

# Klausur zu Ökonometrie (Master)

Technische Universität Dortmund

Fakultät Wirtschaftswissenschaften

2. Oktober 2023

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.

Alle Antworten sind zu begründen.

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.

Sie finden Tabellen mit kritischen Werten verschiedener Verteilungen im Anhang dieser Klausur.

**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

MLR 1

Es sei das lineare Regressionsmodell  $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{u}$  gegeben, wobei  $\mathbf{y}, \mathbf{u} \in \mathbb{R}^n$ ,  $X \in \mathbb{R}^{n \times (k+1)}$  und  $\boldsymbol{\beta} \in \mathbb{R}^{k+1}$ .

Der Vektor  $\mathbf{y}$  und die Matrix  $X$  seien Beobachtungen, der Vektor  $\boldsymbol{\beta}$  beinhalte unbekannte Parameter und der Vektor  $\mathbf{u}$  sei ein unbeobachtbarer Störterm.

Unter bestimmten Voraussetzungen ist  $\hat{\boldsymbol{\beta}}$  bzw.  $\tilde{\boldsymbol{\beta}}$  BLUE.

- Wofür steht das Akronym BLUE und was bedeutet es? *siehe NT 2022*
- Nennen Sie die Annahmen an das Regressionsmodell, sodass  $\hat{\boldsymbol{\beta}}$  und  $\tilde{\boldsymbol{\beta}}$  „LUE“ sind!
- Unter welcher zusätzlichen Voraussetzung ist  $\hat{\boldsymbol{\beta}}$  BLUE?
- Unter welcher zusätzlichen Voraussetzung ist  $\tilde{\boldsymbol{\beta}}$  BLUE?

b)  $\left. \begin{array}{l} \text{MLR 1} \quad \mathbf{y} = X\boldsymbol{\beta} + \mathbf{u} \\ \text{MLR 3} \quad \text{rk}(X) \\ \text{MLR 4} \quad E[\mathbf{u}|X] = 0 \end{array} \right\} \Rightarrow \hat{\boldsymbol{\beta}}, \tilde{\boldsymbol{\beta}} \text{ LUE}$

c)  $\underbrace{\text{MLR 5} \quad \mathbb{E}(\mathbf{u}|X) = \sigma^2 \mathbf{I}}_{\text{OLS}} \quad \hat{\boldsymbol{\beta}} \text{ LUE} + \text{B}$

d)  $\underbrace{\mathbb{E}(\mathbf{u}|X) = \sigma^2 \Omega, \Omega \text{ pds}}_{\text{GLS}} \quad \tilde{\boldsymbol{\beta}} \text{ LUE} + \text{B}$

## Aufgabe 2

Gegeben sei folgende Ausgabe der Software Gretl für den Datensatz wage1:

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + u$$

Stichprobengröße:  $n = 526$

	Koeffizient	Std. Fehler	t-Quotient	p-Wert
const	0,216854	0,108595	$t_0$	$p_0$
educ	0,0979356	0,00762240	$t_1$	$p_1$
exper	0,0103469	0,00155514	$t_2$	$p_2$
Mittel abhängige Var.	1,623268	Stdabw. abhängige Var.	0,531538	
Summe quad. Residuen	111,3447	Stdfehler Regression	$\hat{\sigma}$	
$R^2$	$R^2$	Korrigiertes $R^2$	$\bar{R}^2$	
$F(2, 523)$	F	P-Wert(F)	$p_F$	

Fehlende Angaben sind durch ein Rechteck mit dem jeweiligen Bezeichner gekennzeichnet.

Die Konstante habe den Index 0, die Regressoren educ und exper die Indizes 1 und 2. Der Regressand sei mit  $y$  bezeichnet.

Wie lauten die folgenden Größen? (Runden Sie je auf drei Nachkommastellen genau.)

Je 0,5 Punkte pro richtiger Antwort:

$\hat{\beta}_0$ :	$\hat{\beta}_1$ :	$\hat{\beta}_2$ :
$se(\hat{\beta}_0)$ :	$se(\hat{\beta}_1)$ :	$se(\hat{\beta}_2)$ :
$\bar{y}$ :	$s_y$ :	SSR:

Je 1 Punkt pro richtiger Antwort:

Tragen Sie die Antworten für die p-Werte bitte folgendermaßen ein:

$0 \leq p_j < 0,01 \rightarrow (***)$ ,  $0,01 \leq p_j < 0,05 \rightarrow (**)$ ,  $0,05 \leq p_j < 0,1 \rightarrow (*)$

$t_0$ :	$t_1$ :	$t_2$ :
$p_0$ :	$p_1$ :	$p_2$ :
SST:	$\hat{\sigma}$ :	$p_F$ :

Je 3 Punkte pro richtiger Antwort:

$R^2$  :  $F$  :

Es sind maximal 18 Punkte für diese Aufgabe zu erreichen.

## Aufgabe 3

Betrachtet sei der Datensatz `wage2`. Hierzu seien drei Modelle mit der Software Gretl geschätzt worden:

$$\text{Modell 1: } \widehat{\text{lwage}} = 5,49670 + 0,0748638 \text{ educ} + 0,0153285 \text{ exper} + 0,0133748 \text{ tenure}$$

(0,11053)            (0,0065124)            (0,0033696)            (0,0025872)

$$\text{Modell 2: } \widehat{\text{lwage}} = 5,49670 + 0,0595353 \text{ educ} + 0,0153285 (\text{exper} + \text{educ}) + 0,0133748 \text{ tenure}$$

(0,11053)            (0,0057910)            (0,0033696)            (0,0025872)

$$\text{Modell 3: } \widehat{\text{lwage}} = 5,49670 + 0,090192 \text{ educ} - 0,0153285 (\text{exper} - \text{educ}) + 0,0133748 \text{ tenure}$$

(0,11053)            (0,0086021)            (0,0033696)            (0,0025872)

Für alle Modell gilt  $n = 935$ ,  $R^2 = 0,155$  und  $\hat{\sigma} = 0,38773$ ; Standardfehler werden in Klammern angegeben.

- a) Geben Sie mit Hilfe einer geeigneten Tabelle aus dem Anhang dieser Klausur eine möglichst kleine obere Schranke für den kritischen Wert einer  $t$ -verteilten Zufallsvariablen mit 932 Freiheitsgraden zum Signifikanzniveau  $1 - \alpha/2$  bei  $\alpha = 1\%$  für einen zweiseitigen Test an!
- b) Benutzen Sie die Informationen einer geeigneten Modellgleichung um je ein zweiseitiges Konfidenzintervall zum Signifikanzniveau  $\alpha = 1\%$  für `educ` und für `exper` zu konstruieren! Runden Sie hierbei die Intervallgrenzen auf drei Nachkommastellen.

Betrachten Sie für die folgenden zwei Teilaufgaben folgende Nullhypothese: „Die Variablen `educ` und `exper` haben den gleichen Effekt auf `lwage`!“

- c) Argumentieren Sie, warum die Informationen aus Modellgleichung 1 nicht ausreichen, um die obige Nullhypothese zu testen!
- d) Wählen Sie eine geeignete Modellgleichung, um die obige Nullhypothese zu testen! Begründen Sie Ihre Wahl und Ihre Testentscheidung!

## Aufgabe 4

Betrachtet sei der Datensatz `fertil3`. Für `cgfr` (change general fertility rate, Veränderung der Gesamtfertilitätsrate; Kinder pro Frau) und `cpe` (change personal tax exemption, Veränderung der individuellen Steuerbefreiung) in den Jahren 1916-1984 wurde folgendes Modell per OLS geschätzt:

$$\text{Modell 1:} \quad cgfr_t = \gamma_0 + \delta_0 cpe_t + \delta_1 cpe_{t-1} + \delta_2 cpe_{t-2} + u_t$$

Der Output von Gretl lautet (Standardfehler in Klammern):

$$\widehat{cgfr} = -0,963679 - 0,0362021 cpe - 0,0139706 cpe\_1 + 0,109990 cpe\_2$$

(0,46776)                      (0,026774)                      (0,027554)                      (0,026880)

Die Residuen dieser Regression wurden als Variable `uhat1` gespeichert und es wurde folgendes Modell per OLS geschätzt:

$$\text{Modell 2:} \quad \hat{u}_t = \alpha_0 + \rho \hat{u}_{t-1} + \epsilon_t$$

Der Output von Gretl lautet (Standardfehler in Klammern):

$$\widehat{uhat1} = 0,0180209 + 0,291809 uhat1\_1$$

(0,44425)                      (0,11798)

Im nächsten Schritt werden folgende Variablen definiert:

- $\widetilde{cgfr}_t = cgfr_t - 0,292 \cdot cgfr_{t-1}$
- $\widetilde{cpe}_t = cpe_t - 0,292 \cdot cpe_{t-1}$
- $\tilde{u}_t = u_t - 0,292 \cdot u_{t-1}$

Nun wird Modell 1 mit den Daten  $\widetilde{cgfr}_t$  und  $\widetilde{cpe}_t$  per OLS geschätzt und die Residuen `uhat3` gespeichert. Mit diesen Residuen wird nun Modell 2 geschätzt und wir erhalten folgenden Output (Standardfehler in Klammern):

$$\widehat{uhat3} = 0,00685406 + 0,0145004 uhat3\_1$$

(0,44999)                      (0,12428)

- a) Erklären Sie, was ein AR(1) Prozess ist und für welchen Datentyp dieser zu erwarten ist!
- b) Testen Sie, ob die Störterme  $u_t$  von Modell 1 ein AR(1) Prozess sind! Benutzen Sie hierfür den Output für Modell 2 und beschreiben Sie Ihr Vorgehen!
- c) Was ist unter dem Vorgang der Quasidifferenzierung zu verstehen und welchem Zweck dient diese?
- d) Testen Sie, ob die transformierten Störterme  $\tilde{u}_t$  autokorreliert sind! Beschreiben Sie hierfür Ihr Vorgehen!

## Aufgabe 5

Es sei folgendes Modell gegeben:

$$y_i = \beta_0 + \beta_1 x_{1i} + u_i, \quad i = 1, \dots, n = 500$$

Sie vermuten, dass  $x_{1i}$  mit dem Störterm  $u_i$  korreliert ist. Sie verfügen außerdem über die Daten  $(z_1, \dots, z_n)$ , wobei Sie wissen, dass  $z_i$  nicht mit  $u_i$  korreliert ist. Die Stichprobenvarianz von  $z_i$  sei  $s_z^2 = 10$  und die Stichprobenkovarianz von  $z_i$  mit  $y_i$  sei  $s_{zy} = 100$ .

- Welches Problem tritt auf, wenn das Modell per OLS geschätzt wird?
- Unter welcher Bedingung ist  $z_i$  eine geeignete Instrumentalvariable?
- Eine OLS-Schätzung ergibt  $\hat{x}_{1i} = \hat{\gamma}_0 + \hat{\gamma}_1 z_i$ , wobei  $\hat{\gamma}_1 = -10$  und der Standardfehler von  $\hat{\gamma}_1$  den Wert  $se(\hat{\gamma}_1) = 2$  hat. Testen Sie, ob  $x_{1i}$  und  $z_i$  korreliert sind! Liegt hier das „weak instruments problem“ vor?
- Ermitteln Sie den IV-Schätzer für  $\beta_1$  unter Verwendung von  $z_i$  als Instrumentalvariable.





















# Kritische Werte der $t$ -Verteilung

		Signifikanzniveau				
		10%	5%	2,5%	1%	0,5%
einseitig:		10%	5%	2,5%	1%	0,5%
zweiseitig:		20%	10%	5%	2%	1%
Freiheits- grade	1	3,078	6,314	12,706	31,821	63,657
	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	
$\infty$	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
	$\infty$	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
	$\infty$	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

## Kritische Werte der $\chi^2$ -Verteilung

		Signifikanzniveau		
		10%	5%	1%
Freiheits- Grade	1	2,71	3,84	6,63
	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