen Modell vorhergesagten Wertes, $\log \tilde{WL}_i$ der direkt aus dem logarithmierten Modell vorhergesagte Wert.

- $H_0 : \delta_{LIN} = 0$

e) Sie haben die Vermutung, dass die Fehlertermvarianz mit steigenden Temperaturen variiert:
 $\text{var}(\varepsilon_i) = \sigma_i^2 = \sigma^2 \cdot \text{JanTemp}_i$.

e1) Beschreiben Sie eine Vorgehensweise, Modell A so zu transformieren, dass die Heteroskedastie eliminiert wird. Stellen Sie formal die Auswirkung der Transformation auf die Varianz des Störterms dar. (3 Punkte)

- Für das GLS-Verfahren wird jede Variable aus dem Ausgangsmodell mit $1/\sqrt{\text{JanTemp}_i}$ als Gewichtungsfaktor h_i multipliziert.

-
$$\frac{W_i}{\sqrt{\text{JanTemp}_i}} = \frac{b_0}{\sqrt{\text{JanTemp}_i}} + \dots + \underbrace{\frac{e_i}{\sqrt{\text{JanTemp}_i}}}_{u_i}$$

- Der neue Störterm ist homoskedastisch, da $\frac{e_i}{\sqrt{\text{JanTemp}_i}} = u_i \sim N(0, \sigma^2)$, wenn $\varepsilon_i \sim N(0, \sigma^2 \cdot \text{JanTemp}_i)$.

e2) Zeigen Sie, dass im Fall von Heteroskedastie $V\{b|X\} \neq \sigma^2(X'X)^{-1}$. (4 Punkte)

- $V\{b|X\} = V\{(X'X)^{-1} X'y|X\} = V\{(X'X)^{-1} X'\varepsilon|X\} = (X'X)^{-1} X'V\{\varepsilon|X\} X(X'X)^{-1}$

- Nur wenn $V\{\varepsilon|X\} = \sigma^2 \cdot I$ lässt sich das zu $V\{b|X\} = \sigma^2(X'X)^{-1}$ kürzen, das ist bei Heteroskedastie aber nicht gegeben.

e3) Welche Vor- und Nachteile hat die Berechnung von White-Standardfehlern gegenüber dem FGLS-Verfahren? (3 Punkte)

- Vorteile:

1) Es ist einfacher: Schätzwerte aus normalem KQ können verwendet werden;

2) Es muss keine genaue Struktur der Heteroskedastie unterstellt werden → Potenzielle Fehlerquellen werden vermieden.

- Nachteil: nicht BLUE / nicht effizient.

Aufgabe 2:

[8 Punkte]

Der Bürgermeister einer Großstadt steht vor der Aufgabe, Maßnahmen zur Kriminalitätsbekämpfung zu ergreifen. Um Kosten und Nutzen verschiedener Maßnahmen gegeneinander abwägen zu können, interessiert er sich für den marginalen Effekt einer Aufstockung des Polizeipersonals auf die Kriminalitätsrate.

a) Er schätzt mittels der KQ-Methode ein lineares Modell, das die erwartete Kriminalitätsrate in Abhängigkeit der Zahl der beschäftigten PolizistInnen und einiger anderer Variablen x beschreibt: (3 P)

$$\text{Kriminalitätsrate}_i = \beta_0 + \beta_1 \cdot \text{PolizistInnen}_i + x_i \gamma + \varepsilon_i$$

a1) Könnte in der Beispielschätzung ein Endogenitätsproblem auftreten? Begründen Sie inhaltlich am Beispiel.

- Mögliche Antworten:

- Realisierte Kriminalitätsrate und Zahl der PolizistInnen werden gemeinsam determiniert durch eine unbeobachtbare „kriminelle Neigung“ in der Bevölkerung, oder

- Es ist unklar, ob Zahl der PolizistInnen auf die Kriminalität wirkt oder Kriminalität auf die Zahl der PolizistInnen.

a2) Was wären gegebenenfalls die Konsequenzen für den resultierenden Schätzer b ?

- Der Schätzer wäre verzerrt und inkonsistent.

b) Dem Bürgermeister fällt ein, dass der Stadtrat in der Vergangenheit oft uneins war, ob mehr Personal für die Polizei oder mehr Personal in den Kindergärten eingestellt werden sollte. Aus diesem Grund wurden immer anteilig PolizistInnen und KindergärtnerInnen eingestellt, sobald aus dem städtischen Haushalt noch Gelder zu verteilen waren. Er instrumentiert daher die Zahl der PolizistInnen mit der Zahl der KindergärtnerInnen. (5 Punkte)

b1) Hat er eine gute Wahl getroffen? Begründen Sie.

- Zahl der KindergärtnerInnen ist ein gutes Instrument, da die Anforderungen erfüllt sind:

- $cov(x_i, z_i)$ möglichst hoch \rightarrow Dies ist nach Behauptung aus Aufgabenstellung erfüllt.

- $E(\varepsilon_i z_i) = 0$, d.h. kein Endogenitätsproblem beim Instrument \rightarrow Es ist unwahrscheinlich, dass die Kriminalitätsrate auf das Kindergartenpersonal wirkt oder beide Größen gemeinsam durch etwas Drittes determiniert werden. (Oder plausible andere Einschätzungen)

b2) Anhand welcher Größe ließe sich die Qualität des Instruments empirisch prüfen? (2 Punkte)

- Der Zusammenhang mit der zu instrumentierenden Variable ($cov(x_i, z_i)$ soll hoch sein) lässt sich mit dem R^2 einer Hilfsregression von x auf z prüfen.

Aufgabe 3:

[15 Punkte]

Sie interessieren sich für Einflussfaktoren auf die Anzahl von Verkehrsunfällen mit Todesfolge. Ihnen stehen für den Zeitraum von 1992 bis 1998 Daten für 48 US Bundesstaaten mit folgenden Variablen zur Verfügung:

<i>fatal</i>	Anzahl von Verkehrsunfällen mit Todesfolge (Tote pro 100,000 Einwohner)
<i>beertax</i>	Biersteuer (in Prozent)
<i>spircons</i>	Konsum von Spirituosen (in Gramm Alkohol)
<i>unrate</i>	Arbeitslosenquote (in Prozent)
<i>pcinc</i>	Pro-Kopf-Einkommen (in 1,000 US-\$)

Die Mittelwerte der Variablen betragen:

	<i>fatal</i>	<i>beertax</i>	<i>spircons</i>	<i>unrate</i>	<i>pcinc</i>
Mittelwert	2.04	0.51	1.75	7.35	13.88

Bei Anwendung unterschiedlicher Verfahren ergeben sich folgende Schätzergebnisse:

Abhängige Variable: $\log(\text{fatal})$				
	gepooltes KQ	between	fixed effects	random effects
constant	4.119*** (0.30)	3.796*** (0.75)	-0.384 (0.42)	2.002*** (0.38)
beertax	0.097 (0.062)	0.0740 (0.15)	-0.048 (0.16)	0.443*** (0.12)
spircons	0.162*** (0.043)	0.300** (0.11)	0.817*** (0.079)	0.302*** (0.064)
unrate	-0.029** (0.013)	0.032 (0.038)	-0.029*** (0.009)	-0.049*** (0.009)
pcinc	-0.158*** (0.017)	-0.184*** (0.042)	0.105*** (0.021)	-0.011 (0.019)

Anmerkungen: Standardfehler in Klammern; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

a) Erläutern Sie kurz verbal, wieso bei der gepoolten KQ-Schätzung das Problem ausgelassener Variablen auftritt. (2 Punkte)

- Modell $y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}$; wenn beobachtungsspezifische Effekte α_i mit den sonstigen im Modell enthaltenen Variablen korreliert sind, schlagen sich die α_i in diesen nieder: ‚omitted variable bias‘.
- b) Interpretieren Sie die Koeffizienten des Parameters *beertax* für die fixed effects und random effects Spezifikation inhaltlich und statistisch. (2 Punkte)**
 - Steigt die Biersteuer um eine Einheit (ein Prozent), so deutet der Koeffizient aus der FE Schätzung auf einen Rückgang der tödlichen Verkehrsunfälle um mehr als 48 Prozent ($e^{0.48} - 1$) hin, allerdings ist der Koeffizient nicht statistisch signifikant.
 - Der Koeffizient aus der RE Schätzung hingegen weist auf Erhöhung der Anzahl tödlicher Verkehrsunfälle um ca. 4% hin. Der Koeffizient ist zudem auf dem 1%-Signifikanzniveau statistisch signifikant.
- c) Erläutern Sie, warum sich die Schätzergebnisse für einerseits between und fixed effects Schätzer und andererseits random effects und fixed effects Schätzer teils erheblich unterscheiden können. (6 Punkte)**
 - Unterschied between – fixed effects:
Der between-Schätzer verwendet gemittelte Daten über alle Beobachtungen i , der fixed effects-Schätzer hingegen verwendet nur Abweichungen vom individuellen Mittelwert; bei unterschiedlichen Daten können sich unterschiedliche Schätzergebnisse einstellen.
 - Unterschied random effects – fixed effects:
„Natur“ des individuenspezifischen Effekts: fixer Effekt (Dummy pro Beobachtungseinheit, Niveauunterschied zwischen Beobachtungen) vs. normalverteilter Bestandteil des Fehlerterms.
Korrelation von α_i mit Kovariaten x_{it} : zulässig bei fixed effects Schätzer; nicht zulässig bei random effects Schätzer. Konsequenz: Wenn α_i mit Kovariaten x_{it} korreliert, ist der random effects Schätzer verzerrt und inkonsistent.
- d) Skizzieren Sie die Idee des Hausman Tests und testen Sie auf dem 1% Signifikanzniveau, ob die Ergebnisse der random effects oder der fixed effects Schätzung vorzuziehen sind; die Teststatistik beträgt 20.28 bei 4 Freiheitsgraden. Geben Sie hierzu den kritischen Wert und Ihre Schlusslogik an. (5 Punkte)**
 - $H_0: \text{cov}(\alpha_i, x_{it}) = 0$; $H_1: \text{cov}(\alpha_i, x_{it}) \neq 0$
 - $\hat{\beta}_{RE}$ ist nur dann konsistent und effizient, wenn α_i und x_{it} unkorreliert sind, $\hat{\beta}_{FE}$ ist unter der Null- und Alternativhypothese konsistent
 - Test, ob Differenz der beiden Schätzer signifikant: $H_0: \text{plim}(\hat{\beta}_{FE} - \hat{\beta}_{RE}) = 0$; wenn signifikant, ist der fixed effects Schätzer heranzuziehen.
 - $\chi_{(df=4, 1\%)} = 13.28 < 20.48$
 - H_0 wird verworfen: Ergebnisse des fixed effects Schätzers sind heranzuziehen.

Aufgabe 4:

[5 Punkte]

Unterstellen Sie ein lineares Modell $y_t = x'_t\beta + \varepsilon_t$.

- a) Sie haben Anlass zur Vermutung, dass in Ihren Daten Autokorrelation 3. Ordnung vorliegt, $\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \rho_3\varepsilon_{t-3} + v_t$, mit $v_t \sim i.i.d.(0, \sigma^2)$. Geben Sie eine Spezifikation des Modells an, bei der Autokorrelation beseitigt ist. (2 Punkte)**

- Transformation des Modells durch Aufnehmen verzögerter abhängiger und unabhängiger Variablen:

$$y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} - \rho_3 y_{t-3} = (x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2} - \rho_3 x_{t-3})' \beta + u_t$$

- b) Erläutern Sie, wie sich die Varianz-Kovarianz-Matrix des Störterms mit AR(1) Fehlern von der des Modells mit MA(1) Fehlern unterscheidet. (3 Punkte)**

- Bei AR(1) Fehlern gibt es keine Elemente der Matrix, die gleich Null sind.
- Bei MA(1) Fehlern sind alle Elemente der Matrix gleich Null, die nicht auf oder neben der Hauptdiagonalen liegen.
- (graphische Darstellung ebenfalls gültig)

Aufgabe 5:

[45 Punkte]

Wahr oder falsch? ...

W	Ein niedriger AIC-Wert weist auf ein besseres Regressionsmodell hin als ein hoher.
W	Der White-Test ist allgemeiner als der Test auf Gleichheit zweier Varianzen (Goldfeldt-Quandt).
F	Bei Messfehlern beschreibt das „noise-to-signal ratio“ das Verhältnis der Varianz des tatsächlichen Wertes zur Varianz des Messfehlers.
W	Der RESET Test nutzt Potenzen der vorhergesagten Werte von y, um ein Modell auf Fehlspezifikation zu überprüfen.
W	Wald Tests können in Form von χ^2 -Tests durchgeführt werden.
F	In ein Modell mit logarithmierter abhängiger Variable können keine Dummy-Variablen als Erklärende eingefügt werden, da $\ln(0) = -\infty$.
W	Unterschiede in Koeffizienten für verschiedene Gruppen lassen sich mittels Interaktionstermen ermitteln.
W	Solange $E(\varepsilon X) = 0$ geben Regressionsmodelle den auf X bedingten Erwartungswert von y an.
W	Wenn irrelevante erklärende Variablen berücksichtigt werden, steigt die Varianz des KQ-Schätzers.
W	Autokorrelation führt nicht zu Inkonsistenz des KQ-Schätzers.
W	Für das Ergebnis eines Chow-Tests ist es unerheblich, ob man für die fraglichen Gruppen getrennt schätzt oder Unterschiede über Interaktionsterme abbildet.
F	Wenn ein Regressionsmodell keine Konstante enthält, ist das angepasste R^2 vor dem einfachen R^2 als Gütemaß zu bevorzugen.
F	Die gemeinsame Signifikanz mehrerer Steigungsparameter lässt sich durch genestete t-Tests überprüfen.
W	Im Rahmen einer GIVE-Schätzung mit mehreren endogenen Variablen und mehreren Instrumenten ist es egal, welcher endogenen Variablen welches Instrument zugeordnet wird.
F	Projiziert man den Vektor y auf den Spaltenraum X, so ist das resultierende Residuum orthogonal zu y.
W	Die Wahrscheinlichkeit eines Typ-II-Fehlers hängt vom wahren Parameterwert ab.
F	Wenn $E(\varepsilon X) = E(\varepsilon) = 0$ und $V(\varepsilon X) = V(\varepsilon) = \sigma^2 \cdot I$ liefert der KQ-Schätzer den wahren Wert.

Musterlösung zur Diplomprüfung Ökonometrie im WS 06/07

W	Der gepoolte KQ-Schätzer nutzt im Panel sowohl die Information der Within-Variation als auch der Between-Variation.
F	Im einfachen Regressionsmodell gilt: Der Vorhersagefehler für eine bestimmte Beobachtung wird umso größer, je weiter die Beobachtung vom Mittelwert der erklärenden Variablen entfernt liegt.
W	Autokorrelation führt nicht zu Verzerrtheit des KQ-Schätzers.
F	Das Auslassen relevanter erklärender Variablen verzerrt die im Modell enthaltenen Koeffizienten nur dann, wenn die ausgelassene Variable mit dem Störterm korreliert ist.
W	Messfehler in erklärenden Variablen führen im einfachen Modell dazu, dass der absolute Wert des geschätzten Steigungsparameters im Erwartungswert kleiner ist als der wahre Wert.
F	Wenn es mehr Instrumente als zu schätzende Parameter gibt, ist das Modell überidentifiziert und kann daher nicht gelöst werden.
W	Die Parameter eines FGLS-Modells müssen wie die Parameter des Original-Modells interpretiert werden.
F	Die FGLS-Schätzung bei heteroskedastischen Störtermen beruht darauf, dass Beobachtungen mit kleiner Varianz im Störterm ein kleineres Gewicht erhalten als Beobachtungen mit großer Varianz im Störterm.
W	Die Varianz eines unverzerrten Schätzers kann höher sein als die Varianz eines inkonsistenten Schätzers.
F	Ein J-Test ist nur für genestete Modelle durchführbar.
F	Beim Random-Effects-Verfahren sind die personenspezifischen Störterme Zufallsvariablen und können daher auch über die Zeit variieren.
F	Im Falle von Heteroskedastie sind die Störterme untereinander korreliert.
W	Nur in einem Modell mit Konstante beschreibt das R^2 den Anteil der erklärten Variation der abhängigen Variable an der gesamten Variation der abhängigen Variable.
F	An der Stichprobenmitte sagt ein KQ-Schätzer den wahren Wert vorher.
W	Für den Durbin-Wu-Hausman-Test auf Endogenität regressiert man die potenziell endogene Variable auf das Instrument.
F	Mit einem in den Parametern linearen Modell lassen sich keine Elastizitäten berechnen.
W	Der RE-Schätzer kann konsistent sein, wenn gleichzeitig auch der FE-Schätzer konsistent ist.
F	Im einfachen Regressionsmodell entspricht der Schätzer des Steigungsparameters dem Verhältnis der Stichprobenkovarianz von x und y zur Stichprobenvarianz von y .
F	Für die Vorhersage von y_i spielt es keine Rolle, ob das Modell linear oder loglinear, geschätzt wird.
W	Im FGLS-Verfahren zur Korrektur von Heteroskedastie wird das transformierte Modell ohne eigentliche Konstante geschätzt.
W	Vom „least squares dummy variables“ Schätzer spricht man, wenn bei Vorliegen von Paneldaten Dummy-Variablen für Beobachtungseinheiten in das Modell aufgenommen werden.
W	Modelle in reduzierter Form enthalten ausschließlich exogene erklärende Variablen.
W	Der p -Wert gibt das kleinstmögliche Signifikanzniveau an, auf dem die Nullhypothese verworfen werden kann.

Musterlösung zur Diplomprüfung Ökonometrie im WS 06/07

W	Autokorrelation kann durch das Auslassen relevanter erklärender Größen verursacht werden.
F	Sind in Paneldatenmodellen die α_i mit den erklärenden Größen korreliert, so ist der fixed effects Schätzer inkonsistent.
F	Bedingung für ein gutes Instrument ist, dass die Instrumentvariable z und die endogene Variable x unkorreliert sind.
F	Das R^2 ist in Modellen mit logarithmierten abhängigen Variablen größer als bei Modellen mit nicht-logarithmierten abhängigen Variablen.
W	Newey-West Standardfehler stellen eine Erweiterung von White-Standardfehlern auf den Fall der Autokorrelation dar.
F	Bei Multikollinearität führen positiv korrelierte erklärende Variablen zu positiv korrelierten Koeffizienten.
W	Der random effects GLS-Schätzer ist das matrix-gewichtete Mittel aus between- und within-Schätzer.
F	Bei moving average Prozessen sind alle Elemente der Varianz-Kovarianz Matrix des Störterms von Null verschieden.
F	Bedingung für einen asymptotisch normalverteilten KQ-Schätzer sind normalverteilte Störterme.
W	Solange gilt $E\{x_t \varepsilon_t\} = 0$ (mit ε_t als Residuum der KQ-Schätzung), ist der KQ-Schätzer bei Modellen, die eine verzögerte abhängige Variable als erklärende Größe beinhalten, konsistent.
W	GMM-Schätzer des linearen Modells treffen keine Annahme hinsichtlich Autokorrelation.
F	Der PE-Test wird herangezogen, um Strukturbrüche in Zeitreihen zu ermitteln.
F	Bei Monte-Carlo Studien greift man auf repräsentative Erhebungen zurück, um Eigenschaften von Schätzern zu untersuchen.
W	Der Breusch-Pagan Test basiert auf einer Regression der quadrierten geschätzten KQ-Residuen auf eine Teilmenge z der erklärenden Größen x .
W	Man spricht von einem stationären autoregressiven Prozess erster Ordnung, wenn gilt $ \rho < 1$.
F	Liegen mehrere endogene Regressoren vor, so benötigt der IV-Schätzer zur Identifikation zwei Instrumente pro endogener Variable.
F	Das GMM-Verfahren schätzt Parameter direkt auf Basis von nichtlinearen Momentenbedingungen.
F	Der Prais-Winsten-Schätzer basiert auf transformierten Daten ohne Berücksichtigung der ersten Beobachtung.
W	Den FGLS-Schätzer kann man bestimmen, wenn man eine konkrete Form der Heteroskedastie unterstellt.
W	Sind Polynome der erklärenden Variablen x im Modell enthalten, so ergibt sich der marginale Effekt der Variablen x als Ableitung des auf x bedingten Erwartungswerts von y nach x .

Aufgabe 6:

[15 Punkte]

Wahr oder falsch? ...

W	Der marginale Effekt erklärender Variablen kann auch von anderen erklärenden Variablen abhängen. → wenn Variablen mit anderen Variablen interagiert sind.
----------	--------------------------------------------------------------------------------------------------------------------------------------------------------------

Lehrstuhl für Statistik und emp. Wirtschaftsforschung, Prof. Regina T. Riphahn, Ph.D.
Musterlösung zur Diplomprüfung Ökonometrie im WS 06/07

F	<p>Bei Vorliegen von Paneldaten sind gepoolte KQ-Schätzungen effizient.</p> <p>→ wenn α_i Bestandteil des Störterms, dann Autokorrelation und ineffiziente KQ-Ergebnisse; GLS-Schätzer wäre effizienter.</p>
W	<p>Das angepasste R^2 ist zum Vergleich genesteter Modelle besser geeignet als das nicht angepasste R^2.</p> <p>→ Korrektur der Freiheitsgrade um die zusätzlich ins Modell aufgenommenen Variablen.</p>
W	<p>In Stichproben, die aus zwei Gruppen mit jeweils homoskedastischen Fehlertermvarianzen bestehen, kann Heteroskedastie ein Problem darstellen.</p> <p>→ wenn die Gruppen unterschiedliche Fehlertermvarianzen aufweisen.</p>
F	<p>In Modellen mit verzögerten abhängigen Variablen als Regressor sind random effects Schätzer konsistent.</p> <p>→ da y_{i-1} mit α_i korreliert.</p>
F	<p>Der KQ-Schätzer ist nur dann konsistent, wenn die Gauss-Markov Bedingungen erfüllt sind.</p> <p>→ auch bei nicht-sphärischen Störtermen (Heteroskedastie, AR) konsistent; → nur A6 und A7 müssen erfüllt sein.</p>
F	<p>Das Modell $g(x_i, \beta) = \beta_1 x_{i1}^{\beta_2} x_{i2}^{\beta_3}$ lässt sich nicht per KQ schätzen.</p> <p>→ Linearisierung der Parameter über Logarithmierung möglich.</p>
F	<p>In Modellen mit logarithmierter abhängiger Variable y lassen sich prozentuale Veränderungen in y bei hohen Absolutwerten der geschätzten Koeffizienten direkt aus diesen ablesen.</p> <p>→ nur bei kleinen Werten gilt approximativ $e^\beta - 1 \approx \beta$</p>
W	<p>Beim Durbin-Watson Test auf negative Autokorrelation können bei gegebenem Signifikanzniveau die gleichen Tabellen wie beim Test auf positive Autokorrelation herangezogen werden.</p> <p>→ kritischer Wert liegt zwischen $4 - d_U$ und $4 - d_L$.</p>
W	<p>In simultanen Gleichungssystemen kann der GIVE-Schätzer auch als zweistufiger Schätzer ermittelt werden.</p> <p>→ 1) Regressieren jeder endogenen Variable auf Instrumente; 2) Verwenden der vorhergesagten Werte aus 1) anstelle der endogenen Variablen in KQ-Schätzung der Strukturgleichungen.</p>