Klausur zu Ökonometrie (Master)

Technische Universität Dortmund | Fakultät Wirtschaftswissenschaften | 1. Oktober 2024

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.

Erlaubte Hilfsmittel:

- Taschenrechner (nicht programmierbar)
- Ein DIN-A4 Blatt mir eigenen handschriftlichen Notizen (Vorder- und Rückseite)

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

Es seien die Beobachtungen $\mathbf{y} \in \mathbb{R}^n$ und $X \in \mathbb{R}^{n \times k+1}$ verfügbar, wobei $\mathbf{y}$ der Regressand und X die Regressormatrix sei.

In der Vorlesung wurden die Annahmen MLR 1, MLR 3, MLR 4 und MLR 5 zur Analyse der linearen Regression mit $k+1$ Regressoren (inklusive Konstante) getroffen.

a) Wie lauten diese vier Annahmen?

b) Geben Sie für jede dieser vier Annahmen ein Beispiel, welches die entsprechende Annahme verletzt.

c) Wie lautet der kleinste Quadrate Schätzer $\hat{\beta}$? Argumentieren Sie, welche der obigen Annahmen notwendig für die Spezifikation und Existenz dieses Schätzers sind.

d) Welchen Erwartungswert und welche Varianz-Kovarianz-Matrix hat der Schätzer $\hat{\beta}$? Argumentieren Sie, welche der obigen Annahmen benötigt werden, um den Erwartungswert und die Varianz-Kovarianzmatrix dieses Schätzers zu bestimmen.

Aufgabe 2

Dieser Aufgabe liegt der in der Vorlesung und Übung analysierte Datensatz `wage1` zugrunde. Ihnen steht folgender Output zur Verfügung:

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

	Koeffizient	Std. Fehler	t-Quotient	p-Wert
const	−3,39054	0,766566	−4,423	0,0000
exper	0,0700954	0,0109776	6,385	0,0000
educ	0,644272	0,0538061	11,97	0,0000

Mittel abhängige Var.	5,896103	Stdabw. abhängige Var.	3,693086
Summe quad. Residuen	5548,160	Stdfehler Regression	3,257044
R^2	0,225162	Korrigiertes R^2	0,222199
$F(2, 523)$	75,98998	P-Wert(F)	1,07e−29

a) Wie sind die Werte in der Spalte „t-Quotient" zu interpretieren? Welche Formel liegt ihnen zugrunde?

b) Welchen Schluss können Sie in Bezug auf den Wert des t-Quotienten des Regressors `educ` ziehen?

c) Wie viel Prozent der Streuung von `wage` wird durch die Streuung von `exper` und `educ` erklärt? Erklären Sie auf Grundlage welchen Werts der Tabelle Sie Ihre Aussage treffen und wie dieser Wert berechnet wird!

d) Welche statistische Kennzahl wird mit SST bezeichnet und wie lautet dieser Wert im vorliegenden Fall? Es gibt zwei Möglichkeiten, diesen auf Grundlage des vorliegenden Outputs zu berechnen. Geben Sie mindestens eine Möglichkeit an!

Aufgabe 3

Es sei Z und Y beliebige Zufallsvariablen mit $|\mathbb{E}[Z]|, |\mathbb{E}[Y]| < \infty$ und $\mathbb{E}[Z^2], \mathbb{E}[Y^2] < \infty$.

a) Definieren Sie die Varianz von Z, $Var(Z)$, und die Kovarianz von Z und Y, $Cov(Z, Y)$, als Erwartungswerte!

Betrachten Sie nun das lineare Regressionsmodell unter den Annahmen MLR 1 bis MLR 6:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$

Es bezeichnen $\hat{\beta}_1$ und $\hat{\beta}_2$ die OLS-Schätzer für β_1 und β_2.

Es bezeichnen $\sigma_1^2 = Var(\hat{\beta}_1)$, $\sigma_2^2 = Var(\hat{\beta}_2)$ und $\sigma_{12}^2 = Cov(\hat{\beta}_1, \hat{\beta}_2)$.

Um die Nullhypothese $\beta_1 + 2\beta_2 = 2$ zu testen wird der Schätzer $\hat{\theta} = \hat{\beta}_1 + 2\hat{\beta}_2$ definiert.

b) Bestimmen Sie $Var(\hat{\theta})$ in Ausdrücken von σ_1^2, σ_2^2 und σ_{12}^2. Wie lautet der Standardfehler $se(\hat{\theta})$ von $\hat{\theta}$?

c) Bestimmen Sie die t-Statistik für die Nullhypothese $\mathbb{E}[\hat{\theta}] = 2$!

d) Definieren Sie $\theta = \beta_1 + 2\beta_2$ und $z = x_2 - 2x_1$. Stellen Sie eine Regressionsgleichung mit den Parametern β_0, θ, β_2 und β_3 und den Regressoren x_1, z und x_3 auf, welche äquivalent zur obigen Regressionsgleichung ist und deren OLS-Schätzung Ihnen erlaubt $se(\hat{\theta})$ direkt abzulesen.

Aufgabe 4

a) Was bedeutet der Ausdruck „Heteroskedastizität“ in Bezug auf die unbeobachtbaren Störterme u_i, $i = 1, \ldots, n$? Welche der Annahmen MLR 1 bis MLR 6 wird oder werden durch Heteroskedastizität verletzt?

b) Warum sind heteroskedastische Störterme in Bezug auf das OLS-Schätzverfahren problematisch?

c) Nennen Sie ein Testverfahren, mit welchem Heteroskedastizität festgestellt werden kann und erläutern Sie kurz dessen Vorgehen!

Es sei angenommen, dass die auf X bedingte Varianz-Kovarianz-Matrix der Störterme bekannt und wie folgt gegeben sei:

$$\Sigma\left(\mathbf{u}|X\right) = \sigma^2 \begin{pmatrix} \frac{\sigma_1^2}{\sigma^2} & 0 & \ldots & 0 \\ 0 & \frac{\sigma_2^2}{\sigma^2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{\sigma_n^2}{\sigma^2} \end{pmatrix}$$

d) Wie müssten der Regressand $\mathbf{y}$ und die Regressoren X transformiert werden, so dass OLS anwendbar ist und Hypothesen an die Parameter valide getestet werden können? Zeigen Sie, dass die transformierten Störterme einen Erwartungswert von null haben und homoskedastisch sind!

Aufgabe 5

In einem Zeitreihenmodell sei der unbeobachtbare Störterm u_t, $t = 1, \ldots, n$ ein autoregressiver Prozess erster Ordnung, kurz $AR(1)$.

a) Welche Formel definiert den $AR(1)$-Prozess u_t für ein gegebenes u_0?

b) Unter welchen Bedingungen ist der $AR(1)$-Prozess $\{u_t\}_{t=1}^n$ stationär?

c) Welche der Annahmen MLR 1 bis MLR 6 ist oder sind verletzt? Warum ist dies in Bezug auf das OLS-Schätzverfahren problematisch?

d) Nehmen Sie an, die Korrelation zwischen den Störtermen u_t und u_{t-1} sei bekannt und durch den Wert ρ gegeben. Wie müssen der Regressand $\mathbf{y}$ und die Regressoren X transformiert werden, so dass OLS anwendbar ist und Hypothesen an die Parameter valide getestet werden können? Zeigen Sie, dass die transformierten Störterme einen Erwartungswert von null haben und seriell unkorreliert sind!

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

Lösung für Aufgabe 1

a) Die Annahmen lauten wie folgt:

MLR 1 $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{u}$
(Der Erwartungswert von $\mathbf{y}$ ist linear in den Parametern $\boldsymbol{\beta}$.)

MLR 3 $rk(X) = k + 1$
(Die Regressormatrix hat vollen Spaltenrang; kein Regressor kann durch eine Linearkombination der anderen Regressoren ausgedrückt werden.)

MLR 4 $E[\mathbf{u}|X] = \mathbf{0}$
(Strikte Exogenität)

MLR 5 $\Sigma(\mathbf{u}|X) = \sigma^2 \cdot I_n$
(homoskedastische und seriell unkorellierte Störterme)

b)

MLR 1 ist zum Beispiel verletzt, falls $y_i = \beta_0 + x_i + u_i$

MLR 3 ist zum Beispiel verletzt, falls $x_{i0} = 1$, $x_{i1} = d_i$ und $x_{i2} = 1 - d_i$ für $i = 1, \ldots, n$ lauten. (Dann würde gelten $x_{\bullet 1} + x_{\bullet 2} = \iota$ und $rk(x) \leq k <$ k+1).

MLR 4 ist zum Beispiel verletzt, falls u einen Regressor enthält, der nicht im Regressionsmodell enthalten ist, aber mit mindestens einem vorhandenen Regressor korreliert.

MLR 5 ist zum Beispiel verletzt, falls $Var(u_i) = \sigma^2 \cdot x_i$ für einen Regressor x oder falls $Cov(u_t, u_{t+1}) \neq 0$.

c) Der kleinste Quadrate Schätzer lautet $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{y}$. Diese Spezifikation ist nur sinnvoll, falls MLR 1: $\mathbf{y} = X\boldsymbol{\beta} + \mathbf{u}$ erfüllt ist. Der Schätzer existiert nur dann, falls $(X'X)^{-1}$ existiert. Dies ist der Fall, wenn X den vollen Spaltenrang hat, also wenn MLR 3: $rk(X) = k + 1$ gilt.

d) Der Erwartungswert von $\hat{\boldsymbol{\beta}}$ lautet $E[\hat{\boldsymbol{\beta}}|X] = (X'X)^{-1}X'E[\mathbf{y}|X]$ unter MLR 1 gilt $E[\mathbf{y}|X] = X\boldsymbol{\beta} + E[\mathbf{u}|X]$ und unter MLR 4 gilt dann $E[\hat{\boldsymbol{\beta}}|X] = (X'X)^{-1}X'X\boldsymbol{\beta} + (X'X)^{-1}X'E[\mathbf{u}|X] = \boldsymbol{\beta}\ \forall\ \boldsymbol{\beta} \in \mathbb{R}^{k+1}$.

Die Varianz-Kovarianz-Matrix von $\hat{\boldsymbol{\beta}}$ lautet $\Sigma(\hat{\boldsymbol{\beta}}|X) = (X'X)^{-1}X'\Sigma(\mathbf{y}|X)\left((X'X)^{-1}X'\right)'$. Mit MLR 1 gilt $\Sigma(\mathbf{y}|X) = \Sigma(\mathbf{u}|X)$ und mit MLR 5 gilt $\Sigma(\mathbf{u}) = \sigma^2 I$. Daraus folgt $\Sigma(\hat{\boldsymbol{\beta}}|X) = \sigma^2(X'X)^{-1}$.

Lösung für Aufgabe 2

a) Unter der Nullhypothese, dass der Parameter β_j für den j-ten Regressor gleich null ist, $H_0 : \beta_j = 0$, und falls die Modellannahmen MLR 1 bis MLR 6 erfüllt sind, gibt der t-Quotient den Wert der t-verteilten Zufallsvariablen $t_j = \frac{\hat{\beta}_j - 0}{se(\hat{\beta}_j)}$ an. Ist dieser Wert stark von null verschieden, insbesondere betragsmäßig größer als der kritische Wert $c_{t_{n-k-1},1-\frac{\alpha}{2}}$, so kann die Nullhypothese H_0 zum Signifikanzniveau α verworfen werden.

b) Falls die Modellannahmen MLR 1 bis MLR 6 erfüllt sind, kann die Hypothese $\beta_{educ} = 0$ zum Signifikanzniveau $\alpha = 1\%$ verworfen werden, da $c_{523;0,995} \approx 2,6 < 11,97 = t_{educ}$.

c) Auf Grundlage von $R^2 = 0,225162$ werden 22,5% der Streuung von `wage` durch die Streuungen von `exper` und educ erklärt. Es gilt $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$.

d) Es gilt $SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$, wobei y_i die i-te Beobachtung des Regressanden bezeichnet. SST ist also die Summe der quadrierten Abweichungen des Regressanden von seinem arithmetischen Mittel.

Möglichkeit 1: Dem Output ist zu entnehmen, dass die Summe der quadrierten Residuen durch $SSE = 5548,160$ gegeben ist. Nach Umformung der Formel für R^2 ergibt sich:

$$R^2 = 1 - \frac{SSR}{SST} \Leftrightarrow SST = \frac{SSR}{1 - R^2} = \frac{5548,160}{1 - 0,225162} = 7160,41$$

Möglichkeit 2: Dem Output ist zu entnehmen, dass die Standardabweichung der abhängigen Variable $se(y) = 3,693086$ beträgt. Die Formel für die Standardabweichung lautet:

$$se(y) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} = \sqrt{\frac{1}{n-1}SST}$$

Formt man diese nach SST um, ergibt sich:

$$(se(y))^2 = \frac{1}{n-1}SST \Leftrightarrow SST = (n-1)\,(se(y))^2$$

Wird $n = 526$ und $se(y) = 3,693086$ eingesetzt, erhält man:

$$SST = 525(3,693086)^2 = 7160,41$$

Lösung zu Aufgabe 3

a) $Var(Z) = \mathbb{E}[(Z - \mathbb{E}[Z])^2]$, $Cov(Z,Y) = \mathbb{E}[(Z - \mathbb{E}[Z])(Y - \mathbb{E}[Y])]$

b)

$$\begin{aligned}
Var(\hat{\theta}) &= Var(\hat{\beta}_1 + 2\hat{\beta}_2) \\
&= Var(\hat{\beta}_1) + 2Cov(\hat{\beta}_1, 2\hat{\beta}_2) + Var(2\hat{\beta}_2) \\
&= Var(\hat{\beta}_1) + 4Cov(\hat{\beta}_1, \hat{\beta}_2) + 4Var(\hat{\beta}_2) \\
&= \sigma_1^2 + 4\sigma_{12}^2 + 4\sigma_2^2
\end{aligned}$$

$$se(\hat{\theta}) = \sqrt{\sigma_1^2 + 4\sigma_{12}^2 + 4\sigma_2^2}$$

c)

$$t_\theta = \frac{\hat{\theta} - 2}{se(\hat{\theta})} = \frac{\hat{\beta}_1 + 2\hat{\beta}_2 - 2}{\sqrt{\sigma_1^2 + 4\sigma_{12}^2 + 4\sigma_2^2}}$$

d)

$$\begin{aligned}
y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u \\
&= \beta_0 + \beta_1 x_1 + 2\beta_2 x_1 - 2\beta_2 x1 + \beta_2 x_2 + \beta_3 x_3 + u \\
&= \beta_0 + (\underbrace{\beta_1 + 2\beta_2}_{\theta})x_1 + \beta_2(\underbrace{x_2 - 2x_1}_{z}) + \beta_3 x_3 + u
\end{aligned}$$

Lösung zu Aufgabe 4

a) Die unbeobachtbaren Störterme u_i sind heteroskedastisch, falls die Varianzen $Var(u_i)$ für mindestens zwei Beobachtungen unterschiedlich sind. Annahme MLR 5 fordert $Var(u_i) = \sigma^2$ für alle $i = 1, \ldots, n$ ist unter Heteroskedastizität verletzt. Annahme MLR 6 fordert $\mathbf{u} \sim \mathcal{N}(0, \sigma^2 I)$ und ist unter Heteroskedastizität ebenfalls verletzt.

b) Wird die Heteroskedastizität ignoriert, so werden die Standardfehler der Schätzer falsch geschätzt und die Teststatistiken falsch berechnet. Es kann also nicht valide getestet werden.

c) Der Breusch-Pagan-Test testet die Störterme auf Heteroskedastizität. Hierbei wird zunächst eine normale OLS-Regression durchgeführt und die quadrierten Residuen gespeichert. Unter der Nullhypothese der Homoskedastizität sollte die Streuung der Störterme nicht durch die Streuung der Regressoren erklärt werden. Der Breusch-Pagan-Test regressiert nun die quadrierten Residuen auf die Regressoren und testet global, ob die Regressoren einen Erklärungsgehalt für die quadrierten Residuen haben. Ist dies der Fall, wird die Nullhypothese der homoskedastischen Störterme abgelehnt.

d) Werden alle Daten der Beobachtung i durch $\sigma_i/\sigma = \sqrt{\sigma_i^2/\sigma^2}$ geteilt, so sind die transformierten Störterme homoskedastisch. Definiere $\tilde{y}_i = \sigma \cdot y_i/\sigma_i$, $\tilde{x}_{ij} = \sigma \cdot x_{ij}/\sigma_i$ und $\tilde{u}_i = \sigma \cdot u_i/\sigma_i$. Dann gilt:

$$\mathbb{E}[\tilde{u}_i] = \mathbb{E}[\sigma \cdot u_i/\sigma_i] = \sigma \cdot \mathbb{E}[u_i]/\sigma_i = \sigma \cdot 0/\sigma_i = 0 \checkmark$$

$$Var(\tilde{u}_i) = Var(\sigma \cdot u_i/\sigma_i) = \sigma^2 \cdot Var(u_i)/\sigma_i^2 = \sigma^2 \cdot \sigma_i^2/\sigma_i^2 = \sigma^2 \checkmark$$

Lösung zu Aufgabe 5

a) $u_t = \rho \cdot u_{t-1} + \epsilon_t$, wobei $\rho \neq 0$ und $\epsilon_t \overset{iid}{\sim} (0, \sigma_\epsilon^2)$.

b) $\{u_t\}_{t=1}^n$ ist stationär, falls $\mathbb{E}[u_t] = 0$ für alle t und $Var(u_t) = \sigma^2$ für alle t. Dies ist der Fall, wenn $|\rho| < 1$.

c) Annahme MLR 5' fordert $Cov(u_t, u_s) = 0$ für alle $t \neq s$. Für einen $AR(1)$-Prozess gilt aber $Cov(u_t, u_{t-s}) = \rho^s \sigma^2$, also ist MLR 5' verletzt. MLR 6 ist entsprechend ebenfalls verletzt.

d) Die Transformationsmatrix lautet

$$P = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -\rho & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -\rho & 1 \end{pmatrix} ,$$

die tranformierten Daten lauten:

$$\tilde{y} = P\mathbf{y}, \; \tilde{X} = PX, \; \tilde{u} = P\mathbf{u} \; .$$

Konkret gilt für die transformierten Störterme:

$$\tilde{u}_t = u_t - \rho u_{t-1} = \rho u_{t-1} + \epsilon_t - \rho u_{t-1} = \epsilon_t$$

Per Annahme an ϵ gilt dann $\mathbb{E}[\tilde{u}_t] = \mathbb{E}[\epsilon_t] = 0$, $Var(\tilde{u}_t) = Var(\epsilon_t) = \sigma_\epsilon^2$ und $Cov(u_t, u_s) = Cov(\epsilon_t, \epsilon_s) = 0$ für alle $t \neq s$.