Aufgabe 1: **[25 Punkte]**

Ihnen liegen 28155 Beobachtungen aus dem Current Population Survey aus dem Jahr 1988 vor. Unter anderem sind folgende Variablen enthalten:

wage	**Lohn in US-Dollar,**
exper	**Berufserfahrung in Jahren,**
educ	**Schulbildung in Jahren,**
nonwhite	**Dummyvariable (0 = Weiße(r) und 1 = andere ethnische Zugehörigkeit) ,**
male	**Dummyvariable (0 = Frau und 1 = Mann).**

Sie interessieren sich für die Determinanten des Lohnes einer Person *i* und schätzen hierzu folgendes Modell:

$$log(wage_i) = \beta_0 + \beta_1 exper_i + \beta_2 exper_i^2 + \beta_3 educ_i + \beta_4 nonwhite_i + \beta_5 male_i + \varepsilon_i .$$

Die Auswertung der Stichprobe ergibt folgenden Output:

```
Call:
lm(formula = log(wage) ~ exper + I(exper^2) + educ + nonwhite + male)

Coefficients:
                     Estimate Std.Err. t value  Pr(>|t|)
(Intercept)  4.321   0.019    225.38   < 2e-16
exper        0.077   0.0088    8.750   < 2e-16
I(exper^2)  -0.001   0.00001  -69.31   < 2e-16
educ          ???    0.001     67.34   < 2e-16
nonwhite    -0.243    ???     -18.84   < 2e-16
male         0.184   0.087     2.115     0.017
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5839 on 28149 degrees of freedom
Multiple R-Squared: 0.3347,    Adjusted R-squared: ???
F-statistic: ??? on 5 and 28149 DF,  p-value: < 2.2e-16
```

a) Berechnen Sie unter Angabe des Rechenwegs **(6 Punkte)**

a1) den geschätzten Koeffizienten für β_3 ;

- $b_3 = t_{b_3} \cdot se(b_3) = 67.34 \cdot 0.001 = 0.06734$

a2) den Standardfehler für b_4 ;

- $se(b_4) = \dfrac{b_4}{t_{b_4}} = \dfrac{-0.243}{-18.84} = 0.0129$

a3) ein 95% Konfidenzintervall für b_1 ;

- $[b_1 \pm 1.96 \cdot se(b_1)] = [0.077 \pm 1.96 \cdot 0.0088] = [0.0597 ; 0.0942]$

a4) das korrigierte R^2 ;

- $\bar{R}^2 = 1 - \dfrac{N-1}{N-K}(1-R^2) = 1 - \dfrac{28155-1}{28155-6}(1-0.3347) = 0.3346$

a5) die F-Statistik des Gesamtmodells.

- $F = \frac{R^2/(K-1)}{(1-R^2)/(N-K)} = \frac{0.3347/(6-1)}{(1-0.3347)/(28155-6)} = 2832.247$

b) Sie schätzen ein weiteres Modell ohne *exper* und *exper*^{2}, das R^2 betrage 0.2953. Führen Sie unter Angabe des Rechenwegs auf dem 5%-Niveau einen F-Test auf gemeinsame Signifikanz der beiden Koeffizienten der Berufserfahrung durch. Geben Sie hierzu auch die Null- und die Alternativhypothese sowie den kritischen Wert an. (4 Punkte)

- $F = \frac{(R_1^2 - R_0^2)/J}{(1-R_1^2)/(N-K)} = \frac{(0.3347-0.2953)/2}{(1-0.3347)/(28155-6)} = 833.5116$
- $H_0: \ \beta_{exper} = \beta_{exper^2} = 0$
- H_1: mindestens einer der beiden geschätzten Koeffizienten $\neq 0$
- $F_{krit} = F_{2;28151} = 3.00$
- 833.5116 > 3.00, H_0 wird abgelehnt.

c) Berechnen Sie den marginalen Effekt eines zusätzlichen Jahres Berufserfahrung am Stichprobenmittelwert von 18 Jahren und interpretieren Sie diesen kurz inhaltlich. (3 Punkte)

- $\frac{\partial \, log(wage_i)}{\partial \, exper_i} = 0.077 - 2 \cdot 0.001 \cdot exper = 0.077 - 2 \cdot 0.001 \cdot 18 = 0.041$
- Am Stichprobemittelwert von 18 Jahren Berufserfahrung geht c.p. ein weiteres Jahr Berufserfahrung mit gut 4 % höherem erwartetem Lohn einher.

d) Interpretieren Sie den Koeffizienten b_5 inhaltlich und statistisch. (1 Punkt)

- Der mittlere Lohn ist für Männer unter sonst gleichen Bedingungen im Vgl. zum Lohn für Frauen um über 18% höher (exakt: exp(0.184)-1 = 0.202, also gut 20%)
- Der Koeffizient ist auf dem 5%-Signifikanzniveau statistisch signifikant von Null verschieden, (genauer: auf dem 2%-Niveau signifikant, da p-Wert von 0.017).

e) Sie vermuten, dass die Fehlertermvarianz mit Bildung variiert: $var(\varepsilon_i) = \sigma^2 \cdot educ_i$.

e1) Man kann die Schätzgleichung so transformieren, dass homoskedastische Fehlerterme resultieren. Stellen Sie am Beispiel diese Transformation dar und zeigen Sie formal die Auswirkung auf die Varianz des Störterms. (4 Punkte)

- Modell transformieren, indem alle Variablen mit $1/\sqrt{educ_i}$ multipliziert werden:
- $\frac{log(wage_i)}{\sqrt{educ_i}} = \frac{\beta_0}{\sqrt{educ_i}} + \beta_1 \frac{exper_i}{\sqrt{educ_i}} + \beta_2 \frac{exper_i^2}{\sqrt{educ_i}} + \beta_3 \frac{educ_i}{\sqrt{educ_i}} + \beta_4 \frac{nonwhite_i}{\sqrt{educ_i}} + \beta_5 \frac{male_i}{\sqrt{educ_i}} + \frac{\varepsilon_i}{\sqrt{educ_i}}$
- Varianz des transformierten Störterms ist homoskedastisch:

$$var\left\{\frac{\varepsilon_i}{\sqrt{educ_i}}\right\} = \frac{1}{educ_i} \cdot var\{\varepsilon_i\} = \frac{1}{educ_i} \cdot \sigma^2 \cdot educ_i = \sigma^2$$

e2) Welche weitere Möglichkeit haben Sie, um dem Problem heteroskedastischer Fehlerterme zu begegnen? Benennen Sie diese und geben Sie je einen Vor- und Nachteil an. (3 Punkte)

- Berechnung von White-korrigierten Standardfehlern (nicht erwartet) mittels $\hat{var}\{b\} = \left(\sum_{n=1}^{N} x_i x_i' \right)^{-1} \cdot \sum_{n=1}^{N} e_i^2 x_i x_i' \cdot \left(\sum_{n=1}^{N} x_i x_i' \right)^{-1}$, wobei e_i der KQ-Störterm ist.

- Vorteil(e):

  1. Die Korrektur kann angewendet werden, ohne die genaue Form der Heteroskedastie zu kennen, es ist also ein allgemein anwendbares Verfahren.

  oder

  2. Die Korrektur kann aus KQ-Schätzwerten berechnet werden.

- Nachteil: Kennt man die genaue Form der Heteroskedastie, so sind andere Schätzer, wie z.B. der FGLS-Schätzer, effizienter als KQ mit White-Standardfehlern.

f) Sie vermuten weiterhin, dass ein Spezifikationsproblem vorliegen könnte und wollen dies mit einem RESET-Test überprüfen. Erläutern Sie am Beispiel zunächst die Vorgehensweise dieses Tests und beschreiben Sie die zugrunde liegende Test-Idee. (4 Punkte)

- Vorgehensweise: KQ schätzen, vorhergesagte Werte berechnen, Polynom bilden und als Regressoren mit ins Modell aufnehmen, z. B.

$$log(wage_i) = \beta_0 + \beta_1 exper_i + \beta_2 exper_i^2 + \beta_3 educ_i + \beta_4 nonwhite_i + \beta_5 male_i + \beta_6 predict(log(wage_i))^2 + \varepsilon_i,$$

  anschließend β_6 per *t*- oder F-Test auf statistische Signifikanz testen. Besteht diese, ist das Modell fehlspezifiziert.

- Test-Idee: Solange die vorhergesagten Werte noch Erklärungsgehalt haben, ist der Erklärungsgehalt von *X* noch nicht vollständig ausgeschöpft.

Aufgabe 2: [8 Punkte]

Sie untersuchen den Zusammenhang zwischen Arbeitslosigkeit und Armut anhand von Jahresdaten für den Zeitraum von 1980 bis 2003. Sie schätzen folgendes Modell:

$$poverty_t = \beta_0 + \beta_1 unempl_t + \varepsilon_t,$$

mit
***poverty*$_t$ Armutsrate in Jahr *t*, in %,**
***unempl*$_t$ Arbeitslosigkeitsrate in Jahr t, in %.**

a) Sie vermuten, dass die Störterme autokorreliert sein könnten. Der geschätzte Autokorrelationskoeffizient $\hat{\rho}$ betrage 0.2876. Approximieren Sie die Teststatistik des Durbin-Watson Tests und testen Sie auf dem 5% Niveau, ob positive Autokorrelation erster Ordnung vorliegt. Geben Sie hierzu auch die Nullhypothese, die Alternativhypothese sowie die Freiheitsgrade und die kritischen Werte an. (5 Punkte)

- $dw \approx 2 - 2 \cdot \hat{\rho} = 2 - 2 \cdot 0.2876 = 0.5752 = 1.4248$

- $H_0 : \rho \leq 0 ; H_1 : \rho > 0$

- mit den Freiheitsgraden K = 2 (inkl. Konstante) und T = 24 Beobachtungen erhält man: $d_L = 1.27$; $d_U = 1.45$.

- 1.27 < *dw* < 1.45. Die Teststatistik liegt im Inkonklusionsbereich, es kann also am 5%-Niveau keine Aussage über positive Autokorrelation erster Ordnung gemacht werden.

b) Wie ändern Sie das Testprozedere, wenn zusätzlich die verzögerte abhängige Variable als Regressor im Modell berücksichtigt wird? Erläutern Sie Ihre Vorgehensweise. (3 Punkte)

- Asymptotischer Test: Die Residuen aus t, ε_t, werden regressiert auf $unempl_t$, $poverty_{t-1}$ sowie ε_{t-1};

- Zwei Alternativen:

  1) Mit einem *t*-Test kann sodann getestet werden, ob der Koeffizient zu ε_{t-1}, ρ, statistisch signifikant von Null verschieden ist; dies würde auf Autokorrelation schließen lassen.

  2) R^2 der Hilfsregression heranziehen und Teststatistik $(T\text{-}1)\cdot R^2$ berechnen. Diese ist χ^2 -verteilt mit einem Freiheitsgrad. Ist die Teststatistik größer als der kritische Wert, wird H_0: $\rho = 0$ verworfen.

Aufgabe 3: [5 Punkte]

Sie wollen den Zusammenhang zwischen Abiturnoten (*Abinote*$_i$) und Familieneinkommen (*Fameink*$_i$) untersuchen und formulieren hierzu das folgende Modell:

$$Abinote_i = \beta_0 + \beta_1 Fameink_i + \upsilon_i .$$

Sie vermuten, dass das Familieneinkommen durch die einzig beobachtbare Variable H_i (Haushaltseinkommen) fehlerhaft gemessen wird: $H_i = Fameink_i + u_i$. Es gelte hierbei, dass υ_i und u_i sowie *Fameink*$_i$ und u_i unkorreliert sind.

a) Zeigen Sie kurz, welche Auswirkung der Messfehler in der erklärenden Variable auf die Kovarianz zwischen geschätztem Fehlerterm des Modells und erklärender Variable hat; unterstellen Sie hierfür einen positiven Einfluss des Familieneinkommens auf die Abiturnote. (3 Punkte)

- Unter Beachtung des Messfehlers lautet die Funktion

$$Abinote_i = \beta_0 + \beta_1 (H_i - u_i) + \upsilon_i = \beta_0 + \beta_1 H_i + \underbrace{\upsilon_i - \beta_1 u_i}_{\varepsilon_i} = \beta_0 + \beta_1 H_i + \varepsilon_i$$

- $E\{H_i \varepsilon_i\} = E\{(Fameink_i + u_i)(\upsilon_i - \beta_1 u_i)\} = ... = -\beta_1 \sigma_u^2$

- Bei positivem Einfluss des Familieneinkommens auf die Abiturleistung, $\beta_1 > 0$, sind H_i und ε_i negativ miteinander korreliert.

b) Erläutern Sie kurz die Auswirkungen des Messfehlers auf den KQ-Schätzer. (2 Punkte)

- Der KQ-Schätzer ist inkonsistent.

- “attenuation bias”:

  (nicht erwartet:) $plim\, b_1 = \beta_1 \left(1 - \frac{\sigma_u^2}{\sigma_{Fameink}^2 + \sigma_u^2}\right)$, mit $\sigma_u^2 > 0$ folgt $|b_1| < |\beta_1|$.

- Verzerrung zur Null hin, wahrer Effekt wird unterschätzt.

Aufgabe 4: [10 Punkte]

Sie interessieren sich für die Determinanten des Monatsverdienstes W_i eines Abiturienten *i*. Ihnen liegen dazu Daten über die Punktezahl im Abiturzeugnis S_i und das Alter A_i vor. Sie wollen folgendes Modell schätzen:

$$W_i = \beta_0 + \beta_1 S_i + \beta_2 A_i + \varepsilon_i .$$

Sie vermuten, dass im vorliegenden Fall die Variable Punktezahl im Abiturzeugnis endogen ist und beschließen, einen IV-Schätzansatz anzuwenden.

a) Stellen Sie die drei Momentenbedingungen für den IV-Schätzer auf und erläutern Sie inhaltlich, warum hier möglicherweise eine Instrumentvariable benötigt wird. (4 Punkte)

- $E\{\varepsilon_i\} = E\{W_i - \beta_0 - \beta_1 S_i - \beta_2 A_i\} = 0$
- $E\{\varepsilon_i A_i\} = E\{(W_i - \beta_0 - \beta_1 S_i - \beta_2 A_i)A_i\} = 0$
- $E\{\varepsilon_i z_i\} = E\{(W_i - \beta_0 - \beta_1 S_i - \beta_2 A_i)z_i\} = 0$
- Es könnte sein, dass ein weiterer, unbeobachteter Faktor, der im Fehlerterm aufgefangen wird, sowohl Schulbildung als auch Verdienst beeinflusst (z.B. Disziplin, Motivation), damit aber wäre die Schulbildung mit dem Fehlerterm korreliert, der KQ-Schätzer wäre inkonsistent.

b) Erläutern Sie kurz den zentralen Unterschied zwischen dem IV-Schätzer und dem GIVE-Schätzer. (1 Punkt)

- IV: Eine Instrumentvariable pro endogener Variable
- GIVE: Mehrere Instrumentvariablen pro endogener Variable (nicht erwartet: die mit unterschiedlicher Gewichtung in die Schätzung einbezogen werden).

c) Als Instrumentvariablen stehen Ihnen die Schulbildung der Mutter M_i, die Schulbildung des Vaters, V_i, sowie die Entfernung des Wohnorts zum nächsten Gymnasium, E_i, zur Verfügung. E_i hat sich als Instrumentvariable für die vorliegende Fragestellung bewährt. Wie gehen Sie vor, um die Eignung von M_i und V_i als Instrumente zu überprüfen? (5 Punkte)

- Voraussetzung: Der IV-Schätzer ist mit E_i identifiziert, $R_1 = K$, dies ist per Annahme gegeben.
- GIVE-Schätzung mit allen Instrumenten; Hilfsregression der Residuen auf alle Instrumente und Berechnen der Teststatistik, ξ_R, aus R^2 der Hilfsregression; anschließend GIVE-Schätzung mit den als valide bekannten Instrumenten und Bestimmen von ξ_{R_1} aus analoger Hilfsregression der Residuen auf die R_1 Bedingungen;
- Die Teststatistik für die $R - R_1$ unsicheren Instrumente kann dann berechnet werden als $\xi_R - \xi_{R_1}$;
- $\xi_R - \xi_{R_1}$ ist unter H_0 (= Vorliegen von R gültigen Momentenbedingungen) χ^2 -verteilt mit $R - R_1$, hier also 2 Freiheitsgraden.
- Wird H_0 abgelehnt, so ist das Modell nicht überidentifiziert, es kann aber nicht gesagt werden, ob M_i oder V_i oder beide ungültige Instrumente sind.

Aufgabe 5: [12 Punkte]

Sie interessieren sich für den Einfluss des Stundenlohns auf das Arbeitsangebot. Sie formulieren hierzu folgendes Modell:

$$\log(hrs_{it}) = \alpha_i + \beta_0 \log(wage_{it}) + \beta_1 male_i + \varepsilon_{it},$$

wobei
hrs **Arbeitsstunden pro Jahr,**
wage **Bruttostundenlohn,**
male **Dummyvariable (= 1, wenn Person *i* ein Mann, = 0 sonst).**

Panel-Analysen mit Daten von 532 Erwerbstätigen, die von 1989 bis 1998 befragt wurden, ergeben folgende Schätzergebnisse (Standardfehler in Klammern):

	(1) pooled	(2) between	(3) fixed effects	(4) random effects
constant	4.218***	3.223***	–	3.430***
	(0.12)	(0.38)		(0.32)

log(wage)	0.133**	0.138**	0.105*	0.128**
	(0.055)	(0.068)	(0.062)	(0.058)
male	0.242***	0.251***	–	0.237***
	(0.026)	(0.027)		(0.023)

a) Interpretieren Sie den geschätzten Koeffizienten des fixed effects Modells für β_0 statistisch und inhaltlich. (2 Punkte)

- Elastizität, da logarithmierte Werte auf beiden Seiten der Gleichung
- Steigt der Lohn um 1%, so steigt das Arbeitsangebot um gut 0.1%.
- Der Koeffizient ist auf dem 10%-Niveau statistisch signifikant von Null verschieden.

b) Erläutern Sie formal – z.B. für ein Modell mit zwei Perioden –, wieso bei der fixed effects Schätzung keine Koeffizienten für *male* ausgewiesen werden. (4 Punkte)

- Informationsbasis und Grundlage für den fixed effects Schätzer ist die Variation über die Zeit. Damit folgt aber, dass zeitkonstante Größen heraus gerechnet werden:

$$\log(hrs_{i2}) = \alpha_i + \beta_0 \log(wage_{i2}) + \beta_1 male_i + \varepsilon_{i2}$$

$$-\ \log(hrs_{i1}) = \alpha_i + \beta_0 \log(wage_{i1}) + \beta_1 male_i + \varepsilon_{i1}$$

$$= \log(hrs_{i2}) - \log(hrs_{i2}) = \beta_0 (\log(wage_{i2}) - \log(wage_{i1})) + \varepsilon_{i2} - \varepsilon_{i1}$$

c) Wieso können sich die Ergebnisse der between und fixed effects Schätzung für β_0 unterscheiden? (2 Punkte)

- Der between-Schätzer basiert auf gemittelten Daten über alle Beobachtungen *i*; der fixed effects-Schätzer auf Abweichungen vom individuellen Mittelwert;
- Bei unterschiedlicher Datengrundlage können unterschiedliche Ergebnisse zustande kommen.

d) Die Teststatistik eines hier relevanten Hausman Tests sei 3.86. Geben Sie die Nullhypothese sowie die Anzahl der Freiheitsgrade an und testen Sie auf dem 1% Signifikanzniveau, ob der random effects oder der fixed effects Schätzer heranzuziehen ist. (4 Punkte)

- H_0: $b_{FE} = b_{RE}$; oder H_0: cov(α_i, log($wage_i$)) = 0
- Anzahl der Freiheitsgrade = 1
- $\chi^2_{krit} = 6.635$
- 3.86 < 6.635, H_0 wird nicht verworfen, d.h. es sollten die Schätzergebnisse des random effects Modells herangezogen werden.

Aufgabe 6: [45 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 0.75 Punkte, für jede falsche Antwort werden 0.75 Punkte abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

F	Ein Spaltenvektor ergibt sich als die Quadratwurzel eines Zeilenvektors.
W	Der "PE Test" verwendet *t*-Teststatistiken.
F	Ein Parameterschätzer ist effizient, wenn er gegen seinen Erwartungswert konvergiert.
W	Die Varianz des KQ-Schätzers lässt sich auf Basis von Information über die Störtermvarianz und die $X'X$ Matrix berechnen.

F	Der Prais-Winsten-Schätzer basiert auf transformierten Daten ohne Berücksichtigung der ersten drei Beobachtungen.
F	In ein Modell mit logarithmierter abhängiger Variable können keine Dummy-Variablen als erklärende Variablen eingefügt werden.
W	Der Goldfeld-Quandt Test ist ein Test auf die Gleichheit der Varianz zweier Teilstichproben.
W	Die Normalverteilung ist eine symmetrische Verteilungsfunktion.
F	Wird die Nullhypothese des Sargan-Tests verworfen, so ist der Störterm heteroskedastisch.
F	GMM Verfahren benötigen Verteilungsannahmen für die abhängige Variable.
W	Für die Vorhersage von y_i spielt es eine Rolle, ob das Modell linear oder loglinear geschätzt wird.
F	Der RESET Test nutzt Wurzeln der Residuen, um ein Modell auf Fehlspezifikation zu überprüfen.
F	Bei Messfehlern in der abhängigen Variable sind die Parameterschätzer zu groß.
F	Liegen in einem Modell statt einer drei endogene Variable vor, so reicht eine zusätzliche Instrumentvariable aus, um das Problem zu lösen.
F	Monte Carlo Studien erfordern Paneldatensätze.
F	Das "nonlinear least squares" Verfahren erlaubt es, solche Spezifikationen zu schätzen, die nichtlinear in den Parametern sind.
W	Modelle in reduzierter Form enthalten auf der rechten Seite keine endogenen erklärenden Variablen.
W	Newey-West Standardfehler korrigieren für Heteroskedastie unbekannten Ursprungs ebenso wie für Autokorrelation.
W	Die Nullhypothese H_0: $\beta \leq c$ wird bei 1500 Freiheitsgraden am 5 Prozentniveau verworfen, wenn als Teststatistik der t-Wert größer als 2.79 ist.
F	Enthält das lineare Regressionsmodell eine verzögerte endogene Variable (y_{t-1}), dann sollte für den Test auf Autokorrelation erster Ordnung des Störterms der Durbin-Watson Test verwendet werden.
F	Wenn autokorrelierte Störterme vorliegen, ist der Feasible-GLS Schätzer BLUE.
F	Für das Ergebnis eines Chow-Tests ist es unerheblich, ob die Störmtermvarianz korrekt spezifiziert ist.
F	Vom „least squares dummy variables“ Schätzer spricht man, wenn bei Vorliegen von Paneldaten Dummy-Variablen für die Perioden in das Modell aufgenommen werden.
F	Das angepasste R^2 einer Schätzung kann nicht negativ werden.
W	Die Standardnormalverteilung hat einen Erwartungswert von Null.
F	An der Parametermitte sagt ein KQ-Schätzer den effizienten Wert vorher.
W	Wenn statt eines Cochrane-Orcutt Schätzers ein iterativer Cochrane-Orcutt Schätzer verwendet wird, steigt die Effizienz der Schätzung von ρ.
F	Bei moving average Prozessen im Störterm sind alle Elemente der Varianz-Kovarianz Matrix des Störterms von Null verschieden.
F	Im Gegensatz zum KQ-Schätzer optimieren GMM (generalized method of moments) Schätzer eine quadratische Zielfunktion.
W	Durch das Hinzufügen weiterer erklärender Variablen kann der angepasste R^2 Wert sinken.
W	Das BIC Kriterium fällt umso günstiger aus, je kleiner die Fehlerquadratsumme bei gegebener Parameter- und Beobachtungszahl ist.
F	Positive Autokorrelation kommt seltener vor als negative.

W	Die Summe quadrierter, standardnormalverteilter Zufallsvariablen ist Chi-quadrat verteilt.
F	AR(1) Störterme sind heteroskedastisch.
F	Simultane Gleichungssysteme in struktureller Form können mit dem KQ-Schätzer unverzerrt geschätzt werden, solange jede Gleichung einzeln betrachtet wird.
F	Der Durbin-Watson Test verallgemeinert den PE Test.
W	Im linearen Regressionsmodell wird unterstellt, dass die abhängige Variable eine Zufallsvariable ist.
W	Der Durbin-Wu-Hausman Test auf Endogenität einer erklärenden Variablen wird durchgeführt, indem der Regressionsgleichung eine zusätzliche erklärende Variable hinzugefügt wird.
W	Die Annahme der statistischen Unabhängigkeit zweier Zufallsvariablen ist stärker als die Annahme mittlerer bedingter Unabhängigkeit, also z.B. $E\{u \mid X\} = 0$.
W	Mithilfe hedonischer Preisfunktionen lassen sich einzelne Eigenschaften eines Gutes bewerten.
W	Je nach Wahl der Gewichtungsmatrix W des GMM Schätzers ergeben sich unterschiedliche, aber stets konsistente Schätzer.
F	Der Breusch-Pagan Test auf Heteroskedastie ist eine Verallgemeinerung des Durbin-Watson Tests.
F	Auch bei exakter Multikollinearität kann der Kleinstquadrateschätzer unverzerrt geschätzt werden.
W	Mithilfe eines linearen Regressionsmodells lassen sich Elastizitäten schätzen.
W	Die Inkonsistenz eines Steigungsparameters führt zur Inkonsistenz der gleichzeitig geschätzten Regressionskonstanten.
F	Die Dichtefunktion der *t*-Verteilung hat ihr Minimum bei 1,96.
W	Das Auslassen einer relevanten erklärenden Variablen kann zu verzerrten Schätzern führen.
W	Der nichtlineare Kleinstquadrateschätzer bestimmt diejenigen Parameter, die die Summe der quadrierten Störterme minimieren.
F	Auf Basis linearer Modelle geschätzte Koeffizienten können nie als Kausaleffekte interpretiert werden.
W	Damit der fixed effects-Schätzer konsistent ist, müssen die erklärenden Variablen strikt exogen sein.
W	Mit steigender Zahl von Freiheitsgraden konvergiert die t-Verteilung zur Normalverteilung.
W	Die kritischen Werte des Durbin-Watson-Tests sind für Tests auf positive Autokorrelation erster Ordnung anwendbar.
F	Um *k* Parameter zu identifizieren, benötigt man *k*-1 Momentenbedingungen.
F	Die within Transformation modifiziert alle Beobachtungen, indem der individuelle Störterm von den erklärenden Variablen abgezogen wird.
F	Bei der Ableitung des KQ-Schätzers im linearen Modell erhält man so viele Normalgleichungen wie Beobachtungen vorliegen.
F	Zur Unverzerrtheit des fixed effects-Schätzers ist keine Aussage möglich.
F	Je nachdem, ob die Schätzung mit oder ohne Konstante durchgeführt wird, spricht man bei Paneldaten von fixed oder random effects Schätzungen.
F	Das R^2 ist in Modellen mit großer Beobachtungszahl größer als bei Modellen mit kleiner Beobachtungszahl.
F	Die optimale Gewichtungsmatrix des GMM Modells W entspricht der Varianz-Kovarianzmatrix der abhängigen Variablen.
F	Der quadrierte Korrelationskoeffizient für die abhängige Variable und ihren vorhergesagten Wert ist immer größer als das R^2 der zugehörigen Schätzung.

Aufgabe 7: **[15 Punkte]**

Wahr oder falsch? Begründen Sie Ihre Auffassung (Bsp.: "Stimmt, weil..." bzw. "Stimmt nicht, weil..."). Nur bei korrekter Begründung erhält jede richtige Antwort 1.5 Punkte; Angaben **ohne Begründung** werden **nicht gewertet**.

W	**Bei Paneldaten ist der between-Schätzer ineffizient.** ➔ Die zeitliche Variation wird nicht berücksichtigt, d.h. es werden nicht alle Informationen genutzt.
W	**Beim random effects Modell besteht der Störterm aus zwei Zufallsvariablen.** ➔ Ein Teil des Störterms ist ein individuenspezifischer Term, der als Zufallsvariable geschätzt wird.
F	**Der KQ-Schätzer im einfachen Modell minimiert die Abstände zwischen *y* und seinen vorhergesagten Werten.** ➔ Der KQ-Schätzer minimiert die <u>quadrierte</u> Summe der Residuen.
F	**Die Parameter β_j im Modell $g(x_i,\beta)=\beta_1 x_{i1}^{\beta_2} x_{i2}^{\beta_3}$ lassen sich nicht per KQ schätzen.** ➔ Mittels Logarithmierung kann das Modell linearisiert und mit KQ geschätzt werden.
W	**Alle Fragestellungen der Querschnittsdaten können auch mit Paneldatenanalyse beantwortet werden.** ➔ Mit Paneldaten kann die Panel-, d.h. Zeitdimension genutzt werden, muss aber nicht.
F	**Das kritische Signifikanzniveau einseitiger Hypothesentests variiert mit der Beobachtungszahl.** ➔ Das Signifikanzniveau wird exogen festgelegt.
W	**Die Berücksichtigung einer verzögerten endogenen Variable als erklärende Variable kann zur Inkonsistenz des KQ-Schätzers führen.** ➔ Wenn gleichzeitig Autokorrelation vorliegt; dann automatisch Korrelation der erklärenden Variable mit dem Störterm.
F	**Der Durbin-Watson Test ist ein F-Test auf die Signifikanz der Koeffizienten von Polynomen der vorhergesagten abhängigen Variable.** ➔ Der beschriebene Test ist der RESET-Test. Oder: ➔ Der Durbin-Watson Test testet H_0: $\rho = 0$ z.B. über den bounds-Test.
W	**Bei gegebenem *N* ist die Wahrscheinlichkeit eines Typ II Fehlers umso höher, je niedriger die Wahrscheinlichkeit eines Typ I Fehlers.** ➔ Ein Typ II Fehler kann nur gemacht werden, wenn H_0 nicht verworfen wird. Fällt α, steigt diese Wahrscheinlichkeit.
F	**Der Cochrane-Orcutt Schätzer nutzt mehr Beobachtungen als der Prais-Winsten Schätzer.** ➔ Umgekehrt, Prais-Winsten macht z.B. t = 1 nutzbar.