Eksamen på økonomi-studiet vinter 2021-2022

Økonometri I

Vejledende besvarelse

Opgave 1

(a) Datasættet er et ubalanceret panel datasæt med 197 banker for 1992-2015. Der er i alt 3633 observationer. Datasættet indeholder variable, der varrierer over bank og tid: logaritmen til bankudlån, den pct.vise årlige vækst i bankudlån samt en variabel med forholdet mellem udlån og indlån i banken. Desuden indeholder datasættet makro variable, der kun varierer over tid. Det drejer sig om logaritmen til BNP og CIBOR renten samt begge variables førstedifferenser. Over perioden har været en gennemsnitlig vækst i udlånene på 8,2 pct. og en gennemsnitlig vækst i BNP på 1,8 pct. CIBOR renten har i perioden svinget mellem -0,7 pct. og 12,8 pct. med et gennemsnit på 4,2 pct.

	mean	sd	min	max
År	2001.8858	6.421	1992.000	2015.000
$CIBOR_t$	0.0422	0.032	-0.007	0.128
$\Delta CIBOR_t$	-0.0044	0.018	-0.059	0.030
$\log(BNP_t)$	14.3142	0.112	14.102	14.473
$\Delta\log(BNP_t)$	0.0179	0.019	-0.050	0.052
$\left(\frac{\text{udlan}}{\text{indlan}}\right)_{it}$	1.1029	1.602	-0.549	18.055
$\Delta\log(\text{udlån}_{it})$	0.0821	0.411	-8.177	1.829
Bank id nr	99.9166	56.889	1.000	197.000
No. obs	3633			

Table 1: Opgave 1.1

(b) I figur 1 er tidsserierne for $\log(BNP_t)$, $\Delta\log(BNP_t)$ vist. Det er oplagt, at $\log(BNP_t)$ trender og dermed er ikke-stationær. Væksten i BNP ser stationær ud bortset fra den store outlier i 2009 i forbindelse med finanskrisen. I figur 2 er $CIBOR_t$ og $\Delta CIBOR_t$ vist. Førstnævnte er helt åbenbart ikke stationær, da den er faldende over tid. Middelværdien for ændringerne i CIBOR renten ser konstant ud over perioden, men der lader til at være en større varians i 1992-1994. Dvs. både med $\Delta\log(BNP_t)$ og $\Delta CIBOR_t$ er vi ikke helt overbeviste om, at tidsserierne er stationære.

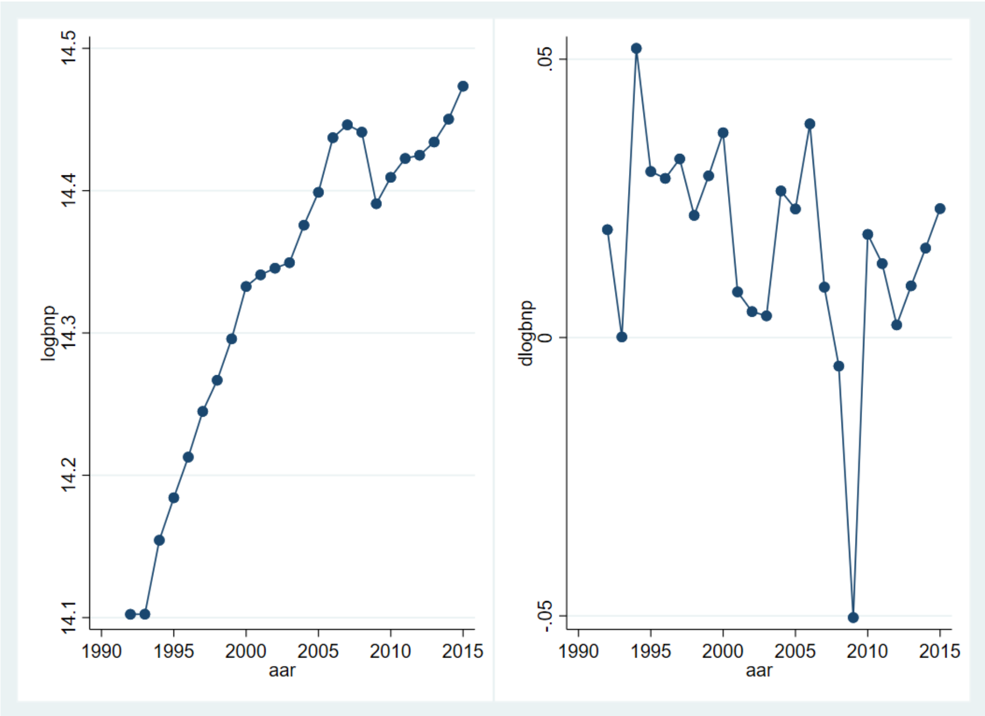

Figure 1: Opg. 1.2: Venstre: log(BNP). Højre: $\Delta\log$(BNP)

Opgave 2

(a) Vi estimerer den følgende regressionsmodel med pooled OLS

$$\begin{aligned}\Delta\log(\text{udlån}_{it}) = \beta_0 + \beta_1\Delta\log(\text{BNP}_{t-1}) + \beta_2\Delta\text{CIBOR}_{t-1} \\ +\beta_3\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1} + u_{it}\end{aligned} \tag{1}$$

Vi kan fortolke β_1 som pct.point ændring i bankernes udlån, når BNP stiger med 1 pct. Vi finder, jf. første kolonne i tabel 2, at når BNP stiger med 1 pct.point, så stiger udlånet med 0,8 pct.point. På lignende vis finder vi, at en stigning i CIBOR renten med på 1 pct.point får bankernes udlån til at falde med 2,5 pct. Tilslut har vi, at når udlån/ indlån raten stiger med 1, så falder udlånet med

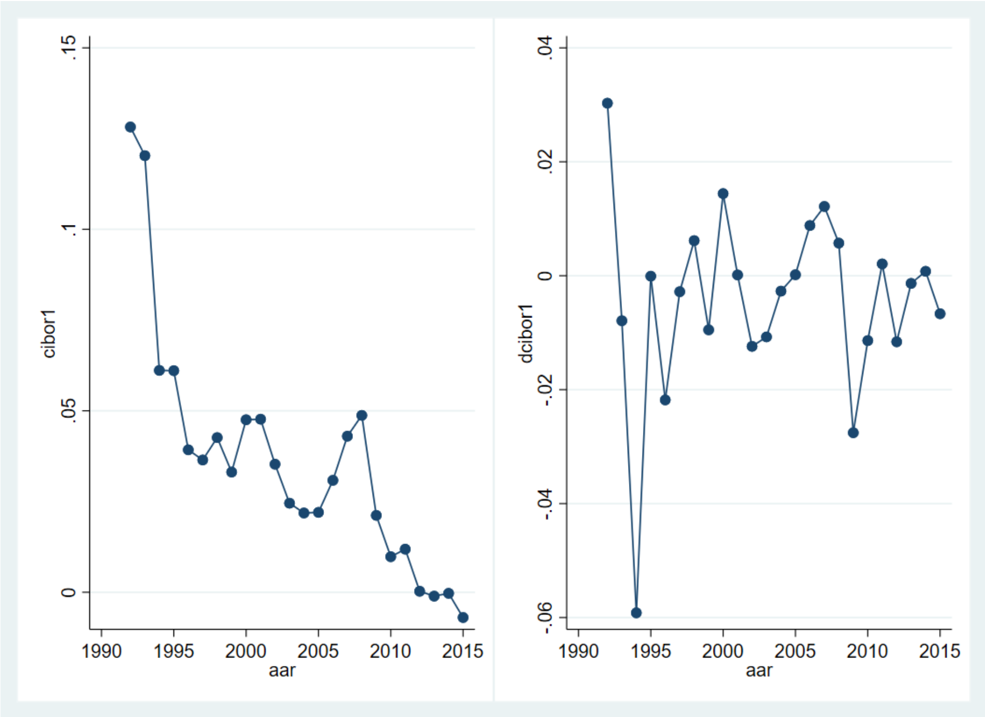

Figure 2: Opg. 1.2: Venstre: CIBOR. Højre: ΔCIBOR

8,2 pct.point. Alle disse fortegn er som forventede, men da vi ikke ved, om der er heteroskedasticitet kan vi ikke fortolke på signifikansen af effekterne.

(b) Det efterspurgte Breusch-Pagan test implementeres i STATA ved at skrive `estat hettest, rhs iid`. Nulhypotesen er, at fejlledet er homoskedastisk. Vi finder en LM test statistik på 2,12, som er χ^2-fordelt med 3 frihedsgrader. P-værdien er 0,55, så vi kan ikke afvise nulhypotesen om homoskedasticitet. Alligevel bliver vi bedt om at benytte robuste standardfejl, hvilket vi hermed gør i denne og de resterende estimationer. Vi ser i anden kolonne af tabel 2, at alle effekterne, der blev beskrevet i det foregående spørgsmål, er signifikante på et 5 pct. niveau. Forskellen mellem de almindelige og de robuste standardfejl er forholdsvis lille, og det ligger også i tråd med, at vi jo ikke kan afvise antagelsen om homoskedasticitetet. Det forhold, at indlån-udlånsraten påvirker væksten i udlånet negativt er ikke evidens for *bankudlånsmekanismen.*

(c) I kolonne 3 i tabel 2 er fixed effects estimaterne vist. Parameter estimaterne ligner OLS estimaterne meget, så det lader ikke til, at den tidsinvariante uobserverbare heterogenitet er korreleret med de forklarende variable. Derfor foretrækker vi en estimation, der både benytter *within* og *between* variationen i data. Valget står dermed mellem pooled OLS og random effects estimatoren, og der

	2.1 OLS	2.2 OLS robust	2.3 FE	2.4 OLS interaktion
$\Delta CIBOR_{t-1}$	-2.519***	-2.519***	-2.642***	-1.557**
	(0.369)	(0.366)	(0.379)	(0.522)
$\Delta \log(BNP_{t-1})$	0.783*	0.783*	0.969**	0.770*
	(0.345)	(0.314)	(0.326)	(0.312)
$\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1}$	-0.082***	-0.082***	-0.089***	-0.084***
	(0.004)	(0.006)	(0.014)	(0.006)
$\Delta CIBOR_{t-1} \times \left(\frac{\text{udlan}}{\text{indlan}}\right)_{it-1}$				-0.844*
				(0.353)
Constant	0.148***	0.148***	0.152***	0.150***
	(0.010)	(0.010)	(0.016)	(0.010)
Obs	3436	3436	3436	3436
Rsq	0.114	0.114	0.061	0.117

Table 2: Opgave 2.1-2.4

vil vi foretrække den sidstnævnte, da den er mest efficient.

(d) Vi tilføjer nu en interaktionseffekt mellem ΔCIBOR_{t-1} og $\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1}$ og estimerer modellen nedenfor med OLS

$$\begin{aligned}\Delta \log(\text{udlån}_{it}) = \delta_0 + \delta_1 \Delta \log(\text{BNP}_{t-1}) + \delta_2 \Delta\text{CIBOR}_{t-1} \\ + \delta_3 \left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1} + \delta_4 \left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1} \Delta\text{CIBOR}_{t-1} + \varepsilon_{it}\end{aligned} \tag{2}$$

Nulhypotesen for om interaktionseffekten mellem ΔCIBOR_{t-1} og $\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1}$ er signifikant kan skrives som: $H_0 : \delta_4 = 0$ mod alternativet $H_A : \delta_4 \neq 0$. t-værdien er givet ved $t = \frac{-0{,}844-0}{0{,}353} = -2,391$ med p-værdien $0,017$. Vi afviser derfor nulhypotesen om, at effekten på bankens vækst i udlånet af interaktionen mellem renten og den individuelle banks udlån-indlånsrate er lig med 0.

Vi tester tosidet, men man kunne også have valgt et enkeltsidet test, idet man nok ville forvente, at banker med højere rate mellem udlån og indlån ikke vil reagere mindre stærkt over for rentestigninger end banker med en lav rate mellem udlån og indlån. Den signifikant negative interaktionseffekt støtter hypotesen om *bankudlånsmekanismen*, eftersom det betyder, at effekten af en rentestigning ikke kun skyldes en lavere låneefterspørgsel blandt bankens kunder. Tværtimod viser interaktionen, at der er heterogene effekter på bankniveau af renteændringer.

Vi kan ikke direkte sammenligne estimatet af β_2 fra model (1) med estimaterne af δ_2, da der kun i den sidste model er et interaktionsled. Dvs. ud over den negative effekt af en renteændring på $-1,557$ er der også en yderligere negativ effekt fra interaktionsleddet mellem renten og udlån-indlånsraten. Hvis man f.eks. udregner effekten af ΔCIBOR_{t-1} for gennemsnittet af $\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1}$ finder vi, at effekten er $\delta_1 + \delta_4 avg(\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1}) = -1,557 - 0,844 \cdot 1,108 = -2,493$, som svarer nogenlunde til parameterestimatet af β_1 i model (1).[1]

Opgave 3

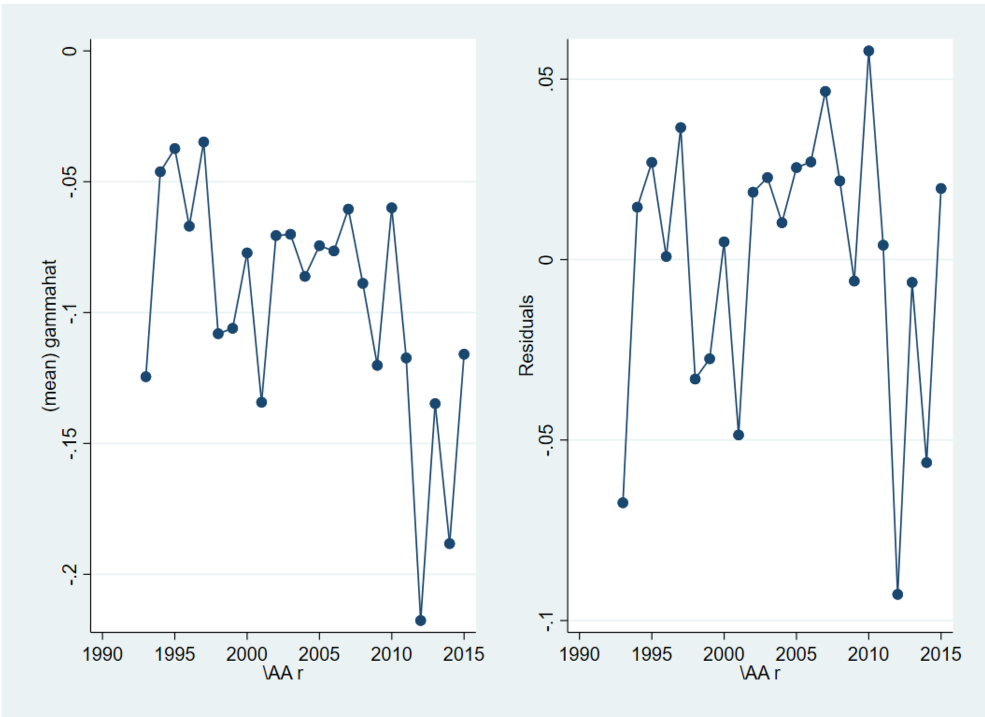

Figure 3: Opg. 3.1: Første trin $\hat{\gamma}$

(a) I den venstre side af figur 3 er $\hat{\gamma}$ vist. Tidsserien ser ikke stationær ud, da det ser ud til, at der kunne være et skift i niveau for 2012-2015. I det hele taget kunne $\hat{\gamma}$ måske godt være faldende

[1] Vi rammer ikke den samme effekt, da parameter estimatet for $\left(\frac{\text{udlån}}{\text{indlån}}\right)_{it-1}$ også er ændret mellem estimation af model (1) og model (2).

over tid. I den højre side af figuren vises tidsserien for $\hat{\gamma}$, hvor vi har taget højde for en lineær tidstrend. Det er fortsat ikke helt klart, men tidsserien ser trods alt mere stationær ud, efter vi har fjernet trenden. Når vi estimerer modellen i ligning (3), bør man derfor som minimum undersøge, hvad det betyder, når man kontrollerer for en lineær tidstrend.

(b) Vi estimerer den følgende model med og uden en lineær tidstrend

$$\hat{\gamma}_t = \mu_0 + \mu_1 \Delta \log(\text{BNP}_{t-1}) + \mu_2 \Delta \text{CIBOR}_{t-1} + e_t \tag{3}$$

Estimaterne er vist i tabel 3. I overensstemmelse med analysen i opgave 2, finder vi i begge kolonner af tabel 3, at effekten af udlån-indlån raten afhænger signifikant negativt af ændringen i CIBOR renten. Med andre ord jo højere rente, desto mere vil banker med høj udlån-indlånsrate reducere væksten i udlånet. I den anden kolonne ser vi, at den lineære tidstrend er signifikant på et 5 pct. niveau. Tidstrenden får effekten af CIBOR renten til at blive en anelse mindre. Effekten af væksten i BNP er insignifikant i begge modeller. Når vi skal teste for autokorrelation benytter vi Breusch-Godfrey testet ved hjælp kommandoen `estat bgodfrey, lag(1)`. Princippet bag testet er at regressere residualet mod de forklarende variable og det laggede residual. I tilfældet uden en tidstrend er LM-statistikken givet ved 4,8, og med en tilhørende p-værdi på 0,03 forkaster vi nulhypotesen om ingen autokorrelation. I tilfældet med en tidstrend er LM-statistikken til gengæld givet ved 1,6. Med den tilhørende p-værdi på 0,20 kan vi ikke forkaste nulhypotesen om ingen autokorrelation. Vi foretrækker modellen med en lineær trend, da vi i dette tilfælde ikke kan afvise, at der ingen autokorrelation er. Desuden var vi jo også klart mere tilbøjelige til at konkludere, at $\hat{\gamma}$ er trend-stationær end stationær.

(c) De to modeller er ikke ens. Hovedligheden mellem de to er, at vi i begge tilfælde kontrollerer for den generelle effekt af CIBOR renten for at isolere den del af rente effekten, der gå via bankernes forskellige karakteristika (målt ved udlån-indlånsraten) for at kunne finde evidens for bankudlånsmekanismen. I et-trins proceduren sker det via et interaktionsled, mens det i to-trinsproceduren sker ved at undersøge om tidvariationen i effekten fra udlån-indlånsraten kan forklares ved CIBOR rentens udvikling i en ren tidsrække regression.

Der er flere grunde til at modellerne ikke er ens. For det første har det betydning for den generelle effekt af CIBOR renten, hvorledes der kontrolleres for bankernes udlån-indlånsrater. I regressionen i det første trin af de to trin (resultaterne er ikke gengivet) finder vi en mindre generel effekt af CIBOR renten i forhold til når ligning (2) estimeres. For det andet tillader vi i andet trin, at effekten af γ_t afhænger af BNP væksten, mens der ikke er nogen interaktion mellem BNP væksten og udlån-indlån-raten i ligning (2).

	3.2 2nd step	3.2 2nd step time-trend
$\Delta CIBOR_{t-1}$	-1.172*	-1.047*
	(0.524)	(0.441)
$\Delta \log(BNP_{t-1})$	0.207	-0.322
	(0.445)	(0.410)
År		-0.004**
		(0.001)
Constant	-0.105***	7.460**
	(0.012)	(2.451)
Obs	23	23
Rsq	0.206	0.471

Table 3: Opgave 3.2

Opgave 4

(a) Opskriv OLS estimatoren og indsæt $y_t = \beta_0 + \beta_1 x_t + u_t$

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{t=1}^{T}(x_t - \bar{x})y_t}{\sum_{t=1}^{T}(x_t - \bar{x})x_t} \Leftrightarrow \\ &= \frac{\sum_{t=1}^{T}(x_t - \bar{x})(\beta_0 + \beta_1 x_t + u_t)}{\sum_{t=1}^{T}(x_t - \bar{x})x_t} \Leftrightarrow \\ &= \beta_1 + \frac{\sum_{t=1}^{T}(x_t - \bar{x})u_t}{SST_x}\end{aligned}$$

hvor $SST_x = \sum_{t=1}^{T}(x_t - \bar{x})x_t$. Tag nu den betingede varians af $\hat{\beta}_1$.

$$\begin{aligned}\mathrm{Var}(\hat{\beta}_1|\boldsymbol{x}) &= \mathrm{Var}\left(\frac{\sum_{t=1}^{T}(x_t - \bar{x})u_t}{SST_x}|\boldsymbol{x}\right) \\ &= SST_x^{-2}\,\mathrm{Var}\left(\sum_{t=1}^{T}(x_t - \bar{x})u_t|\boldsymbol{x}\right)\end{aligned}$$

(b) Vi opskriver ligningen for u_3 og indsætter tidligere perioders u_t succesivt.

$$\begin{aligned} u_3 &= \rho u_2 + v_3 \\ &= \rho(\rho u_1 + v_2) + v_3 \\ &= \rho(\rho(\rho u_0 + v_1) + v_2) + v_3 \\ &= \rho^3 u_0 + \rho^2 v_1 + \rho v_2 + v_3 \\ &= \rho^2 v_1 + \rho v_2 + v_3 \end{aligned}$$

Vi tager udgangspunkt i den næstsidste ligning, hvorefter vi udnytter, at vi har antaget at $u_0 = 0$

$$\begin{aligned} E(u_3 u_1) &= E((\rho^3 u_0 + \rho^2 v_1 + \rho v_2 + v_3)(\rho u_0 + v_1)) \\ &= E((\rho^2 v_1 + \rho v_2 + v_3)(v_1)) \\ &= E(\rho^2 v_1^2 + \rho v_1 v_2 + v_1 v_3) \\ &= \rho^2 E(v_1^2) + \rho E(v_1 v_2) + E(v_1 v_3) \\ &= \rho^2 \sigma_v^2 \end{aligned}$$

(c) Når vi har seriel korrelation i u'erne, så har vi ikke, at variansen af en sum er lig med summen af varianserne. Faktisk får vi kovarianser mellem u_t og u_{t-j}. Det forventes ikke, at man får skrevet den følgende ligning præcist op.

$$\begin{aligned} \mathrm{Var}(\hat{\beta}_1|\boldsymbol{x}) &= SST_x^{-2}\,\mathrm{Var}\left(\sum_{t=1}^{T}(x_t-\bar{x})u_t|\boldsymbol{x}\right) \\ &= SST_x^{-2}\left[\sum_{t=1}^{T}\mathrm{Var}((x_t-\bar{x})u_t|\boldsymbol{x}) + 2\sum_{t=1}^{T}\sum_{t=1}^{T-t}E((x_t-\bar{x})u_t(x_{t+j}-\bar{x})u_{t+j}|\boldsymbol{x})\right] \\ &= SST_x^{-2}\left[\sum_{t=1}^{T}(x_t-\bar{x})^2\,\mathrm{Var}(u_t|\boldsymbol{x}) + 2\sum_{t=1}^{T}\sum_{t=1}^{T-t}(x_t-\bar{x})(x_{t+j}-\bar{x})E(u_t u_{t+j}|\boldsymbol{x})\right] \\ &= SST_x^{-2}\left[\sigma_u^2\sum_{t=1}^{T}(x_t-\bar{x})^2 + 2\sum_{t=1}^{T}\sum_{t=1}^{T-t}(x_t-\bar{x})(x_{t+j}-\bar{x})E(u_t u_{t+j}|\boldsymbol{x})\right] \\ &= \frac{\sigma_u^2}{SST_x} + 2SST_x^{-2}\sum_{t=1}^{T}\sum_{t=1}^{T-t}(x_t-\bar{x})(x_{t+j}-\bar{x})E(u_t u_{t+j}|\boldsymbol{x}) \end{aligned}$$

Vi bemærker, at det første led er den almindelige varians i fraværet af seriel korrelation, mens det andet således er biasen, når der er seriel korrelation. Jo mere seriel korrelation, som målt ved ρ (jf. resultatet i 4 (b)), desto større bias i den almindelige OLS varians (medmindre der ingen seriel korrelation er i x'erne).

Opgave 5

(a) I tabel 4 er resultaterne af simulationseksperimentet vist. Vi ser, at OLS estimatoren er middelret, da gennemsnittet ligger meget tæt på den sande værdi. Autokorrelationen i fejlledet gør tilgengæld, at de almindelige standardfejl er biased. Vi ser, at vi forkaster den sande nulhypotese for ofte, dvs. i 8,1 pct. af gangene. Det betyder, at de almindelige standardfejl er for små, hvilket stemmer godt overens med resultatet i opgave 4 (c), idet $\rho > 0$.

(b) Når vi benytter HAC standardfejlene, der tager højde for autokorrelation (og heteroskedasticitet), så har vi naturligvis det helt samme estimat for β_1, da det kun er standardfejlene, der beregnes anderledes. Som forventet får vi større standardfejl, og vi forkaster således den sande nulhypotese mindre end med de almindelige OLS standardfejl, men vi forkaster stadigvæk i 7,2 pct., som jo ligger et stykke fra det ønskede 5 pct.

(c) Tilslut estimerer vi ADL modellen. Denne model fremkommer ved at trække ρy_{t-1} fra $y_t = \beta_0 + \beta_1 x_t + u_t$. Vi får således

$$\begin{aligned} y_t - \rho y_{t-1} &= \beta_0 + \beta_1 x_t + u_t - \rho(\beta_0 + \beta_1 x_{t-1} + u_{t-1}) \\ y_t &= (1-\rho)\beta_0 + \beta_1 x_t + \rho\beta_1 x_{t-1} + \rho y_{t-1} + u_t - \rho u_{t-1} \\ y_t &= \alpha_0 + \beta_1 x_t + \alpha_2 x_{t-1} + \rho y_{t-1} + \varepsilon_t \end{aligned}$$

hvor $\alpha_0 = (1-\rho)\beta_0$, $\alpha_2 = \rho\beta_1$, $\alpha_3 = \rho$ og $\varepsilon_t = u_t - \rho u_{t-1} = v_t$. Ved at inkludere lags af y og x på højresiden, så får vi fjernet autokorrelationen i fejleddet. Det betyder, at denne estimation er mere præcis (jf. standardfejlene og min og max af estimatet på β_1). Vi har desuden, at andelen af de 1000 simulationer, hvor vi forkaster nulhypotesen, er 4,8 pct., der er tæt på de ønskede 5 pct.

	mean	sd	min	max
bols	1.9932	0.136	1.483	2.493
seols	0.1203	0.008	0.096	0.142
rejectols	0.0810	0.273	0.000	1.000
bnewey	1.9932	0.136	1.483	2.493
senewey	0.1261	0.011	0.093	0.169
rejectnewey	0.0720	0.259	0.000	1.000
badl	1.9940	0.116	1.583	2.391
seadl	0.1161	0.006	0.098	0.136
rejectadl	0.0480	0.214	0.000	1.000
No. obs	1000			

Table 4: Opgave 5 (a)-(c)