

Eksamen på Økonomistudiet sommer 2019

**Økonometri I**

Tag-hjem eksamen

13. juni 2019

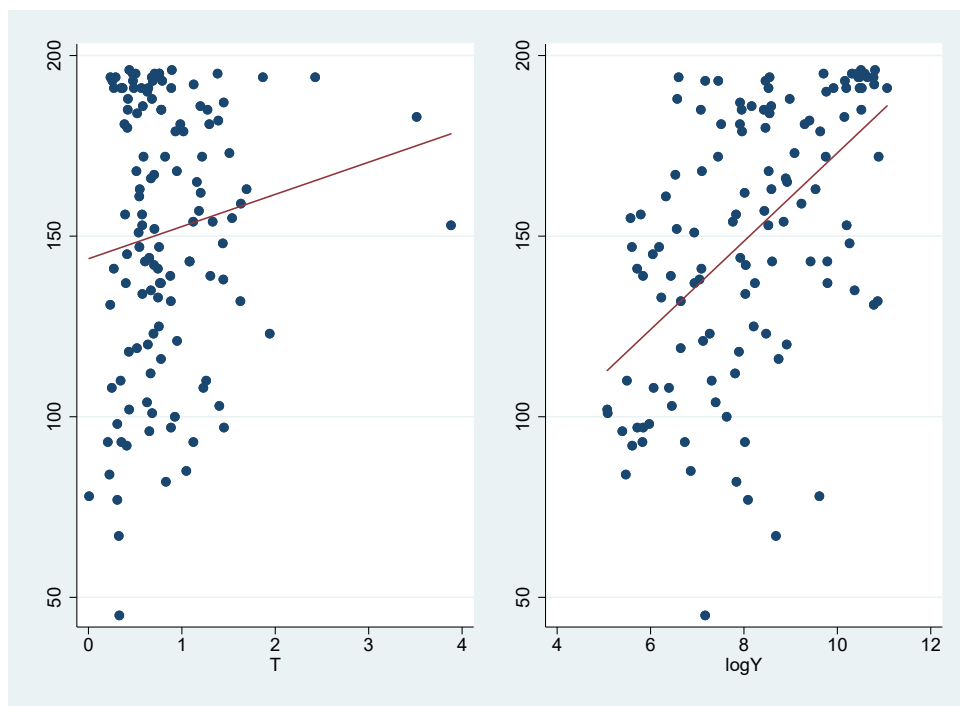
**VEJLEDENDE BESVARELSE**

## Problem 1

Table 1: Summary Statistics

	mean	p50	min	max
Y	11273.6	3512.6	159.6	63918.1
logY	8.215	8.164	5.073	11.07
T	0.839	0.703	0.00665	3.881
P	47.55	10.48	0.297	1304.5
logP	2.510	2.349	-1.215	7.174
A	934342.3	241038	316	17075400
logA	12.18	12.39	5.756	16.65
pTi	0.168	0.151	0.0198	0.700
pTa	0.177	0.163	0.0240	0.717
N	151.2	155	45	196
$N$	125			

Tabel 1 viser at det gns. BNP per capita i data er \$11274. Spredning på BNP per capita er markant, med den laveste indkomst i Etiopien (\$160) og den højeste indkomst i Norge (\$63918). Handelsindekset er i gns. 0.84, og det bemærkes at Hong Kong og Singapore har index større end 3.5, hvorimod Somalia har et indeks tæt på nul. Disse eksempler illustrerer, at indekset ikke er begrænset til intervallet (0,1). Således vil små økonomier, som også er vigtige handelstationer, have høje værdier for  $T$ . Et land har i gns. 151 handelspartnere, hvilket er tæt medianen på 155. Bhutan og Botswana har færre end 75 handelspartnere, mens Tyskland, Italien og England har over 195 handelspartnere. Der observeres også markant variation i befolkningsstørrelse og landeareal på tværs af observationerne i data.



Den første figur viser en positiv, men insignifikant, sammenhæng ml. handelsindekset og antal handelspartnere. Den anden figur viser en positiv og signifikant sammenhæng ml. log BNP per capita og antal handelspartnere. Rige lande er således mere globale.

## Problem 2

Table 2: OLS-resultater

	(1)	(2)	(3)	(4)
	OLS-HCSE	OLS	BP-test	RESET
T	0.220 (0.262)	0.220 (0.308)	-1.047* (0.506)	-54.63 (44.55)
logP	-0.0898 (0.137)	-0.0898 (0.133)	0.0192 (0.218)	22.14 (18.19)
logA	-0.101 (0.114)	-0.101 (0.114)	-0.00831 (0.188)	25.07 (20.50)
yhat2				28.48 (24.14)
yhat3				-1.083 (0.957)
_cons	9.489* (1.315)	9.489* (1.356)	3.528 (2.230)	-1629.0 (1356.6)
N	125	125	125	125
R <sup>2</sup>	0.052	0.052	0.046	0.086

Standard errors in parentheses

\*  $p < 0.05$

1. *Svar:* OLS-estimerne for model 1 er angivet i kolonne 1-2 i Tabel 2, hvor de robuste standardfejl er rapporteret i den første kolonne, mens de konventionelle standardfejl er angivet i den anden kolonne. Resultaterne viser, at hvis handelsindekset øges med en enhed, hvilket svarer til at importen og eksporter øges med et beløb svarende til landets BNP, forventes BNP per capita at stige med  $100 \cdot \hat{\beta}_1 = 22$  pct., alt andet lige. Resultaterne viser også, at mere befolkede og fysisk større lande har lavere BNP per capita, alt andet lige.
2. *Svar:* Breusch-Pagan testet undersøger om fejlleddet  $u$  er heteroskedasticitet på baggrund af følgende hjælperegression:

$$\hat{u}_i^2 = \delta_0 + \delta_1 T_i + \delta_2 \log P_i + \delta_3 \log A_i + \eta_i$$

hvor  $\hat{u}_i^2$  er de kvadrerede OLS-residualer beregnet ud fra kolonne 1-2 i Tabel 2. Nulhypotesen opstilles som:

$$H_0 : \delta_1 = \delta_2 = \delta_3 = 0$$

Hjælperegressionen estimeres af OLS og man finder at  $R^2 = 0.0463$ , jf. kolonne 3 i Tabel 2. Ergo kan næsten 5% af variationen i de kvadrerede OLS-residualer forklares af regressorerne i hjælperegressionen. Teststørrelsen beregnes som  $LM = 125 \times R^2 = 5.782$ . Under  $H_0$  er teststørrelsen  $\chi^2$ -fordelt med 3 frihedsgrader. Vælges

et signifikansniveau på 5%, er den kritiske værdi  $c = 7.81$ . Da  $LM < c$  kan nulhypotesen ikke afvises. Konklusionen er derfor at heteroskedasticitet ikke er et problem. Denne testprocedure er gyldig såfremt MLR.1-4 er opfyldt og  $n \rightarrow \infty$ .

3. *Svar:* RESET undersøger om den funktionelle form af model (1) er korrekt specificeret. Den relevante model er nu:

$$\log Y = \beta_0 + \beta_1 T + \beta_2 \log P + \beta_3 \log A + \theta_1 (\widehat{\log Y})^2 + \theta_2 (\widehat{\log Y})^3 + u$$

hvor  $\widehat{\log Y}$  er den prædikterede log BNP per capita beregnet ud fra Tabel 2, kolonne 2. Nulhypotesen opstilles som:

$$H_0 : \theta_1 = \theta_2 = 0$$

Ifølge kolonne 4 i Tabel 2 er forklaringsgraden,  $R^2 = 0.086$ , i den udvidede model, mens forklaringsgraden for den oprindelige model er  $R^2 = 0.052$ . Den relevante teststørrelse er et ordinært  $F$ -test, da antagelsen om homoskedasticitet er opfyldt. Teststørrelsen beregnes som  $F = \frac{(R_{ur}^2 - R_r^2)/q}{(1 - R_{ur}^2)/(n - k - 1)} = \frac{(0.086 - 0.052)/2}{(1 - 0.086)/(125 - 5 - 1)} = 2.19$  (når alle betydende cifre inkluderes, se STATA do-file). Under  $H_0$  er teststørrelsen  $F$ -fordelt med (2,119) frihedsgrader. Vælges et signifikansniveau på 5%, er den kritiske værdi  $c = 2.35$  for en  $F_{2,120}$ -fordeling. Da  $F < c$  kan nulhypotesen ikke afvises. Det bekræftes af  $p$ -værdien som beregnes til 0.1159. Konklusionen er derfor at  $(\widehat{\log Y})^2$  og  $(\widehat{\log Y})^3$  samlet set ikke indgår signifikant i den udvidede model. Derfor er (1) korrekt specificeret i forhold til den funktionelle form. Denne testprocedure er gyldig såfremt MLR.1-5 er opfyldt og  $n \rightarrow \infty$ .

4. *Svar:* Hypotesen opstilles som

$$H_0 : \beta_1 = 0 \quad \text{versus} \quad H_A : \beta_1 > 0$$

Teststørrelsen beregnes som  $\frac{0.220}{0.308} = 0.71$ . Under  $H_0$  er teststørrelsen  $T$ -fordelt med  $125 - 3 - 1 = 121$  frihedsgrader. Vælges et signifikansniveau på 5%, er den kritiske værdi  $c = 1.658$ . Nulhypotesen kan derfor ikke afvises og konklusionen er derfor at BNP per capita ikke afhænger af handelsindekset.  $P$ -værdien beregnes som  $p = P(T > 0.71) = 0.238$ . Konklusionen er derfor den samme uanset om den kritiske værdi eller  $p$ -værdien anvendes. Denne testprocedure er gyldig såfremt MLR.1-5 er opfyldt og  $n \rightarrow \infty$ .

### Problem 3

Table 3: IV-resultater

	(1)	(2)	(3)	(4)	(5)	(6)
	1st	1st	IV1	IV2	Exo,pTi	Exo,pTa
pTi	1.089* (0.429)					
logP	0.0368 (0.0386)	0.0426 (0.0387)	-0.279 (0.215)	-0.300 (0.240)	-0.279 (0.149)	-0.300 (0.158)
logA	-0.135* (0.0327)	-0.142* (0.0328)	0.483 (0.364)	0.548 (0.452)	0.483 (0.252)	0.548 (0.298)
pTa		0.887* (0.424)				
T			3.587 (1.929)	3.961 (2.454)	3.587* (1.338)	3.961* (1.617)
e1					-3.546* (1.373)	
e2						-3.877* (1.646)
_cons	2.207* (0.383)	2.305* (0.387)	0.0247 (5.620)	-1.028 (7.072)	0.0247 (3.896)	-1.028 (4.660)
$N$	125	125	125	125	125	125
$R^2$	0.310	0.298	.	.	0.102	0.094

Standard errors in parentheses

\*  $p < 0.05$

1. *Svar*: Relevansen af de to instrumenter undersøges i to 1st stage regressioner:

$$T_i = \delta_0^1 + \delta_1^1 pTi_i + \delta_2^1 P_i + \delta_3^1 A_i + \varepsilon_i^1$$

$$T_i = \delta_0^2 + \delta_1^2 pTa_i + \delta_2^2 P_i + \delta_3^2 A_i + \varepsilon_i^2$$

Resultaterne i Tabel 3, kolonne 1 og 2, viser at begge instrumenter påvirker handelsindekset positivt. I begge tilfælde er  $\delta_1^j$  signifikant forskellig fra 0, da  $t_{\delta_1^1} = 2.54$  og  $t_{\delta_1^2} = 2.09$ . Instrumenterne vurderes derfor til at være relevante i begge tilfælde, selvom IV-variablene er på grænsen til at være svage instrumenter. Vi kan ikke teste om instrumenterne er ukorreleret med fejledet i model (1). IV-variablene forsøger at skabe eksogen variation i handelsindekset ved at anvende geografiske karakteristika for landene. Det synes plausibelt at geografiske karakteristika i sig selv påvirker landenes indkomstniveau via  $u$ . Derfor er eksogenitetsantagelsen problematisk. Det gælder især for  $pTi$ -variablen, som både afhænger af  $N$  og de geografiske forhold. I opgave 1 påvises en positiv sammenhæng ml.  $N$  og log indkomst. Det kan tænkes at rige lande har bedre rets- og finansinstitutioner, som gør det lettere at handle

med andre lande. Disse institutioner forventes i sig selv at påvirke BNP per capita, hvorfor  $N$  ikke er eksogen i forhold til  $u$ . Derfor vurderes  $pTa$ -variablen i højere grad at være eksogen, da  $N$  er konstant for alle lande. IV-estimerne er præsenteret i Tabel 3, kolonne 3 og 4. Det bemærkes, at IV-estimerne er mere end 10 gange så høje som OLS-estimerne. Resultaterne viser, at hvis handelsindekset øges med en enhed, forventes BNP per capita at stige med 359-396 pct., alt andet lige. Undersøges den samme hypotese som i opgave 2.4, kan man beregne en teststørrelse på  $t = 1.61$  ved brug af det foretrukne instrument,  $pTa$ . Dermed kan nulhypotesen om ingen effekt fortsat ikke afvises, selvom vi er tæt på drage den modsatte konklusion.

2. *Svar:* Hvorvidt  $T$  er en endogen variabel eller ej, undersøges ved hjælperegressionen:

$$\log Y = \beta_0 + \beta_1 T + \beta_2 \log P + \beta_3 \log A + \rho^j \hat{e}^j + \varepsilon \quad j = 1, 2$$

Ideen bag testet er, at den endogene variabel kan dekomponeres i to dele:  $T = \hat{T} + \hat{e}$ , hvor  $\hat{T}$  og  $\hat{e}$  er hhv. den prædikterede variabel og residualerne fra 1st stage regressionen. Hvis  $\hat{T}$  beregnes ud fra eksogene variable, må der gælde, at  $\text{cov}(u, \hat{e}) \neq 0$ , for at  $T$  er en endogen variabel. Under antagelse at  $u = \rho^j \hat{e}^j + \varepsilon$  kan hjælperegressionen bruges til at undersøge om  $T$  er endogen eller ej. Med to instrumenter kan der implementeres to endogenitetstest, hvor 2 refererer til  $pTa$ -instrumentet. Resultaterne er præsenteret i Tabel 3, kolonne 5 og 6. I begge tilfældes afvises  $H_0 : \rho^j = 0$  til fordel for alternativhypotesen  $H_A : \rho^j \neq 0$ . Konklusionen er derfor at  $T$  er en endogen variabel. IV-resultaterne baseret på  $pTa$ -instrumentet foretrækkes derfor. Dog kan OLS-resultaterne foretrækkes i stedet, hvis der sås tvivl om eksogeniteten af instrumentvariablen. Uanset hvilken metode der foretrækkes, er konklusionen, at handelsindekset er insignifikant i forhold til BNP per capita.

## Problem 4

1. *Svar:* OLS-estimatorens asymptotiske bias for model 1:

$$\begin{aligned} \text{plim } \hat{\beta}_1^{M1} - \beta_1 &= \beta_2 \frac{\text{cov}(x_1, x_2)}{\text{var}(x_1)} + \beta_3 \frac{\text{cov}(x_1, x_3)}{\text{var}(x_1)} + \frac{\text{cov}(x_1, u)}{\text{var}(x_1)} \\ &= \frac{\beta_2 \sigma_{12} + \beta_3 \sigma_{13}}{\sigma_1^2} \end{aligned}$$

idet  $E(u|x_1, x_2, x_3) = 0$ . Bemærk at bias ikke afhænger af  $\sigma_{23}$ ,  $\sigma_2^2$  eller  $\sigma_3^2$ . Biasudtrykket kan være negativt, nul eller positivt, og afhænger af  $x_1$ 's kovarians med de to udeladte variable ( $x_2$  og  $x_3$ ), samt sidstnævntes partielle effekt på  $y$ . Variansen af  $x_1$  indgår også i biasudtrykket, men påvirker kun størrelsen af biasen, og ikke retningen.

2. *Svar:* OLS-estimatoren for model 2 kan opstilles:

$$\hat{\beta}_1^{M2} = \frac{\sum_i \hat{r}_i y_i}{\sum_i \hat{r}_i^2}$$

hvor  $\hat{r}_i = x_{1i} - \hat{a} - \hat{b}x_{2i}$  er OLS-residualet fra en regression af  $x_1$  på  $x_2$ . Det følger af Frisch-Waugh teoremet og er vist i fagets første obligatoriske opgave.

OLS-estimatorens grænsesandsynlighed er derfor:

$$\begin{aligned}
\text{plim } \hat{\beta}_1^{M2} &= \frac{\text{cov}(\hat{r}, y)}{\text{var}(\hat{r})} \\
&= \frac{\text{cov}(\hat{r}, \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u)}{\text{var}(\hat{r})} \quad [\text{MLR.1}] \\
&= \frac{\beta_1 \text{cov}(\hat{r}, x_1) + \beta_3 \text{cov}(\hat{r}, x_3)}{\text{var}(\hat{r})} \quad [\text{cov}(\hat{r}, x_2) = 0 \text{ og MLR.4}] \\
&= \beta_1 + \beta_3 \frac{\text{cov}(\hat{r}, x_3)}{\text{var}(\hat{r})} \quad [\text{cov}(\hat{r}, x_1) = \text{cov}(\hat{r}, \hat{r} + a + bx_2) = \text{var}(\hat{r})] \\
&= \beta_1 + \beta_3 \delta_1
\end{aligned}$$

hvor  $\delta_1$  er fra hjælperegressionen,  $x_3 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \epsilon$ , som beskriver den indbyrdes afhængighed blandt  $x$ -variablene. Med udgangspunkt i kovarians-matricen, kan vi karakterisere  $\hat{\delta}_1$ :

$$\begin{aligned}
\text{plim } \hat{\delta}_1 &= \frac{\text{cov}(\hat{r}, x_3)}{\text{var}(\hat{r})} = \frac{\text{cov}(x_1 - \frac{\sigma_{12}}{\sigma_2^2} x_2, x_3)}{\text{var}(x_1 - \frac{\sigma_{12}}{\sigma_2^2} x_2)} \quad [\text{plim } \hat{b} = \frac{\sigma_{12}}{\sigma_2^2}] \\
&= \frac{\sigma_{13} - \frac{\sigma_{12}}{\sigma_2^2} \sigma_{23}}{\sigma_1^2 + \frac{\sigma_{12}^2}{(\sigma_2^2)^2} \sigma_2^2 - 2 \frac{\sigma_{12}}{\sigma_2^2} \sigma_{12}} = \frac{\sigma_{13} - \frac{\sigma_{12}}{\sigma_2^2} \sigma_{23}}{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}} \\
&= \frac{\sigma_{13} \sigma_2^2 - \sigma_{12} \sigma_{23}}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} = \delta_1
\end{aligned}$$

OLS-estimatorens asymptotiske bias for model 2 er derfor:

$$\text{plim } \hat{\beta}_1^{M2} - \beta_1 = \beta_3 \left( \frac{\sigma_{13} \sigma_2^2 - \sigma_{12} \sigma_{23}}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} \right)$$

Biasen afhænger ikke af  $\beta_2$  eller  $\sigma_3^2$ , men er ellers en funktion af alle andre kovarians-parametre. Udtrykket viser at udeladt variabel-biasen for en multipel regressions-model afhænger af den indbyrdes sammenhæng mellem *alle* variable ( $\sigma_{12}, \sigma_{13}, \sigma_{23}$ ), samt variansen af de inkluderede regressorer ( $\sigma_1^2, \sigma_2^2$ ).

- 3.i *Svar*: Det er vanskeligt at sammenligne bias-udtrykkene direkte. Det kan dog konkluderes at udeladt variabel-biasen for model 1 skyldes to udeladte variable, og er blot en simpel udvidelse af biasudtrykket fra pensum vedrørende en udeladt variabel i en model med en inkluderet regressor. Udeladt variabel-biasen for model 2 afhænger af  $\beta_3$  og den partielle korrelation mellem den udeladte variabel,  $x_3$ , og den forklarende variabel,  $x_1$ . Sidstnævnte er vanskelig at fortegnbestemme, da den afhænger af den indbyrdes samvariation blandt alle  $x$ -variable. Konklusionen er derfor, at OLS-estimatoren kan være mere biased for model 1 end model 2 for visse parameterværdier, mens det modsatte kan være tilfælde for andre parameterværdier. Dette kan illustreres ved et par eksempler. Hvis  $\beta_3 = 0$ , foretrækkes model 2-estimatoren, da den er konsistent, mens model 1-estimatoren lider af udeladt variabel-bias. Det samme er tilfældet, hvis variansen af  $x_2$  er således at  $\sigma_2^2 \approx (\sigma_{12} \sigma_{23}) / \sigma_{13}$ , da det medfører at  $\delta_1 \approx 0$ . I denne situation vil model 2-estimatoren være konsistent, hvorimod model-1 estimatoren fortsat lider af udeladt

variabel-bias. Hvis  $\sigma_{12} = 0$ , vil begge estimators indeholde samme bias, og dermed være lige problematiske. Hvis model 2-estimatoren lider af udeladt variabel-bias, er det muligt at vælge  $\beta_2$  således at model 1-estimatoren er konsistent ( $\beta_2$  indgår ikke i biasudtrykket for model 2-estimatoren). Altså kan inddragelsen af  $x_2$  både forværre, forbedre eller have ingen effekt på en bias af OLS-estimatoren som skyldes udeladte variable.

3.ii *Svar:* Hvis  $\sigma_{12} = 0$  gælder der:

$$\text{plim } \hat{\beta}_1^{M1} = \beta_1 + \beta_3 \frac{\sigma_{13}}{\sigma_1^2} = \text{plim } \hat{\beta}_1^{M2}$$

I dette tilfælde er udeladt variabel-biasen for OLS-estimatoren den samme uanset om  $x_2$  er inkluderet eller ej — og biasen kan fortegnbestemmes ud fra  $\beta_3$  og  $\sigma_{13}$ . Altså vil biasen i dette tilfælde kun afhænge af kovariansen mellem  $x_1$  og den udeladte variabel,  $x_3$ , samt den partielle effekt af sidstnævnte på  $y$ . Dette resultat illustrerer, at det simple udtryk for udeladt variabel-bias kun gælder for en multipel regressionsmodel hvor de inkluderede regressorer er ukorreleret. Hvis  $\sigma_{12} \neq 0$  kan det simple biasudtryk lede til misvisende konklusioner, da  $\delta_1 \neq \frac{\sigma_{13}}{\sigma_1^2}$ . Hvis  $\beta_2 = 0$  gælder der:

$$\text{plim } \hat{\beta}_1^{M1} = \beta_1 + \beta_3 \frac{\sigma_{13}}{\sigma_1^2} \neq \text{plim } \hat{\beta}_1^{M2}$$

Model 1-estimatoren lider i dette tilfælde kun af udeladt variabel-bias pga.  $x_3$ . Derfor gælder det simple biasudtryk for OLS-estimatoren, når  $\beta_2 = 0$ . Model 2 medtager her en irrelevant variabel, som samtidig er korreleret med  $x_1$ . Udredningen af udeladt variabel-bias kræver derfor at man tager højde for den indbyrdes afhængighed blandt  $x$ -variablene, hvorfor det simple biasudtryk ikke kan anvendes.

## Problem 5

Tabel 4 præsenterer resultaterne fra simulationseksperimentet. I det første scenarie er de gns. OLS-estimer for model 1 lig den sande parameter,  $\beta_1 = 0.5$ . Model 1-estimatoren er dog biased pga. to udeladte variable ( $x_2$  og  $x_3$ ), men i det første scenarie, neutraliserer de to biaskilder hinanden, så den samlede bias er lig nul. Tilsvarende er de gns. OLS-estimer for model 2 lig 0.5. Model 2-estimatoren er ikke biased, selvom  $x_3$  fortsat er en udeladt variabel. Dette skyldes at  $x_1$  ikke samvarierer med  $x_3$ , når der også tages højde for  $x_2$ . Sagt med andre ord: Den partielle korrelation mellem  $x_3$  og  $x_1$  er nul, selvom kovariansen mellem de to er forskellig fra nul. Tabel 4 rapporterer resultater for,  $\hat{\delta}_{x_1}$ , som stammer fra en regression af  $x_3$  på  $x_1$  og  $x_2$ . Resultaterne er således i fuld overensstemmelse med de teoretiske resultater i forrige opgave.

I scenarie 2 er  $\text{cov}(x_1, x_3) = 0$ . De gns. model 1-estimer er 0.187. Derfor er OLS negativt biased for model 1. I dette tilfælde er  $x_2$  den eneste kilde til bias. Da der gælder at  $\beta_2 < 0$  og  $\text{cov}(x_1, x_2) > 0$ , må biasen være negativ, hvilket understøttes af simulationsresultaterne. Model 2-estimerne er i gns.  $-0.151$ . Altså er OLS også negativt biased for model 2, og i endnu højere grad end model 1-estimatoren. Denne negative bias skyldes at den partielle korrelation mellem  $x_3$  og  $x_1$  er negativ og  $\beta_3 > 0$ . Dette resultat kan virke kontraintuitivt, men illustrerer omvendt forskellen på kovariansen ml. to variable (som her



er nul) versus deres partielle korrelation (som her er negativ når der tages højde for  $x_2$ ).

I scenarie 3 er  $cov(x_2, x_3) = 0$ . I det tredje scenarie er de gns. OLS-estimer for model 1 lig den sande parameter,  $\beta_1 = 0.5$ . Som i scenarie 1 indeholder model 1-estimatoren to biaskilder som dog neutraliserer hinanden. Resultater viser derfor at kovariansen mellem  $x_2$  og  $x_3$  ikke påvirker model 1-estimatorens egenskaber i forhold til bias. Dette resultat er at forvente på baggrund af de teoretiske konklusioner i opgave 4. Model 2-estimerne er i gns. lig 1.15. OLS-estimatorer er dermed positivt biased, når  $x_2$  inkluderes. Denne positive bias skyldes at den partielle korrelation mellem  $x_3$  og  $x_1$  er positiv og  $\beta_3 > 0$ .

I det sidste tilfælde er  $cov(x_1, x_2) = 0$ . De gns. model 1-estimer er 0.814, hvilket også er gennemsnittet for model 2-estimerne. I scenarie 4 er OLS-estimatoren dermed positivt biased og størrelsen på biasen er nøjagtigt den samme for begge modeller. For model 1 er  $x_3$  den eneste kilde til bias, og afhænger kun af  $\beta_3, \sigma_{13}, \sigma_1^2$ . Da biasen er nøjagtigt den samme for model 2, må den også kun afhænge af førnævnte parametre. Dette resultat viste vi formelt i opgave 4, dvs.  $\hat{\delta}_1 = \frac{\sigma_{13}}{\sigma_1^2}$ , hvis og kun hvis  $\sigma_{12} = 0$ . Resultaterne illustrerer således at den simple udeladt-variabel biasformel kun gælder for en multipel regression, hvor de inkluderede regressorer er indbyrdes urelateret.

OLS-estimatoren for model 1 foretrækkes frem for model 2 i scenarie 2 og 3, hvorimod biasen er den samme i scenarie 1 og 4. Ergo foretrækkes en simpel model af  $y$  på  $x_1$  frem for en model som også inddrager  $x_2$ . Dette er dog ikke en generel konklusion, men afspejler i høj grad specifikationen af den datagenererende proces. Dog illustrerer eksperimentet at man bør udvise forsigtighed i forhold til at fortegnbestemme en udeladt variabel-bias for en multipel lineær regressionsmodel.

Table 4: Simulationsresultaterne

		Opg.4	Mean	Std.dev.	Min	Max
1. Baseline $\Sigma$ :	Model 1- $\hat{\beta}_{x_1}$	0.5	0.501	0.057	0.337	0.706
	Model 2- $\hat{\beta}_{x_1}$	0.5	0.501	0.084	0.217	0.785
	$\hat{\delta}_{x_1}$	0	-0.000	0.048	-0.136	0.167
2. $\sigma_{13} = 0$ :	Model 1- $\hat{\beta}_{x_1}$	0.1875	0.187	0.053	0.019	0.364
	Model 2- $\hat{\beta}_{x_1}$	-0.1522	-0.151	0.070	-0.383	0.064
	$\hat{\delta}_{x_1}$	-0.522	-0.522	0.030	-0.607	-0.418
3. $\sigma_{23} = 0$ :	Model 1- $\hat{\beta}_{x_1}$	0.5	0.504	0.071	0.262	0.734
	Model 2- $\hat{\beta}_{x_1}$	1.1522	1.153	0.083	0.872	1.432
	$\hat{\delta}_{x_1}$	0.522	0.522	0.046	0.389	0.685
4. $\sigma_{12} = 0$ :	Model 1- $\hat{\beta}_{x_1}$	0.8125	0.814	0.053	0.662	1.001
	Model 2- $\hat{\beta}_{x_1}$	0.8125	0.814	0.053	0.666	1.000
	$\hat{\delta}_{x_1}$	0.25	0.251	0.029	0.160	0.339