Klausur zu Ökonometrie (Master)

Technische Universität Dortmund | Fakultät Wirtschaftswissenschaften | 4. Oktober 2022

Bitte tragen sie ihre Daten sorgfältig und leserlich ein:

Matrikelnummer ☐☐☐☐☐☐ Nachname ____________________

Studiengang ____________________ Vorname ____________________

Bearbeitungshinweise:

Diese Klausur besteht aus fünf Aufgaben, welche alle zu bearbeiten sind.

Bitte verwenden sie einen Kugelschreiber oder nicht zu starken Filzstift.

Nicht-programmierbare Taschenrechner sind als Hilfsmittel für diese Klausur zugelassen.

Für jede der fünf Aufgaben sind maximal je 18 Punkte zu erreichen.

Bei 36 von maximal 90 erreichbaren Punkten ist die Klausur bestanden.

Die Bearbeitungszeit beträgt 90 Minuten.

Viel Erfolg!

Vom Prüfer auszufüllen:

Punkte Aufgabe 1 ☐ / 18

Punkte Aufgabe 2 ☐ / 18

Punkte Aufgabe 3 ☐ / 18

Punkte Aufgabe 4 ☐ / 18

Punkte Aufgabe 5 ☐ / 18

Gesamtpunkte ☐ / 90 **Note:**

Aufgabe 1

a) Benennen sie die Annahmen MLR 1, MLR 3, MLR 4 und MLR 5 aus der Vorlesung.

MLR 1 $\quad y = X\beta + U \quad$ Linear in β

MLR 3 $\quad rk(X) = k+1 \quad$ alle Regressoren linear unabhängig

MLR 4 $\quad E[U \mid X] = 0 \quad$ strikte Exogenität (der Regressoren)

MLR 5 $\quad \Sigma(U) = \sigma^2 \cdot I \quad$ $var(u_i) = \sigma^2 \; \forall i$, $cov(u_i, u_j) = 0 \; \forall i \neq j$ $\quad$ Störterme homoskedastisch & seriell unkorreliert

b) Erklären sie die Bedeutung des Akronyms BLUE.

Estimator (Schätzer)

Unbiased (unverzerrt) $\quad E[\hat{\beta}] = \beta$ für alle $\beta \in \mathbb{R}^{k+1}$

Linear $\quad \hat{\beta}$ ist linear in y, $\hat{\beta} = A \cdot y$

Best (effizient) $\quad \Sigma(\hat{\beta}) \leq \Sigma(\tilde{\beta}) \quad \forall \hat{\beta}, \tilde{\beta}$ linear & unverzerrt

c) Geben sie einen erwartungstreuen Schätzer $\hat{\sigma}^2$ für die Varianz σ^2 der Störterme u_i, $i = 1, \ldots, n$ an. Nehmen sie hierbei an, dass MLR 1, MLR 3, MLR 4 und MLR 5 gelten. Ist $\hat{\sigma}^2$ BLUE? Begründen sie ihre Antwort.

$$\hat{\sigma}^2 = \frac{1}{n-k-1} \hat{U}'\hat{U} = \frac{1}{n-k-1} \sum_{i=1}^{n} \hat{U}_i^2 = \frac{1}{n-k-1} SSR$$

Ist $\hat{\sigma}^2$ BLUE? $\quad$ E ✓ $\quad$ U ✓ $\quad$ L: nein!

$\hat{\sigma}^2$ ist quadratisch in y

Aufgabe 2

Gegeben sei folgende Ausgabe einer Regressionsanalyse:

Modell 1: KQ, benutze die Beobachtungen 1–1388 ($n = 1191$)
Fehlende oder unvollständige Beobachtungen entfernt: 197
Abhängige Variable: bwght

$\hat{\beta}_1$

	Koeffizient	Std. Fehler	t-Quotient	p-Wert
const	114,524	3,72845	30,72	0,0000
cigs	−0,595936	0,110348	−5,401	0,0000
parity	1,78760	0,659406	2,711	0,0068
faminc	0,0560414	0,0365616	1,533	0,1256
motheduc	−0,370450	0,319855	−1,158	0,2470
fatheduc	0,472394	0,282643	1,671	0,0949

Teststatistik für globalen F-Test

Mittel abhängige Var.	119,5298	Stdabw. abhängige Var.	20,14124
Summe quad. Residuen	464041,1	Stdfehler Regression	19,78878
R^2	0,038748	Korrigiertes R^2	0,034692
$F(5, 1185)$	9,553500	P-Wert(F)	5,99e−09 $< 0{,}01$
Log-Likelihood	−5242,220	Akaike-Kriterium	10496,44
Schwarz-Kriterium	10526,94	Hannan–Quinn	10507,93

Die Variablen haben folgende Bedeutungen:

bwght: Geburtsgewicht in Unzen (1 Unze ≈ 28.3 Gramm), *cigs*: Zigaretten pro Tag während Schwangerschaft, *parity*: 1., 2., 3., ... Kind, *faminc*: Familieneinkommen in 1000$ in 1988, *motheduc* / *fatheduc*: #Jahre Bildung der Mutter / des Vaters

a) Betrachten sie folgende Aussage: „Zwei zusätzliche Zigaretten pro Tag während der Schwangerschaft gehen einher mit einer durchschnittlichen Reduktion des Geburtsgewichts um eine Unze, wenn alle anderen Faktoren gleich bleiben.“ Stellen sie die entsprechende Nullhypothese auf und erläutern sie das anzuwendende Testverfahren. Zu welchem Ergebnis kommen sie? $H_0 \quad 2\beta_1 = -1 \quad \Leftrightarrow \quad \beta_1 = -\frac{1}{2}$

$t = \frac{\hat{\beta}_1 - (-\frac{1}{2})}{se(\hat{\beta}_1)} = \frac{-0,6+0,5}{0,11} = \frac{-0,1}{0,11} = -0,91 \qquad c_{t_{1185}, \underbrace{1-\frac{0,1}{2}}_{0,95}} = 1,645$, $|t| < c \Rightarrow H_0$ nicht verworfen.

b) Führen sie einen globalen F-Test durch und erläutern sie dabei ihr Vorgehen und ihre Schlussfolgerung.

$F = 9,55$ (siehe Output), P-Wert $< 1\%$ $\Rightarrow$ $H_0: \beta_1 = 0, \beta_2 = 0 \cdots \beta_5 = 0$ verworfen (zu $\alpha = 1\%$)

c) Welche Bedeutung hat der p-Wert von 0,1256 für *faminc*?

Mit W'keit 12,56% würden wir $H_0: \beta_3 = 0$ zurückweisen, falls sie wahr ist.
(W'keit für Fehler 1. Art)

SSR_r

d) sie führen eine zweite Regression durch, bei welcher sie auf die Variablen *motheduc* und *fatheduc* verzichten. Die Summe der quadrierten Residuen des restringierten Modells beträgt 465166,8. Benutzen sie diese Informationen und eine geeignete Tabelle im Anhang dieser Klausur um zu testen, ob die Koeffizienten von *motheduc* und *fatheduc* gemeinsam signifikant von null verschieden sind. Erläutern sie hierbei ihr Vorgehen und ihre Schlussfolgerung.

F-Test! F-Statistik $F = \frac{(SSR_r - SSR_{ur})/2}{SSR_{ur}/(n-k-1)} = \frac{(465166,8 - 464041,1)/2}{464041,1/1185}$

$c_{F_{2,1185}, 10\%} = 4,61 \qquad = \frac{562,9}{391,6} = 1.437 < c$

$\Rightarrow H_0$ nicht verwerfen

Aufgabe 3

Betrachten sie ein lineares Regressionsmodell unter den Annahmen MLR 1 bis MLR 6:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i \ , \ i = 1, \ldots, n$$

$$-2\beta_3 x_{i2} + 2\beta_3 x_{i2}$$

Es seien $\hat{\beta}_0, \ \hat{\beta}_1, \hat{\beta}_2$ und $\hat{\beta}_3$ die OLS-Schätzer für die unbekannten Parameter $\beta_0, \beta_1, \beta_2$ und β_3.

a) Benennen sie $var(\hat{\beta}_2 - 2\hat{\beta}_3)$ in Ausdrücken $var(\hat{\beta}_2)$, $var(\hat{\beta}_3)$ und $cov(\hat{\beta}_2, \hat{\beta}_3)$.

$$Var(\hat{\beta}_2 - 2\hat{\beta}_3) = Var(\hat{\beta}_2) - 2\,cov(\hat{\beta}_2, 2\hat{\beta}_3) + Var(2\hat{\beta}_3) = Var(\hat{\beta}_2) - 4\,Cov(\hat{\beta}_2, \hat{\beta}_3) + 4\,Var(\hat{\beta}_3)$$

$$\left(\widehat{Var}(\hat{\beta}_2 - \hat{\beta}_3) = \ \wedge \quad \wedge \quad \wedge \right)$$

b) Benennen sie die t-Statistik für die Hypothese H_0 : $\beta_2 - 2\beta_3 = 2$. Verwenden sie hierfür die Ausdrücke $\widehat{var}(\hat{\beta}_2)$, $\widehat{var}(\hat{\beta}_3)$ und $\widehat{cov}(\hat{\beta}_2, \hat{\beta}_3)$.

$$t = \frac{\hat{\beta}_2 - 2\hat{\beta}_3 - 2}{\sqrt{\widehat{Var}(\hat{\beta}_2) - 4\,\widehat{Cov}(\hat{\beta}_2, \hat{\beta}_3) + 4\,\widehat{Var}(\hat{\beta}_3)}}$$

c) Definieren sie $\theta = \beta_2 - 2\beta_3$ und $\hat{\theta} = \hat{\beta}_2 - 2\hat{\beta}_3$. Stellen sie eine Regressionsgleichung mit β_0, β_1, θ und β_3 auf, welche ihnen erlaubt, $\hat{\theta}$ und den Standardfehler von $\hat{\theta}$ aus den Regressionsergebnissen der Software abzulesen. Definieren sie ggf. einen weiteren Regressor $z_i, i = 1, \ldots, n$.

$$y_i = \beta_0 + \beta_1 x_{i1} + \underbrace{(\beta_2 - 2\beta_3)}_{\theta} x_{i2} + \beta_2 \underbrace{(x_{i3} + 2x_{i2})}_{z_i} + u_i$$

Software: $\hat{\theta}$, $se(\hat{\theta}) = \sqrt{\widehat{Var}(\hat{\beta}_2 - 2\hat{\beta}_3)}$

Aufgabe 4

Es sei $\mathbf{u}$ ein Zufallsvektor mit n Komponenten, welcher der Normalverteilung unterliege. Der Erwartungswert von $\mathbf{u}$ betrage $E[\mathbf{u}] = \mathbf{0} \in \mathbb{R}^n$ und die Varianz-Kovarianz-Matrix von $\mathbf{u}$ sei durch $var(\mathbf{u}) = \sigma^2\Omega$ gegeben, wobei σ^2 ein positiver Skalar sei und $\Omega \in \mathbb{R}^{n\times n}$ positiv definit symmetrisch sei. Es gelten außerdem die Annahmen MLR 1, MLR 3 und MLR 4.

Es gelte

$$\mathbf{y} = X\boldsymbol{\beta} + \mathbf{u} \ ,$$

wobei $X \in \mathbb{R}^{n\times(k+1)}$ eine bekannte und nicht stochastische Matrix mit vollen Spaltenrang und $\boldsymbol{\beta} \in \mathbb{R}^{k+1}$ ein unbekannter und nicht stochastischer Vektor seien.

a) Geben sie $E[\mathbf{y}]$, $var(\mathbf{y})$ und die Verteilung von $\mathbf{y}$ an.

- $E[\mathbf{y}]$

- $var(\mathbf{y})$

- Verteilung von $\mathbf{y}$:

b) Betrachten sie nun $\tilde{\boldsymbol{\beta}} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\mathbf{y}$. Geben sie $E[\tilde{\boldsymbol{\beta}}]$, $var(\tilde{\boldsymbol{\beta}})$ und die Verteilung von $\tilde{\boldsymbol{\beta}}$ an.

- $E[\tilde{\boldsymbol{\beta}}]$

- $var(\tilde{\boldsymbol{\beta}})$

- Verteilung von $\tilde{\boldsymbol{\beta}}$:

Nehmen sie nun an, dass $\Omega = diag(\boldsymbol{\omega})$, wobei $\boldsymbol{\omega} \in \mathbb{R}^n$ ein nicht-stochastischer Vektor sei. Ω sei also eine ihnen bekannte Diagonalmatrix, deren i-te Komponente auf der Diagonale durch die i-te Komponente ω_i von $\boldsymbol{\omega}$ gegeben sei.

c) Beschreiben sie, wie die Daten $\mathbf{y}$ und X zu transformieren sind, sodass bei gegebenen transformierten Daten $\tilde{\boldsymbol{y}}$ und $\tilde{X}$ der OLS-Schätzer $\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\boldsymbol{y}}$ BLUE ist.

Aufgabe 5

Ihnen liegt der Datensatz wage1 der Vorlesung vor und sie führen eine OLS-Regression mit folgender Ausgabe durch:

Modell 1: KQ, benutze die Beobachtungen 1–526
Abhängige Variable: lwage

	Koeffizient	Std. Fehler	t-Quotient	p-Wert
const	0,216854	0,108595	1,997	0,0464
educ	0,0979356	0,00762240	12,85	0,0000
exper	0,0103469	0,00155514	6,653	0,0000

Mittel abhängige Var.	1,623268	Stdabw. abhängige Var.	0,531538
Summe quad. Residuen	111,3447	Stdfehler Regression	0,461407
R^2	0,249343	Korrigiertes R^2	0,246473
$F(2, 523)$	86,86167	P-Wert(F)	2,68e−33
Log-Likelihood	−338,0094	Akaike-Kriterium	682,0188
Schwarz-Kriterium	694,8147	Hannan–Quinn	687,0290

Die Variablen haben hierbei folgende Bedeutungen: $lwage$: $log(wage)$, wobei $wage$ den durchschnittlichen Stundenlohn bezeichnet. $educ$: Jahre Bildung, $exper$: Jahre Berufserfahrung

a) Definieren sie reine Heteroskedastizität in Bezug auf die Störterme. Welche der Annahmen MLR 1 bis MLR 6 fordern Homoskedastizität der Störterme?

b) Nennen sie eine Eigenschaft des OLS-Schätzers, die auch unter heteroskedastischen Störterme gültig ist und nennen sie eine Methode, welche in Bezug auf den OLS-Schätzer und Heteroskedastizität nicht valide ist.

c) Erläutern sie, wie sie in Bezug auf die obige Regression vorgehen würden, um die Daten auf Homoskedastizität zu testen. Unter welchen Umständen würden sie die Nullhypothese homoskedastischer Störterme verwerfen?

Kritische Werte der t-Verteilung

		Signifikanzniveau				
einseitig:		10%	5%	2,5%	1%	0,5%
zweiseitig:		20%	10%	5%	2%	1%
Freiheits-	1	3,078	6,314	12,706	31,821	63,657
grade	2	1,886	2,920	4,303	6,965	9,925
	3	1,638	2,353	3,182	4,541	5,841
	4	1,533	2,132	2,776	3,747	4,604
	5	1,476	2,015	2,571	3,365	4,032
	6	1,440	1,943	2,447	3,143	3,707
	7	1,415	1,895	2,365	2,998	3,499
	8	1,397	1,860	2,306	2,896	3,355
	9	1,383	1,833	2,262	2,821	3,250
	10	1,372	1,812	2,228	2,764	3,169
	11	1,363	1,796	2,201	2,718	3,106
	12	1,356	1,782	2,179	2,681	3,055
	13	1,350	1,771	2,160	2,650	3,012
	14	1,345	1,761	2,145	2,624	2,977
	15	1,341	1,753	2,131	2,602	2,947
	16	1,337	1,746	2,120	2,583	2,921
	17	1,333	1,740	2,110	2,567	2,898
	18	1,330	1,734	2,101	2,552	2,878
	19	1,328	1,729	2,093	2,539	2,861
	20	1,325	1,725	2,086	2,528	2,845
	21	1,323	1,721	2,080	2,518	2,831
	22	1,321	1,717	2,074	2,508	2,819
	23	1,319	1,714	2,069	2,500	2,807
	24	1,318	1,711	2,064	2,492	2,797
	25	1,316	1,708	2,060	2,485	2,787
	26	1,315	1,706	2,056	2,479	2,779
	27	1,314	1,703	2,052	2,473	2,771
	28	1,313	1,701	2,048	2,467	2,763
	29	1,311	1,699	2,045	2,462	2,756
	30	1,310	1,697	2,042	2,457	2,750
	40	1,303	1,684	2,021	2,423	2,704
	60	1,296	1,671	2,000	2,390	2,660
	90	1,291	1,662	1,987	2,368	2,632
	120	1,289	1,658	1,980	2,358	2,617
	∞	1,282	1,645	1,960	2,326	2,576

Kritische Werte der F-Verteilung zum Signifikanzniveau von 1%

		Anzahl der Restriktionen									
		1	2	3	4	5	6	7	8	9	10
$n-k-1$	10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85
	11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54
	12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30
	13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10
	14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94
	15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80
	16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69
	17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59
	18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51
	19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43
	20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37
	21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31
	22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26
	23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21
	24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17
	25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13
	26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09
	27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06
	28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03
	29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00
	30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98
	40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80
	60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63
	90	6,93	4,85	4,01	3,54	3,23	3,01	2,84	2,72	2,61	2,52
	120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47
	∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32

Kritische Werte der F-Verteilung zum Signifikanzniveau von 5%

		Anzahl der Restriktionen									
		1	2	3	4	5	6	7	8	9	10
$n-k-1$	10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
	11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
	12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
	13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
	14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
	15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
	16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
	17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
	18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
	19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
	20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
	21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
	22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
	23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
	24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
	25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
	26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
	27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
	28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
	29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
	30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
	40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
	60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
	90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94
	120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91
	∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

Kritische Werte der χ^2-Verteilung

		Signifikanzniveau		
		10%	5%	1%
Freiheits-	1	2,71	3,84	6,63
Grade	2	4,61	5,99	9,21
	3	6,25	7,81	11,34
	4	7,78	9,49	13,28
	5	9,24	11,07	15,09
	6	10,64	12,59	16,81
	7	12,02	14,07	18,48
	8	13,36	15,51	20,09
	9	14,68	16,92	21,67
	10	15,99	18,31	23,21
	11	17,28	19,68	24,72
	12	18,55	21,03	26,22
	13	19,81	22,36	27,69
	14	21,06	23,68	29,14
	15	22,31	25,00	30,58
	16	23,54	26,30	32,00
	17	24,77	27,59	33,41
	18	25,99	28,87	34,81
	19	27,20	30,14	36,19
	20	28,41	31,41	37,57
	21	29,62	32,67	38,93
	22	30,81	33,92	40,29
	23	32,01	35,17	41,64
	24	33,20	36,42	42,98
	25	34,38	37,65	44,31
	26	35,56	38,89	45,64
	27	36,74	40,11	46,96
	28	37,92	41,34	48,28
	29	39,09	42,56	49,59
	30	40,26	43,77	50,89