**Aufgabe 1:** **[24 Punkte]**

In der Mitte der 1970er versuchte die US-Regierung, Kfz-Treibstoffe anhand von u.a. Verbrauchswerten von Vorperioden und soziodemographischen Charakteristika auf 50 US-Bundesstaaten zu allokiieren. Ein diesem Vorhaben implizit zugrunde liegendes ökonometrisches Modell könnte dabei wie folgt gelautet haben:

$$PCON_i = \beta_0 + \beta_1 \cdot REG_i + \beta_2 \cdot TAX_i + e_i ,$$

wobei

$PCON_i$: Mineralöl-Konsum im i-ten Bundesstaat (in Mio. Energieeinheiten)
$REG_i$: Kfz-Zulassungen im i-ten Bundesstaat (in Tsd.)
$TAX_i$: Mineralöl-Steuer im i-ten Bundesstaat (US-Cents pro Gallone Treibstoff)
$POP_i$: Bevölkerung im i-ten Bundesstaat (in Tsd.)

Mit Daten aus dem Jahr 1985 wurde das Modell in R geschätzt:

```
Call:
lm(formula = PCON ~ REG + TAX)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 551.68799  186.27095    ?     0.00479 **
REG           0.18613      ?      15.883  < 2e-16 ***
TAX              ?      16.85588  -3.179  0.00261 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 253 on 47 degrees of freedom
Multiple R-Squared: 0.8664,     Adjusted R-squared: 0.8607
F-statistic: 152.4 on 2 and 47 DF,  p-value: < 2.2e-16
```

**a) Berechnen Sie unter Angabe des Rechenwegs die im Output fehlenden Werte für** **(2,5 Punkte)**

**a1) den t-Wert für $b_0$;**

$$t_{b_0} = \frac{b_0}{se(b_0)} = \frac{551.68799}{186.27095} = 2.96175$$

**a2) den Standardfehler für $b_1$;**

$$t_{b_1} = \frac{b_1}{se(b_1)} \Rightarrow se(b_1) = \frac{b_1}{t_{b_1}} = \frac{0.18613}{15.883} = 0.01172$$

**a3) den Koeffizienten für $b_2$;**

$$t_{b_2} = \frac{b_2}{se(b_2)} \Rightarrow b_2 = t_{b_2} \cdot se(b_2) = -3.179 \cdot 16.85588 = -53.58484$$

**a4) die geschätzte Fehlertermvarianz;**

$$\hat{\sigma}^2 = (\hat{\sigma})^2 = 253^2 = 64009$$

**b) Worin unterscheiden sich das $R^2$ und das angepasste $R^2$?** **(1,5 Punkte)**

- Das angepasste $R^2$ korrigiert das Gütemaß um die Anzahl der Freiheitsgrade (N-K statt N-1 im Zählerterm).

**c) Interpretieren Sie den Koeffizienten des Parameters $b_1$ statistisch und inhaltlich. (2 Punkte)**

- Der Koeffizient ist auf dem 1%-Niveau (s. p-Wert) positiv statistisch signifikant von Null verschieden;

- Steigt die Anzahl der Kfz-Zulassungen um eine Einheit, also 1000 Personen, so nimmt der Mineralöl-Konsum um 0.186 Millionen Energieeinheiten zu.

**d) Um Fehlspezifikation vorzubeugen, werden sodann die nachfolgenden Schritte durchgeführt. (8 P)**

**d1) Unter der Annahme, dass die Größe der Bevölkerung im i-ten Bundesstaat ebenfalls eine Rolle spiele, wird die Modellspezifikation um den Indikator POP erweitert und erneut geschätzt.**

```
Call:
lm(formula = PCON ~ REG + TAX + POP)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 554.28355  188.09310   2.947  0.00503 **
REG           0.16379    0.05855   2.797  0.00750 **
TAX         -53.63094   17.01040  -3.153  0.00285 **
POP           0.01492    0.03830   0.390  0.69870
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 255.3 on 46 degrees of freedom
Multiple R-Squared: 0.8668,     Adjusted R-squared: 0.8581
F-statistic: 99.79 on 3 and 46 DF,  p-value: < 2.2e-16
```

**d11) Begründen Sie, welche der beiden Modellspezifikationen (aus Teilaufgabe a oder aus d) Sie bevorzugen.**

- Die Spezifikation aus a) ist zu bevorzugen: das korrigierte $R^2$ ist höher (0.8607 > 0.8581); der Parameter ist nicht statistisch signifikant.

**d12) Kommen Sie zur selben Entscheidung, wenn Sie anhand der AIC- und BIC-Werte urteilen? Berechnen Sie diese (Hinweis: Im ersten Modell ist $\log \frac{1}{N}\sum e_i^2 = 11.00491$, im zweiten Modell ist $\log \frac{1}{N}\sum e_i^2 = 11.00162$).**

- Mit $K$=3 und $N$=50 folgt:

$$AIC_1 = \log \frac{1}{N}\sum e_i^2 + \frac{2K}{N} = 11.00491 + \frac{6}{50} = \mathbf{11.12491}$$

$$AIC_2 = \log \frac{1}{N}\sum e_i^2 + \frac{2K}{N} = 11.00162 + \frac{8}{50} = \mathbf{11.16162}$$

$$BIC_1 = \log \frac{1}{N}\sum e_i^2 + \frac{K}{N}\log(N) = 11.00491 + \frac{3}{50}\log(50) = \mathbf{11.23963}$$

$$BIC_2 = \log \frac{1}{N}\sum e_i^2 + \frac{K}{N}\log(N) = 11.00162 + \frac{4}{50}\log(50) = \mathbf{11.31458}$$

- Die AIC/BIC-Werte für die Spezifikation aus Teilaufgabe a) sind kleiner als die für Spezifikation aus d): Die Spezifikation aus a) ist also auch nach den AIC/BIC-Kriterien vorzuziehen.

**d2) Es soll ein RESET-Test für die Spezifikation aus Teilaufgabe a) durchgeführt werden. Stellen Sie die Hilfsregression für eine Variante dieses Tests auf.**

- $PCON_i = \beta_0 + \beta_1 \cdot REG_i + \beta_2 \cdot TAX_i + \gamma \hat{y}_i^2 + e_i$, wobei $\hat{y}_i^2$ der vorhergesagte, quadrierte Wert von *PCON* aus einer KQ-Schätzung ist (dies ist die einfachste Variante; Erweiterung um höhere Potenzen ebenfalls korrekt).

**d3) Das Modell wird in modifizierter Spezifikation geschätzt. Interpretieren Sie die Steigungsparameter der folgenden Gleichung:**

$$\log\left(PCON\right) = 6.15 + 0.009 \cdot REG - 0.042 \cdot \log\left(TAX\right).$$

- *PCON* wird log-linear (auch: semi-elastisch) auf *REG* und log-log auf *TAX* regressiert: Steigt REG um eine Einheit, also 1000 Personen, so nimmt der Mineralöl-Konsum um ca. 0,9% zu; steigt die Steuer um 1% (in US-Cent pro Gallone), so nimmt der Mineralöl-Konsum um ca. 0,042% ab.

**e) Es kann weiterhin argumentiert werden, dass mit ansteigender Zahl der Kfz-Zulassungen die Variation im Mineralöl-Konsum steigt, was vermuten lässt, dass Heteroskedastie vorliegt. Es werden daher für das ursprüngliche, nicht logarithmierte Modell aus Teilaufgabe a) der Breusch-Pagan Test und der White-Test auf Heteroskedastie durchgeführt. (10 Punkte)**

**e1) Der Breusch-Pagan Test hat eine Test-Statistik von 14.03 bei einem Freiheitsgrad von d.f.=1. Testen Sie auf dem 5%-Signifikanzniveau, ob Heteroskedastie vorliegt und dokumentieren Sie Ihre Vorgehensweise.**

- Mit einem Freiheitsgrad liegt der kritische Wert bei $\chi^2_{0.95;1} = 3.84$; die Nullhypothese, dass keine Heteroskedastie vorliegt, kann auf dem 5%-Signifikanzniveau abgelehnt werden.

**e2) Stellen Sie die Gleichung der Hilfsregression des White-Tests auf und testen Sie mit dem ermittelten $R^2$-Wert von 0.6645 auf dem 5%-Signifikanzniveau auf Heteroskedastie. Dokumentieren Sie Ihre Vorgehensweise.**

- $e_i^2 = \beta_0 + \beta_1 \cdot REG_i + \beta_2 \cdot TAX_i + \gamma_1 \cdot REG_i^2 + \gamma_2 \cdot TAX_i^2 + \gamma_3 \cdot REG_i \cdot TAX_i + \upsilon_i$

- Die Test-Statistik $N \cdot R^2$ ist $\chi^2$-verteilt mit *P* Freiheitsgraden, wobei *P* die Anzahl der Regressoren (ohne Konstante) aus der Hilfsregression ist.

- Hier: $N \cdot R^2 = 50 \cdot 0,6645 = 33,225$.

- Der kritische Wert liegt bei $\chi^2_{0.95;5} = 11.07$; die Nullhypothese, dass keine Heteroskedastie vorliegt, kann auf dem 5%-Signifikanzniveau abgelehnt werden.

**e3) Erläutern Sie kurz für das zugrunde liegende Modell die Vorgehensweise eines GLS-Schätzers zur Beseitigung der beschriebenen Heteroskedastie.**

- e) impliziert, dass die Heteroskedastie proportional mit der Anzahl der Kfz-Zulassungen wächst: $V\left\{\varepsilon_i \mid x_i\right\} = \sigma^2 \cdot h_i$, mit REG als Proportionalitätsfaktor $h_i$.
- Die Daten werden also mit $\sqrt{REG_i} = \sqrt{h_i}$ gewichtet:

$$\frac{PCON_i}{\sqrt{REG_i}} = \frac{\beta_0}{\sqrt{REG_i}} + \beta_1 + \beta_2 \cdot \frac{TAX_i}{\sqrt{REG_i}} + e_i$$

- Anschließend wird mit den transformierten Daten eine KQ-Schätzung durchgeführt.

**e4) Zur Korrektur von Heteroskedastie unbekannten Ursprungs werden gelegentlich White-Standardfehler berechnet. Erläutern Sie verbal je einen Vor- und einen Nachteil dieses Verfahrens.**

- Vorteil: Die Korrektur kann angewendet werden, ohne die genaue Form der Heteroskedastie zu kennen, es ist also ein allgemein anwendbares Verfahren;

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

**Aufgabe 2:** **[12 Punkte]**

Unterstellen Sie das folgende einfache Modell der Geldnachfrage:

$$\log(M1_t) = \beta_0 + \beta_1 \cdot \log(GDP_t) + \beta_2 \cdot \log(CPI_t) + e_t,$$

wobei

$M1_t$: nominaler M1-Geldbestand im Jahr t

$GDP_t$: reales Bruttosozialprodukt im Jahr t (in Mrd. US-$)

$CPI_t$: Preisindex im Jahr t.

Das Modell wird mit US-Quartalsdaten von 1950.1 bis 1999.4 geschätzt, ein (hier nicht dargestellter) Residuen-Plot lässt Autokorrelation vermuten. Aus einer KQ-Schätzung ergibt sich:

$$\sum_{t=2}^{T} \hat{e}_t \hat{e}_{t-1} = 1.357 \text{ und } \sum_{t=2}^{T} \hat{e}_{t-1}^2 = 1.381$$

**a) Berechnen Sie den Autokorrelationskoeffizienten $\rho$ und die (approximative) Durbin-Watson-Statistik $d$.** **(4 Punkte)**

- $\rho = \frac{\sum_{t=2}^{T} \hat{e}_t \hat{e}_{t-1}}{\sum_{t=2}^{T} \hat{e}_{t-1}^2} = \frac{1.357}{1.381} = \mathbf{0.9826}$

- $d \approx 2(1-\rho) = 2 - 2 \cdot 0.9826 = \mathbf{0.0348}$

**b) Unterstellen Sie eine Durbin-Watson-Statistik von 1,0 und testen Sie auf dem 5%-Signifikanzniveau auf positive Autokorrelation erster Ordnung. Geben Sie hierzu auch die Nullhypothese, die Alternativhypothese sowie die Freiheitsgrade und die jeweiligen kritischen Werte an.** **(4 Punkte)**

- Nullhypothese: $H_0 : \rho = 0$
- Alternativhypothese: $H_1 : \rho > 0$
- Freiheitsgrade: T=200 und K=3 (inkl. Konstante!)
- Kritische Werte bei T=200 und K=3: $d_L$=1.75; $d_U$=1.79
- Die Durbin-Watson Statistik von 1,0 liegt deutlich unter der unteren Grenze von 1,75; die Nullhypothese, dass keine positive Autokorrelation erster Ordnung vorliegt, $H_0 : \rho = 0$, kann auf dem 5%-Signifikanzniveau verworfen werden.

**c) Zeigen Sie für ein lineares Modell $y_t = x_t'\beta + \varepsilon_t$ mit Autokorrelation zweiter Ordnung, also mit $\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \upsilon_t$, wobei $\upsilon_t \sim i.i.d.(0, \sigma^2)$, wie eine Änderung der Spezifikation das Autokorrelationsproblem beseitigt.** **(4 Punkte)**

- In das lineare Modell sind die verzögerte abhängige Variable und die verzögerten erklärenden Variablen mit aufzunehmen:

$$y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} = \beta_0\left(1-\rho_1-\rho_2\right) + \beta_1\left(x_{1t} - \rho_1 x_{1t-1} - \rho_2 x_{1t-2}\right) + \beta_2\left(x_{2t} - \rho_1 x_{2t-1} - \rho_2 x_{2t-2}\right)$$
$$+ \underbrace{\varepsilon_t - \rho_1\varepsilon_{t-1} - \rho_2\varepsilon_{t-2}}_{=\upsilon_t}$$
$$\Leftrightarrow y_t = \beta_0\left(1-\rho_1-\rho_2\right) + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \beta_1\left(x_{1t} - \rho_1 x_{1t-1} - \rho_2 x_{1t-2}\right) + \beta_2\left(x_{2t} - \rho_1 x_{2t-1} - \rho_2 x_{2t-2}\right) + \upsilon_t$$

**Aufgabe 3:** **[15 Punkte]**

Unterstellen Sie das ‚wahre' Modell $y_{it} = x'_{it}\beta + \alpha_i + \varepsilon_{it}$ mit $\alpha_i$ als nicht-beobachtbarer, nicht-stochastischer zeitkonstanter Größe.

**a) Das Modell werde ohne Berücksichtigung der Panelnatur der Daten im Querschnitt geschätzt. Zeigen Sie, dass der KQ-Schätzer $b = \sum\left(x_{it}x'_{it}\right)^{-1} x_{it}y_{it}$ verzerrt ist. (4 Punkte)**

- Zu zeigen: $E(b) \neq \beta$

$$\begin{aligned} E(b) &= E\left\{\sum\left(x_{it}x'_{it}\right)^{-1} x_{it}y_{it}\right\} \\ &= E\left\{\sum\left(x_{it}x'_{it}\right)^{-1} x_{it}\left(x'_{it}\beta + \alpha_i + \varepsilon_{it}\right)\right\} \\ &= E\left\{\sum\left(x_{it}x'_{it}\right)^{-1} x_{it}x'_{it}\beta + \sum\left(x_{it}x'_{it}\right)^{-1} x_{it}\alpha_i + \sum\left(x_{it}x'_{it}\right)^{-1} x_{it}\varepsilon_{it}\right\} \\ &= \underbrace{E\left\{\sum\left(x_{it}x'_{it}\right)^{-1} x_{it}x'_{it}\beta\right\}}_{=\beta} + \underbrace{E\left\{\sum\left(x_{it}x'_{it}\right)^{-1} x_{it}\alpha_i\right\}}_{\neq 0} + \underbrace{E\left\{\sum\left(x_{it}x'_{it}\right)^{-1} x_{it}\varepsilon_{it}\right\}}_{=0} \end{aligned}$$

- Der erste Term entspricht $\beta$, der letzte Term ist mit den zugrunde liegenden Annahmen der Exogenität gleich Null. Nimmt man an, dass $\alpha_i > 0$, dann ist $E(b) \neq \beta$.

**b) Erläutern Sie kurz formal und verbal, wie man durch die Verwendung von Paneldaten $\beta$ unverzerrt schätzen kann; unterstellen Sie *t* = 1,2. (3 Punkte)**

- Im Zwei-Perioden-Panel lässt sich $\alpha_i$ durch das Bilden erster Differenzen beseitigen:

$y_{it} = \alpha_i + \beta_0 + \beta_1 x_{it} + u_{it}$.

(1) t=2: $y_{i2} = \alpha_i + \beta_0 + \beta_1 x_{i2} + u_{i2}$

(2) t=1: $y_{i1} = \alpha_i + \beta_0 + \beta_1 x_{i1} + u_{i1}$

(1) – (2):

$\left(y_{i2} - y_{i1}\right) = \left(\alpha_i - \alpha_i\right) + \beta_1\left(x_{i2} - x_{i1}\right) + \left(u_{i2} - u_{i1}\right) = \beta_1\left(x_{i2} - x_{i1}\right) + \left(u_{i2} - u_{i1}\right)$; $\alpha_i$ „kürzt" sich also heraus;

- Eine KQ-Schätzung mit den transformierten Daten ergibt unverzerrte Schätzwerte, da die zuvor „unberücksichtigten Variablen" nicht mehr im Modell erscheinen.

**c) Aus dem US-amerikanischen NLSY79 werden Daten von 545 vollzeitbeschäftigten Männern, die 1980 ihren Schulabschluss gemacht haben und über 8 Jahre beobachtet wurden, gezogen. Der logarithmierte Lohn wird auf die Merkmale Schuljahre, Arbeitsmarkterfahrung und quadrierte Arbeitsmarkterfahrung, Gewerkschaftsmitgliedschaft, Familienstand (verheiratet), Beschäftigung im öffentlichen Sektor und ethnische Zugehörigkeit regressiert. Es ergeben sich bei Anwendung unterschiedlicher Verfahren die folgenden Schätzergebnisse: (9 Punkte)**

| Abhängige Variable: ln(Wage) | KQ | between | fixed effects | random effects |
|---|---|---|---|---|
| Constant | -0.0344 | 0.490** | – | -0.104 |
| | (0.065) | (0.22) | | (0.11) |
| Schooling | 0.0994*** | 0.0948*** | – | 0.101*** |
| | (0.0047) | (0.011) | | (0.0089) |
| Experience | 0.0891*** | -0.0502 | 0.116*** | 0.112*** |
| | (0.010) | (0.050) | (0.0084) | (0.0083) |
| Experience^2 | -0.00285*** | 0.00511 | -0.00429*** | -0.00406*** |
| | (0.00071) | (0.0032) | (0.00061) | (0.00059) |
| Union member | 0.180*** | 0.274*** | 0.0812*** | 0.106*** |
| | (0.017) | (0.047) | (0.019) | (0.018) |
| Married | 0.108*** | 0.145*** | 0.0451** | 0.0625*** |
| | (0.016) | (0.041) | (0.018) | (0.017) |
| Black | -0.144*** | -0.139*** | – | -0.144*** |
| | (0.024) | (0.049) | | (0.048) |
| Hispanic | 0.0157 | 0.00548 | – | 0.0197 |
| | (0.021) | (0.043) | | (0.043) |
| Public sector | 0.00355 | -0.0563 | 0.0349 | 0.0302 |
| | (0.037) | (0.11) | (0.039) | (0.036) |
| Anmerkungen: Standardfehler in Klammern; *** p<0.01, ** p<0.05, * p<0.1 | | | | |

**c1) Wieso werden für den fixed effects-Schätzer keine Parameterwerte für die Regressoren Schooling, Black und Hispanic ausgewiesen? Erläutern Sie ausführlich.**

- Alle drei Regressoren sind zeitkonstante Größen und weisen daher keine ‚within'-Variation auf. Genau diese ist aber Grundlage für den fixed effects (within)-Schätzer, da dieser von den jeweiligen Beobachtungen für Person *i* zum Zeitpunkt *t* den personenbezogenen Mittelwert über die Zeit abzieht. Bei zeitkonstanten Größen ergibt sich hier somit Null für die transformierten Daten.

**c2) Erläutern Sie, warum sich die Schätzergebnisse für between- und fixed effects-Schätzer teils erheblich unterscheiden können.**

- 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.

**c3) Geben Sie sowohl die Bedingung an, unter der der random effects-Schätzer konsistent ist als auch die konkrete Hypothese, anhand derer dies im Rahmen des Hausman-Tests überprüft wird.**

- Die Annahme für das random effects-Modell und damit die zu testende Nullhypothese besagt, dass $x_{it}$ und $\alpha_i$ unkorreliert sind ($H_0$: cov($\alpha_i$, $x_{it}$) = 0). Wird die Nullhypothese verworfen, so ist das fixed effects-Modell vorzuziehen, welches Korrelation zwischen $x_{it}$ und $\alpha_i$ zulässt.

**Aufgabe 4:** **[7 Punkte]**

Gegeben sei ein lineares Modell, das das Gewicht von Neugeborenen auf verschiedene soziodemographische Merkmale regressiert:

$$\log\left(birthweight\right) = \beta_0 + \beta_1 male + \beta_2 parity + \beta_3 \log\left(fam\_inc\right) + \beta_4 packs + u ,$$

wobei

*birthweight*: Geburtsgewicht (in Unzen);

| | |
|---|---|
| *male*: | =1, falls das Neugeborene ein Junge ist; |
| *parity*: | das wievielte Kind das Neugeborene ist; |
| *fam_inc*: | Familieneinkommen (in 1988 1000 US-$); |
| *packs*: | durchschnittliche Anzahl der während der Schwangerschaft pro Tag gerauchten Zigaretten (in Packungen). |

**a) Im angeführten Modell könnte Endogenität eine Rolle spielen. Unter welcher Bedingung sind erklärende Variablen exogen? Erläutern Sie verbal und formal. (2 Punkte)**

- Erklärende Variablen sind exogen, wenn sie mit dem Störterm unkorreliert sind; $\operatorname{cov}(x_i, \varepsilon_i) = 0$.

**b) Mit Daten einer US-amerikanischen Stichprobe (*N*=1388) wird das Modell mit KQ und, um möglicher Endogenität Rechnung zu tragen, mit IV geschätzt. Für den IV-Schätzer wird *cigprice,* der Preis pro Packung Zigaretten in 1988 US-Cent im jeweiligen Bundesstaat, als Instrument für *packs* herangezogen. (5 Punkte)**

| Abhängige Variable: | KQ<br>log(birthweight) | IV<br>log(birthweight) | KQ<br>packs |
|---|---|---|---|
| Constant | 4.676***<br>(0.022) | 4.468***<br>(0.26) | 0.137<br>(0.10) |
| male | 0.0262***<br>(0.010) | 0.0298*<br>(0.018) | -0.00473<br>(0.016) |
| parity | 0.0147***<br>(0.0057) | -0.00124<br>(0.022) | 0.0181**<br>(0.0089) |
| log(faminc) | 0.0180***<br>(0.0056) | 0.0636<br>(0.057) | -0.0526***<br>(0.0087) |
| packs | -0.0837***<br>(0.017) | 0.797<br>(1.09) | – |
| cigprice | – | – | 0.000777<br>(0.00078) |
| $R^2$ | 0.04 | – | 0.03 |

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

**b1) Was lernen Sie aus dem Vergleich der Koeffizienten für die Variable *packs* in der KQ- und IV-Schätzung des Geburtsgewichts?**

- In der KQ-Schätzung ist *packs* negativ statistisch signifikant (mit jeder zusätzlich pro Tag gerauchten Packung sinkt das Geburtsgewicht um ca. 8%). Die Ergebnisse der IV-Schätzung hingegen deuten nicht auf einen von Null verschiedenen Effekt hin, sie stützen also nicht die Ergebnisse der KQ-Schätzung. Dies könnte auf Endogenität in der KQ-Schätzung hindeuten, da es mglw. nicht beobachtete Größen gibt, die das Rauchverhalten und das Geburtsgewicht beeinflussen, aber nur im Störterm aufgefangen sind. Hier könnte sodann Korrelation zwischen Störterm und *packs* existieren.

**b2) Diskutieren Sie die Qualität des vorliegenden IV-Schätzers.**

- *cigprice* ist ein schwaches Instrument für p*acks*: Das $R^2$ der Hilfsregression ist mit 0.03 sehr gering und der Koeffizient ist nicht statistisch signifikant.

**Aufgabe 5: [40 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.

| | |
|---|---|
| **W** | Der J-Test wird als *t*-Test durchgeführt. |
| **W** | HAC Standardfehler unterstellen auf eine feste Anzahl von Perioden begrenzte Störtermkorrelationen. |

| | |
|---|---|
| **F** | AR(1) Fehlerterme sind heteroskedastisch. |
| **F** | Schätzgleichungen mit quadratischen erklärenden Variablen erfordern nichtlineare Schätzverfahren. |
| **F** | Der Instrumentvariablenschätzer kann nicht mehr als ein Instrument pro Steigungsparameter nutzen. |
| **W** | Die Qualität von Instrumentvariablenschätzern hängt von der Qualität der Instrumente ab. |
| **W** | Im einfachsten Fall des KQ-Schätzers wird unterstellt, dass die erklärenden Variablen des linearen Modells keine Zufallsvariablen sind. |
| **F** | Beim random effects-Schätzer wird unterstellt, dass der unbeobachtete Störterm der Beobachtungseinheit $\left(\alpha_i\right)$ mit dem Zufallsstörterm $\left(\varepsilon_{it}\right)$ korreliert ist. |
| **F** | Die Koeffizienten endogener erklärender Variablen können als kausale Effekte interpretiert werden. |
| **F** | Zur Unverzerrtheit des fixed effects-Schätzers ist keine Aussage möglich. |
| **W** | Die kritischen Werte des Durbin-Watson-Tests sind für Tests auf positive Autokorrelation erster Ordnung definiert. |
| **W** | Die Varianz des KQ-Schätzers lässt sich auf Basis von Information über die Störterme und die $X'X$ Matrix berechnen. |
| **W** | Bei Autokorrelation in Form von moving average Störprozessen gibt es Fehlerterme, die nicht miteinander korreliert sind. |
| **F** | Der within-Schätzer nutzt ausschließlich Unterschiede zwischen den Beobachtungseinheiten, um die Steigungsparameter zu identifizieren. |
| **W** | Positive Autokorrelation kommt häufiger vor als negative. |
| **W** | Je nach Wahl der Gewichtungsmatrix W des GMM Schätzers ergeben sich unterschiedliche, aber stets konsistente Schätzer. |
| **W** | Der Instrumentvariablenschätzer ist ein Spezialfall des GMM (generalized method of moments) Ansatzes. |
| **F** | Grundidee des KQ-Schätzers ist, eine Linie so durch eine Punktwolke zu legen, dass die Summe der quadrierten horizontalen Abweichungen der beobachteten Werte von der Linie minimiert wird. |
| **W** | Das least squares dummy variable (LSDV) Modell berücksichtigt für jede Beobachtungseinheit eine Dummyvariable. |
| **F** | Der Breusch-Pagan Test auf Heteroskedastie ist eine Verallgemeinerung des White-Tests. |
| **F** | Im Gegensatz zum KQ-Schätzer optimieren GMM (generalized method of moments) Schätzer keine Zielfunktion. |
| **F** | Um K Parameter zu identifizieren, benötigt man K – 1 Momentenbedingungen. |
| **W** | Bei perfekter Multikollinearität ist die $X'X$ Matrix nicht invertierbar. |
| **W** | Der GIVE (generalized instrumental variables estimator) Schätzer kann in einem zweistufigen Verfahren angewendet werden. |
| **W** | Newey-West Standardfehler korrigieren für Heteroskedastie unbekannten Ursprungs ebenso wie für Autokorrelation. |

| | |
|---|---|
| **W** | Wenn $X$ deterministisch ist, gilt $E\{\varepsilon \mid X\} = E\{\varepsilon\}$. |
| **F** | Bei Vorliegen von Paneldaten lassen sich Probleme ausgelassener Variablen durch den random effects Schätzer lösen. |
| **F** | Im linearen Modell gibt die Regressionskonstante den Mittelwert der abhängigen Variable an. |
| **F** | Je größer der Typ II Fehler eines Tests, umso größer muss der Typ I Fehler sein. |
| **F** | Wird die Nullhypothese des Sargan-Tests verworfen, so ist das Schätzmodell überidentifiziert. |
| **W** | Der marginale Effekt einer erklärenden Variablen kann je nach Spezifikation des Modells unterschiedlich ausfallen. |
| **F** | Bei Messfehlern in der abhängigen Variable sind die Parameterschätzer auf Null hin verzerrt. |
| **W** | Der Korrelationskoeffizient für die abhängige Variable und ihren vorhergesagten Wert ist immer größer als das $R^2$ der zugehörigen Schätzung. |
| **W** | Das Residuum der linearen Regression ist orthogonal zu den vorhergesagten Werten der abhängigen Variablen. |
| **F** | Mit dem *t*-Test können mehrere lineare Hypothesen gleichzeitig gestestet werden. |
| **F** | Der Chow-Test prüft, ob die lineare oder loglineare Version der abhängigen Variable dem Modell besser entspricht. |
| **W** | Insignifikante Koeffizienten können größer als 100 sein. |
| **W** | Bei der Ableitung des KQ-Schätzers im linearen Modell erhält man so viele Normalgleichungen wie unbekannte Parameter vorliegen. |
| **F** | Die within Transformation modifiziert alle Beobachtungen, indem die Varianz des individuellen Störterms von den erklärenden Variablen abgezogen wird. |
| **W** | Mithilfe hedonischer Preisfunktionen lassen sich einzelne Eigenschaften eines Gutes bewerten. |
| **W** | Man kann den *t*-Test verwenden, um die Nullhypothese positiver Autokorrelation erster Ordnung zu überprüfen. |
| **W** | Der between-Schätzer ist konsistent, wenn der beobachtungsspezifische Störterm nicht mit den beobachtungsspezifischen Mittelwerten der erklärenden Variablen korreliert ist. |
| **W** | Die Nullhypothese $\beta \geq \beta_k$ wird am 5 Prozent Signifikanzniveau bei 500 Freiheitsgraden verworfen, wenn der empirische *t*-Wert größer als 1,645 ist. |
| **F** | Die Inkonsistenz eines Steigungsparameters führt nicht zur Inkonsistenz der gleichzeitig geschätzten Regressionskonstanten. |
| **W** | Mit dem RESET Test können Probleme bei der Spezifikation des Regressionsmodells aufgedeckt werden. |
| **W** | Monte Carlo Studien stützen sich auf per Zufallsziehung generierte Daten. |
| **W** | Der Durbin-Watson-Test auf AR(1) Fehler ist nicht anwendbar, wenn das Modell verzögerte endogene Variablen als erklärende Variablen aufweist. |
| **W** | Bei einem gegebenen Schätzverfahren und Modell können sich für unterschiedliche Stichproben verschiedene Schätzwerte ergeben. |
| **F** | Mit Hilfe des Goldfeld-Quandt Tests lassen sich fixed- und random effects-Schätzer gegeneinander testen. |

| | |
|---|---|
| **F** | Bei Heteroskedastie ist der KQ-Schätzer verzerrt. |
| **W** | Berücksichtigt man im Modell zu viele erklärende Variablen, so steigt die Streuung der geschätzten Parameter. |
| **F** | Die Dichtefunktion der *t*-Verteilung hat ihr Maximum bei 1,96. |
| **F** | Der KQ-Schätzer erzeugt für Steigungsparameter aber nicht für die Regressionskonstante Zufallsvariablen. |
| **F** | Bei GLS (generalized least squares) Schätzern sind *t*- und F-Tests nicht einmal approximativ anwendbar. |
| **F** | Simultane Gleichungssysteme in struktureller Form können mit dem KQ-Schätzer konsistent geschätzt werden, solange jede Gleichung einzeln betrachtet wird. |
| **W** | Das BIC Kriterium fällt umso günstiger aus, je kleiner die Fehlerquadratsumme bei gegebener Parameter- und Beobachtungszahl ist. |
| **W** | Bei strikter Exogenität der erklärenden Variablen ist der fixed effects-Schätzer konsistent. |
| **F** | Der White-Schätzer benutzt gruppenspezifische Mittelwerte, um heteroskedastische Koeffizienten zu berechnen. |
| **W** | Das angepasste $R^2$ einer Schätzung kann bei Hinzufügen von erklärenden Variablen sinken. |
| **W** | Mithilfe eines linearen Regressionsmodells lassen sich Elastizitäten schätzen. |

**Aufgabe 6:** **[15 Punkte]**

Wahr oder falsch? Erläutern Sie stichwortartig 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.

| | |
|---|---|
| **F** | Die Zielfunktion $S\left(\tilde{\beta}\right)=\sum_{i=1}^{N}\left(y_i-g\left(x_i,\tilde{\beta}\right)\right)^2$ ergibt keinen KQ-Schätzer.<br>→ ist Zielfunktion der nicht-linearen KQ-Schätzung |
| **W** | Wenn a ein konsistenter Schätzer für α ist, dann ist $\sqrt{a}$ ein konsistenter Schätzer für $\sqrt{\alpha}$ .<br>→ da, wenn plim a = α, dann plim $\sqrt{a}$ = $\sqrt{\alpha}$ (da plim g(a) = g(α )) |
| **W** | Die Berücksichtigung einer verzögerten endogenen Variable als erklärende Variable kann zur Inkonsistenz des KQ-Schätzers führen.<br>→ wenn die Störterme autokorreliert sind, z.B. AR(1)-Fehler |
| **W** | Der PE Test überprüft die Spezifikation des Modells.<br>→ unterscheidet zwischen linearer und logarithmierter Form |
| **F** | Unter den Annahmen des Gauss-Markov Theorems hat der KQ-Schätzer die kleinste Varianz unter allen linearen Schätzern.<br>→ unter allen linearen _und unverzerrten_ Schätzern |

| | |
|---|---|
| **F** | Der iterative Cochrane-Orcutt Schätzer wird in zwei Stufen geschätzt.<br>→ iteriert eventuell öfter als zweimal |
| **F** | Unter den Annahmen des Gauss-Markov Theorems sind die mittels KQ-Schätzer erzeugten Vorhersagen verzerrt, aber konsistent.<br>→ unverzerrt, da Parameter unverzerrt |
| **W** | Die Endogenität einer erklärenden Variablen lässt sich mit Hilfe eines *t*-Tests überprüfen.<br>→ (Durbin-Wu-)Hausman-Test geht praktisch so vor |
| **W** | Bei KQ-Schätzern sind *t*- und F-Tests asymptotisch zutreffend, selbst wenn der Störterm der Regression nicht normalverteilt ist.<br>→ da Störterm asymptotisch normalverteilt |
| **W** | random effects-, between- und within-Schätzer stehen in einem linearen Zusammenhang.<br>→ weil z.B. der random effects-Schätzer als gewichteter Mittelwert von within und between-Schätzer bestimmt werden kann |

**Aufgabe 7:** **[6 Punkte]**

Das lineare Modell $y_i = x_i'\beta + \varepsilon_i$ kann genutzt werden, um an konkreten Ausprägungen von $x_0$ Vorhersagen von $y$ zu generieren. Leiten Sie die Varianz des Vorhersagefehlers ab und zeigen Sie, in welcher Form diese von $\sigma^2$ abhängt.

- Der Vorhersagefehler lautet

$$\begin{aligned} y_0 - \hat{y}_0 &= x_0'\beta + \varepsilon_0 - x_0'b \\ &= \varepsilon_0 + x_0'\left(\beta - b\right) \end{aligned}$$

- Die Varianz des Vorhersagefehlers ergibt sich als:

$$\begin{aligned} Var\left\{y_0 - \hat{y}_0\right\} &= Var\left\{\varepsilon_0 + x_0'\left(\beta - b\right)\right\} \\ &= \sigma^2 + Var\left\{x_0'\left(\beta - b\right)\right\} + 2 \cdot Cov\left(\varepsilon_0, x_0'\left(\beta - b\right)\right) \\ &= \sigma^2 + x_0'Var\left\{-b\right\}\, x_0 + 2 \cdot \underbrace{Cov\left(y_0 - x_0'\beta, x_0'\left(\beta - b\right)\right)}_{\substack{x_0' \text{ und } \beta \text{ nicht stochastisch, also: Cov(...)} = 0, \\ \text{wenn } y_0 \text{ bzw. } \varepsilon_0 \text{ und b unkorreliert}}} \\ &= \sigma^2 + x_0'\left(+1\right)\underbrace{Var\left\{b\right\}}_{=\sigma^2\left(X'X\right)^{-1}} x_0 \\ &= \sigma^2 + \sigma^2 x_0'\left(X'X\right)^{-1} x_0 \\ &= \sigma^2\left(1 + x_0'\left(X'X\right)^{-1} x_0\right) \end{aligned}$$

und ist somit eine lineare Funktion von $\sigma^2$.