v. d. Lippe, v. d. Locht (download Nr. 8)

Lösungen zu den "Aufgaben in der Ökonometrieklausur in Duisburg"[1]

Aufgabe 1

a) $\hat{\alpha}=0,\ \hat{\beta}=\frac{s_{xy}}{s_x^2}=\frac{0{,}02}{0{,}01}=2,\ s_y^2=0{,}06,\ r^2=R^2=2/3$

b) R^2 ist ein Varianzverhältnis und Maß für die Güte der Anpassung; r ist invariant gegenüber linearen Transformationen

c) $\frac{S_{\hat{y}\hat{y}}}{T}=r^2s_y^2=0{,}04,\ S_{\hat{y}\hat{y}}=4,\ S_{yy}=Ts_y^2=6 \rightarrow S_{\hat{u}\hat{u}}=2$ geschätzte Varianz der Störgröße $\frac{S_{\hat{u}\hat{u}}}{T-2}=\frac{2}{98}$; Erwartungstreue, Konfidenzintervall mit χ^2-Verteilung berechnen.

Aufgabe 2

a) Der Vektor **b** der KQ-Schätzer für α, β_1 und β_2 ist

$$\mathbf{b}=\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix}=(\mathbf{X'X})^{-1}\mathbf{X'y}=\begin{pmatrix}0.809 & -0.074 & -0.176\\ -0.074 & 0.029 & -0.029\\ -0.176 & -0.029 & 0.154\end{pmatrix}\cdot\begin{pmatrix}45\\ 227\\ 101\end{pmatrix}=\begin{pmatrix}1{,}831\\ 0{,}324\\ 1{,}051\end{pmatrix}$$

b) Da T = 8 ist $\hat{\sigma}^2=\mathbf{\hat{u}'\hat{u}}/(T-3)=3.39/5=0{,}678$; die Varianz-Kovarianz-Matrix lautet

$$\hat{\sigma}^2(\mathbf{X'X})^{-1}=0{,}678\cdot\begin{pmatrix}0.809 & -0.074 & -0.176\\ -0.074 & 0.029 & -0.029\\ -0.176 & -0.029 & 0.154\end{pmatrix}=\begin{pmatrix}0{,}5485 & -0{,}0502 & -0{,}1193\\ -0{,}0502 & 0{,}0197 & -0{,}0197\\ -0{,}1193 & -0{,}0197 & 0{,}1044\end{pmatrix}$$

Interpretation: Varianzen $\hat{\sigma}^2_{\hat{\alpha}}=0{,}5485,\ \hat{\sigma}^2_{\hat{\beta}_1}=0{,}0197,\ \hat{\sigma}^2_{\hat{\beta}_2}=0{,}1044,$

Kovarianzen $\hat{\sigma}_{\hat{\alpha}\hat{\beta}_1}=-0{,}0502$ usw.

Aufgabe 3

Im linearen Regressionsmodell $y_t=\alpha+\beta_1x_t+u_t$ liegt Autokorrelation 1. Ordnung vor, d.h. es gilt $u_t=\rho u_{t-1}+e_t$.

a) Annahmen über den Störterm $e_t \sim N(0,\sigma_e^2)$ und $cov(e_t,e_\tau)=0$ für alle t und τ, also die Annahmen B1 bis B4 sollen für e (natürlich nicht für u) gelten

b) $\hat{\rho}=\frac{\sum_{t=2}^{T}\hat{u}_{t-1}\hat{u}_t}{\sum_{t=2}^{T}\hat{u}_{t-1}^2}=\frac{\sum_{t=2}^{T}\hat{u}_{t-1}\hat{u}_t}{\sum_{t=1}^{T}\hat{u}_t^2-\hat{u}_T}=\frac{21{,}6}{40-4}=0{,}6$ also positive Autokorrelation

c) VKQ und GVKQ

d) Durbin-Watson-Test: Vgl. Download Nr. 4

[1] Übungs- bzw. Klausuraufgaben gemäß download Nr. 6

Aufgabe 4

1 F	**2 R**	**3 R**	**4 F**	**5 R**
6 R	**7 R**	**8 F**	**9 R**	**10 R**

Die erste Frage mag etwas missverständlich sein (in solchen Fällen ruhig verbal Anmerkungen machen): im Modell der Korrelationsanalyse ist y und x eine Zufallsvariable. Im Modell der Regressionsanalyse – was hier (und in der Ökonometrie allgemein) gemeint ist wird r^2 als Maß der Güte der Anpassung berechnet und hier ist y aber *nicht* x eine Zufallsvariable (x ist wegen Annahme C2 keine Zufallsvariable)

Aufgabe 5 (EViews-Output)

a) $\hat{y}_t = 102192{,}4 - 9074{,}674\, N_t + 0{,}354668\, P_t + 1{,}287923\, I_t$ Alle sind plausibel denn

$\hat{\beta}_N < 0$ (- 9074): je mehr Konkurrenten in der Umgebung desto weniger Umsatz y

$\hat{\beta}_P > 0$ (+ 0,355): je mehr Einwohner desto mehr Umsatz

$\hat{\beta}_I > 0$ (+ 1,288) je größer das durchschnittliches Haushaltseinkommen der Einwohner, desto größer der Umsatz des Woody- Restaurants

b) RSS = residual sum of squares = $S_{\hat{u}\hat{u}}$ = 6.13E+09, d.h. 6130000000 (neun Stellen nach 6)

ESS = explained sum of squares = $S_{\hat{y}\hat{y}}$ = 9929450000 Diese Größe muss aus dem Bestimmtheitsmaß R^2 = ESS/TSS und TSS = ESS + RSS berechnet werden

T = 33, K = 3, die F-Statistik ist demnach $\dfrac{ESS/K}{RSS/(T-K-1)} = \dfrac{9929450000/3}{6130000000/29} = 15{,}65$

Interpretation: erklärte Varianz durch Residualvarianz; es gibt auch einen Zusammenhang zwischen F und R^2 = ESS/TSS, denn $F = \dfrac{R^2}{1-R^2} \cdot \dfrac{T-K-1}{K}$

und zwischen F bei restringiertem und vollem Modell (vgl. Aufgabe 8) und der korrigierten Bestimmtheit[2] .

Da F = 15,65 F-verteilt ist mit K = 3 und T-K-1 = 33 - 3 - 1 = 29 Freiheitsgraden, kann man mit dieser Statistik auch die Hypothese der Irrelevanz aller Regressoren testen, also die H_0: $\beta_N = \beta_P = \beta_I = 0$. Der Tabellenwert bei einem Signifikanzniveau von 5% (1%) beträgt hier 2,92 (4,51), also gilt: H_0 wird verworfen, sowohl auf dem 5% als auch dem 1% Niveau. Die Regressoren sind nicht *alle zusammengenommen* irrelevant.

c) 90% Konfidenzintervall für β_P[3]: es ist gem. EViews-Output $\hat{\beta}_P$ = und $\hat{\sigma}_P$ = 0.072681 der Tabellenwert der t-Verteilung bei 10% Signifikanzniveau (zweiseitig) und T - K - 1 = 29 Freiheitsgraden ist 1,699. Für die Grenzen des Konfidenzintervalls erhält man also 0.354668 ± 1,699·0.072681 und damit 0,2312 und 0,4782.

Aufgabe 6

a) Beim Tabellenwert 2,59 für die $t_{T-2} = t_{376}$-Verteilung $\hat{\beta}_2$, nicht aber $\hat{\beta}_1$ signifikant von Null verschieden (bei einem Signifikanzniveau von 5 % [Tabellenwert 1,96] wären beide signifikant).

---

[2] Vgl. L von Auer Abschnitt 13.2.5.

[3] Soll I die Variable P sein, wie es in der Aufgabenstellung heißt, dann ist entsprechend zu rechnen mit 1.287923 (statt 0.354668) und 0.543294 (statt 0.072681).

b) zweiseitiges Konfidenzintervall (99%) für β_1 und β_2

$$\text{für } \beta_1: 0{,}1561 \pm 2{,}59 \cdot \underbrace{\frac{0{,}1561}{1{,}9839}}_{0{,}07868} = 0{,}1561 \pm 0{,}20379 \rightarrow \text{von} -0{,}04769 \text{ bis } 0{,}35988$$

$$\text{für } \beta_2: 0{,}1344 \pm 2{,}59 \cdot \underbrace{\frac{0{,}1344}{2{,}8358}}_{0{,}047394} = 0{,}1344 \pm 0{,}12275 \rightarrow \text{von } \underbrace{0{,}011649}_{\text{größer als } 0} \text{ bis } 0{,}25715$$

c) siehe oben; bei Signifikanz (Ablehnung von H_0) überdeckt das Intervall den Wert 0 nicht, bei Annahme von H_0 überdeckt es den Wert gem. Nullhypothese (also meist $\beta = 0$).

d) Das korrigierte Bestimmtheitsmaß R^2_{adj} berücksichtigt auch T und K und nimmt nicht mit zunehmender Zahl K der Regressoren notwendig zu (bzw. besser: nicht ab).

Aufgabe 7

	Schätzeigenschaft	notwendige Voraussetzung
$\hat{\beta} = f(y_1, y_2, ..., y_r)$	Linearität	a2, Methode der kleinsten Quadrate
$E(\hat{\beta}) = \beta$	Erwartungstreue	b1
$V(\hat{\beta}_{KQ}) < V(\hat{\beta}_{AS})$	Effizienz	b1, b2, b3
$\lim_{n\to\infty} V(\hat{\beta}) = 0$	Konsistenz	b1 – b3 und auch c1

a) Die Stichprobenverteilung von $\hat{\beta}$ ist die t-Verteilung mit T – k Freiheitsgraden wenn u normalverteilt ist, aber die Varianz von u (also σ^2) nicht bekannt ist und mit $\hat{\sigma}^2$ geschätzt werden muss; k = K + 1 ist die Anzahl der zu schätzenden Parameter einschließlich des Absolutglieds α bei K Regressoren $x_1, x_2, \dots x_K$. Bei einfacher Regression gilt K = 1 also k = 2, zu schätzen sind α und β. Die Gestalt der Stichprobenverteilung ist wichtig um Intervallen Wahrscheinlichkeiten zuordnen zu können (einfach gesprochen: um zu wissen in welcher Tabelle man nachschauen muss).

b) Das "Gauss Markoff Theorem" gilt für die Methode der kleinsten Quadrate (KQ Schätzverfahren). Unter den bekannten Voraussetzungen ist ein KQ-Schätzer (oder OLS- Schätzer) der **b**este (minimale Varianz) unter den **l**inearen (libnear in u und damit auch in y) **u**nverzerrten (unbiased) Schätzern (**e**stimators). Andere Schätzverfahren: Maximum Likekihood , orthogonale Regression, verschieden Verfahren der robusten Regression usw.

c)

Zu wenige Regressoren	Erwartungstreue und Effizienz
Zu viele (auch irrelevante) Regressoren	Nur Effizienz

Aufgabe 8

a) $F = \dfrac{\text{(zusätzliche) erklärte Varianz}}{\text{Residualvarianz}}$ ist F_{v_1, v_2} verteilt. T = 50 Beobachtungen.

Die Residualvariation beim restringierte Modell mit nur 3 Regressoren X_3, X_5 und X_6 ist $S^0_{\hat{u}\hat{u}}$ und die im vollen Modell $\hat{y} = \alpha + \beta_1 x_1 + ... + \beta_6 x_6$ mit K = 6 Regressoren ist $S_{\hat{u}\hat{u}}$

Im Vergleich dazu ist sie (für die erkläre Variation gilt dort auch $S_{\hat{y}\hat{y}} = S_{yy} - S_{\hat{u}\hat{u}}$)

F-Test für die Hypothese:

- der Irrelevanz *aller* Regressoren $X_1,\ldots X_6$ also von $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$

  $F = \dfrac{S_{\hat{y}\hat{y}}/K}{S_{\hat{u}\hat{u}}/(T-K-1)} = \dfrac{S_{\hat{y}\hat{y}}/6}{S_{\hat{u}\hat{u}}/(50-6-1)}$ ~ $F_{6,43}$ dann mit Tabellenwert vergleichen, wenn dieser kleiner ist als der errechnete F-Wert wird H_0 abgelehnt

- der Signifikanz (zusätzlicher Erklärungsbeitrag) der zu X_3, X_5 und X_6 *hinzukommenden* Regressoren X_1, X_2 und X_4 (was $H_0: \beta_1 = \beta_2 = \beta_4 = 0$ getestet wird) L = 3; Prüfgröße $F = \dfrac{\left(S^0_{\hat{u}\hat{u}} - S_{\hat{u}\hat{u}}\right)/L}{S_{\hat{u}\hat{u}}/(T-K-1)} = \dfrac{\left(S^0_{\hat{u}\hat{u}} - S_{\hat{u}\hat{u}}\right)/3}{S_{\hat{u}\hat{u}}/43}$ ~ $F_{3,43}$. Man beachte, dass gilt $S^0_{\hat{u}\hat{u}} - S_{\hat{u}\hat{u}} = S_{\hat{y}\hat{y}} - S^0_{\hat{y}\hat{y}}$.

b) Welche Verteilung

	eine Linearkombination	mehrere Linearkombinationen
Wahre Varianz σ^2 bekannt		χ^2-Verteilung
Varianz zu schätzen mit $\hat{\sigma}^2$		F-Verteilung

Aufgabe 9

E (u_t) = = 0, Korrelogramm hat bei $\rho < 0$ alternierendes Vorzeichen, $\hat{\sigma}^2$ ist verzerrt für die Schätzung von $\hat{\beta}$ (bzw. für das Testen von Hypothesen über β) gilt weiter erwartungstreu aber nicht mehr effizient Intervalle zu breit, Test wertlos Gl. 18.6 bei v. Auer, zu 4 (Abhilfe durch)Newey-West Schätzung der var ($\hat{\beta}_i$).

Aufgabe 10 (zugleich quasi ein Lehrtext)

Die folgende Lösungsangabe ist sehr ausführlich, weil die Zusammenhänge vielleicht etwas kompliziert sind.

a) Gl. 1 und Gl. 2 sind beobachtungsäquivalent, d.h. sie haben die gleiche Gestalt. Bei einer Schätzung der Regressionsfunktion $\hat{y}_t = c_0 + c_1 x_t$ kann man nicht einfach $c_0 = \alpha$ und $c_1 = \beta$ setzen und sagen, man habe die Angebotsfunktion (AA') geschätzt, selbst dann nicht wenn c_1 positiv ist, was dafür sprechen würde, dass es die (steigende) Angebotskurve und nicht die (fallende) Nachfragekurve NN' ist). Man könnte mit c_0 und c_1 auch γ und δ geschätzt haben und damit die Nachfragekurve, bzw. keine von beiden (also weder Angebots- noch Nachfragekurve).[4] x_t und u_t und x_t und v_t sind auch nicht unkorreliert, d.h. der Preis kann nicht als exogene Variable betrachtet werden.
Das Gleichungssystem der Aufgabe entspricht dem **partiellen Marktgleichgewicht bei v. Auer (Abschn. 23.5.3)** mit $y_{St} = g^A_t$ und $y_{Dt} = g^N_t$ sowie $x_t = p_t$. Fügt man dem System die Variable y (Einkommen) in der Aufgabenstellung x_{2t} (in Gl. 2) zu, so entsteht aus der strukturellen Form (Gleichungen der Aufgabenstellung in v. Auers Symbolik mit erweiterter Gl. 2 {Nachfragefunktion})
$g^A_t = g_t = \alpha + \beta p_t + u_t, \quad g^N_t = g_t = \gamma + \delta_1 p_t + \delta_2 y_t + v_t$

[4] Mehr dazu (Interpretation mit Konstanz bzw. Veränderlichkeit der Angebots- bzw. Nachfragekurve) in einem download zum Identifikationsproblem (auch auf der Seite der übrigen downloads [auch des hier vorliegenden] zur Vorlesung).

Die reduzierte Form erhält man durch Auflösen des Systems nach den beiden endogenen Variablen g_t und p_t. Konkret geht das so: Gleichsetzen der beiden Gleichungen von g_t^A und g_t^N liefert (23.23a) und das daraus bestimmte p_t eingesetzt in die NN'funktion liefert 23.23b). Man erhält aus der AA'funktion

(23.23a) $\quad p_t = -\frac{\alpha}{\beta} + \frac{1}{\beta} g_t + \left(-\frac{1}{\beta} u_t \right)$ und aus der NN'funktion

(23.23b) $\quad g_t = \gamma + \delta_1 p_t + \delta_2 y_t + v_t$

In dieser Form wird auch deutlich, dass das Modell interdependent ist. Exogene Variablen sind $x_{0t} \equiv 1$ (verantwortlich für das Absolutglied α bzw. γ) und y_t. Die reduzierte Form des Modells 23.23 ist demnach

$$p_t = \frac{\gamma - \alpha}{\beta - \delta_1} + \frac{\delta_2}{\beta - \delta_1} y_t + \frac{v_t - u_t}{\beta - \delta_1} = \pi_1 + \pi_2 y_t + u_t^* \text{ und}$$

$$g_t = \frac{-\alpha\delta_1 + \gamma\beta}{\beta - \delta_1} + \frac{\beta\delta_2}{\beta - \delta_1} y_t + \frac{\beta v_t - \delta_1 u_t}{\beta - \delta_1} = \pi_3 + \pi_4 y_t + v_t^*$$

Man kann aus den Koeffizienten π der reduzierten Form die Koeffizienten der strukturellen Form bei der Angebotskurve (α und β) aber nicht bei der Nachfragekurve (γ, δ_1,δ_2) zurückrechnen (z.B. ist $\beta = \frac{\pi_4}{\pi_2}$): AA' ist identifizierbar, NN' nicht. Man kann mit vier Koeffizienten $\pi_1, \dots, \pi_4$ der reduzierten Form nicht fünf Koeffizienten der strukturellen Form (α,β,γ, δ_1,δ_2) bestimmen.

Arbeitet man auch mit Kosten ($k_t = x_{1t}$ in der Aufgabenstellung) als exogene Variable also dem Modell $g_t^A = \alpha + \beta_1 p_t + \beta_2 k_t + u_t \quad g_t^N = \gamma + \delta_1 p_t + \delta_2 y_t + v_t$ dann ist AA' und NN' identifizierbar.

b) $p_t = \frac{\gamma - \alpha}{\beta - \delta} + \frac{v_t - u_t}{\beta - \delta} = \pi_1 + u_t^*$ und $p_t = \frac{\gamma\beta - \alpha\delta}{\beta - \delta} + \frac{\beta v_t - u_t \delta}{-\delta} = \pi_2 + v_t^*$ (Das Modell ist unteridentifiziert, man kann nicht mit π_1 und π_2 die Parameter α, β, γ und δ bestimmen.

ILS (indirect least squares)[5] bedeutet OLS statt auf die strukturelle Form auf die reduzierte Form anzuwenden und mit den Koeffizienten π_i, der reduzierten Form diejenigen der strukturellen Form zu berechnen.

Beispiele für Multiple Choice Fragen (nur ein Kreuz!)

Ein nicht-konstruktiver Test : **C**

Die Anwendbarkeit des Durbin-Watson-Tests setzt voraus…: **B**

Mit dem Jarque Bera (JB) Test prüft man… **C**

Dass in einer Gleichung die Störgröße mit dem Regressor korreliert ist… **C**

Ein asymptotisch erwartungstreuer Schätzer…**C**

[5] IKQ bei von Auer.

Ein erwartungstreuer Schätzer hat eine kleinere Varianz … **B**

Ein konsistenter Schätzer hat eine kleinere Varianz als ein verzerrter Schätzer **B**

Bei einem überladenen (overfitted) Modell gilt

	$\mathrm{var}(\hat{\beta}_t^*) > \mathrm{var}(\hat{\beta}_t)$	$\mathrm{var}(\hat{\beta}_t^*) = \mathrm{var}(\hat{\beta}_t)$
$E(\hat{u}_t^*) \neq E(\hat{u}_t) = 0$		
$E(\hat{u}_t^*) = E(\hat{u}_t) = 0$	b	

Die Auswahl sollte nicht zu dem Schluss verleiten, dass die Antwortmöglichkeiten A und D nicht vorkommen.