

Institutt for samfunnsøkonomi

## Eksamensoppgave i SØK3001 – Økonometri I

**Faglig kontakt under eksamen: Bjarne Strøm**

**Tlf.: 73 59 19 33**

**Eksamensdato:** 31. mai 2017

**Eksamenstid (fra-til):** 5 timer (09.00-14.00)

**Sensurdato:** 22. juni 2017

**Hjelpemiddelkode/Tillatte hjelpemidler:** C /Flg formelsamling: Knut Sydsæter, Arne Strøm og Peter Berck (2006): Matematisk formelsamling for økonomer, 4utg. Gyldendal akademiske. Knut Sydsæter, Arne Strøm, og Peter Berck (2005): Economists' mathematical manual, Berlin. Calculator Casio fx-82ES PLUS, Casio fx-82EX Citizen SR-270x, SR-270X College eller HP 30S.

**Målform/språk:** Bokmål, nynorsk og engelsk

**Antall sider (uten forside):** 12

**Antall sider vedlegg:** 5 tabeller

**Informasjon om trykking av eksamensoppgave**

**Originalen er:**

**1-sidig**       **2-sidig**

**sort/hvit**       **farger**

**skal ha flervalgskjema**

**Kontrollert av:**

\_\_\_\_\_  
Dato

\_\_\_\_\_  
Sign

## Oppgave 1.

Vi har en enkel lineær økonometrisk modell der  $y$  er avhengig variabel og  $x$  er forklaringsvariabel.

a) Forklar hva som menes med at  $x$  er en endogen forklaringsvariabel i den økonometriske modellen og gjør *kort* rede for mulige årsaker til dette.

Tabell 1 oppgir estimerte koeffisienter, avhengig variabel, forklaringsvariabler, estimeringsmetode, estimerte standardavvik og en del andre opplysninger fra en rekke estimerte sammenhenger basert på datamateriale for variablene  $y$ ,  $x$ ,  $z1$  og  $z2$ .  $z1$  og  $z2$  er instrumentvariable. Variablene  $e2$  og  $e3$  er residualene basert på estimatene i henholdsvis kolonne (2) og kolonne (3).

b) Forklar hvilke forutsetninger som må være oppfylt for at metodene benyttet i henholdsvis kolonne (1) og (2) skal gi konsistente estimatorer for effekten av  $x$  på  $y$ .

c) Benytt informasjonen i Tabell 1 til å konstruere og gjennomføre en F-test for relevansen av de to instrumentvariablene  $z1$  og  $z2$ .

d) Forklar hvordan du kan benytte informasjonen i Tabell 1 til å teste empirisk om  $x$  er en endogen forklaringsvariabel og utfør testen.

e) Hva menes med at den økonometriske modellen er overidentifisert? Forklar hvordan du kan benytte informasjonen i Tabell 1 til å teste for overidentifikasjon og utfør testen.

Tabell 1. Estimeringsresultater. Estimerte standardavvik i parentes under koeffisientene. Avhengig variabel er angitt øverst i hver kolonne.

VARIABLE	(1) y	(2) y	(3) x	(4) y	(5) e2
constant	0.9788 (0.088)	1.1376 (0.116)	0.1947 (0.079)	1.1376 (0.080)	0.0188 (0.106)
x	1.7034 (0.089)	1.0399 (0.194)		1.0399 (0.133)	
z1			0.5700 (0.089)		0.0881 (0.118)
z2			0.2068 (0.077)		0.1817 (0.103)
e3				0.9957 (0.163)	
R <sup>2</sup>	0.785		0.3336	0.845	0.0363
Observasjoner	100	100	100	100	100
Estimeringsmetode	OLS	2SLS	OLS	OLS	OLS

## Oppgave 2.

Tilbudet av ris modelleres som en lineær regresjonsligning mellom logaritmen til arealet tilplantet med ris ( $la$ ) og logaritmen til prisen på ris ( $lp$ ). Tabell 2 presenterer ulike versjoner av tilbudsrelasjonen basert på årsdata fra Bangladesh.  $la_1$  og  $lp_1$  er et års lag i henholdsvis  $la$  og  $lp$ .  $dla$  og  $dlp$  er førstedifferensen av henholdsvis  $la$  og  $lp$ .  $\hat{e}$  er residualene basert på estimerte koeffisienter i kolonne (1), mens  $\hat{e}_D$  er residualene basert på estimerte koeffisienter i kolonne (3).

- Beregn kort- og langsiktig priselastisitet for tilbudet av ris i de estimerte tilbudsrelasjonene i Tabell 2.
- Test en hypotese om at den langsiktige priselastisiteten er lik 0.
- Hva menes med at restleddet er seriekorrelert og pek på årsaker og konsekvenser av seriekorrelasjon.
- Forklar hvordan du kan teste for seriekorrelasjon i tilbudsrelasjonene og utfør testen når du benytter informasjonen gitt i Tabell 2. Bruk disse resultatene sammen med annen informasjon til å evaluere troverdigheten av tilbudsrelasjonene rapportert i Tabell 2.

Tabell 2. Estimeringsresultater. Estimeringsmetode er OLS. Estimerte standardavvik i parentes. Avhengig variabel er angitt øverst i hver kolonne.

VARIABLE	(1) la	(2) ehat	(3) la	(4) ehatD	(5) dla
la_1			0.4043 (0.1666)		-0.5957 (0.1666)
lp	0.7761 (0.2775)	-0.0152 (0.2807)	0.7766 (0.2798)	0.0152 (0.2779)	
lp_1			-0.6109 (0.2966)	0.1011 (0.2844)	0.1658 (0.3844)
ehat_1		0.3986 (0.1673)			
ehatD_1				0.0656 (0.1868)	
dlp					0.7766 (0.2798)
Constant	3.8933 (0.0613)	0.0133 (0.0608)	2.3662 (0.6557)	-0.0228 (0.0713)	2.3662 (0.6557)
Observasjoner	34	33	33	32	33
R <sup>2</sup>	0.196	0.163	0.304	0.009	0.583

## Oppgave 3.

Tabell 3 viser estimerte koeffisienter i ulike varianter av en log-linearisert versjon av en underliggende Cobb-Douglas produktfunksjon estimert med data fra nasjonalregnskapstall fra 82 land fra hvert år for perioden 1960 til 1987.  $ly$  er logaritmen til produksjonen,  $ln$  er logaritmen til sysselsettingen og  $lk$  er logaritmen til kapitalen.

- a) Forklar i) hvilke av de estimerte ligninger ivaretar permanente forskjeller i produksjonsnivå mellom landene, og ii) hvilke av de estimerte ligninger kan sies å ivareta *både* permanente forskjeller i produksjonsnivå mellom landene *og* virkningen av teknologisk endring felles for alle landene.
- b) Forklar hvilke restriksjoner, hvis noen, som er pålagt koeffisientene på tvers av land og over tid i modellene rapportert i Tabell 3.
- c) Hva er elastisiteten av produksjonen med hensyn på sysselsettingen og kapitalen i modellene estimert i Tabell 3, og hvilke av disse modellene vil du foretrekke? Konstruer 95% konfidensintervall rundt elastisiteten av produksjonen med hensyn på kapital og sysselsetting i modellene.
- d) Bruk resultatene i Tabell 3 til å teste hypotesen om konstant skala-avkastning. Hva er din tolking av denne hypotesen?
- e) En kommentator foreslår å utvide alle spesifikasjonene i Tabell 3 med en dummyvariabel DEM, lik 1 hvis landet var et demokrati i 1960 og 0 ellers for å ivareta betydningen av politiske institusjoner for økonomisk utvikling. Kommenter dette forslaget.
- f) En annen kommentator foreslår å utvide modellene i Tabell 3 med et interaksjonsledd mellom  $lk$  og DEM. Kommenter dette forslaget og forklar hvordan du ville tolke koeffisientene i modellene i denne utvidede versjonen.
- g) En tredje kommentator foreslår at du utvider modellene i Tabell 3 med antallet patenter hvert år i hele verden, WP, for å ivareta virkningen av teknologisk utvikling. Kommenter dette forslaget.

Tabell 3. Estimeringsresultater. Metode: OLS. Estimerte standardavvik i parentes. Avhengig variabel er angitt øverst i hver kolonne.

VARIABLE	(1) ly	(2) ly-ln	(3) ly	(4) ly-ln	(5) ly	(6) ly-ln
ln	0.1373 (0.0068)	-0.0003 (0.0054)	0.2908 (0.0277)	-0.1230 (0.0191)	0.1602 (0.0327)	-0.3250 (0.0315)
lk	0.8624 (0.0049)		0.5862 (0.0106)		0.5148 (0.0123)	
lk-ln		0.8624 (0.0049)		0.5862 (0.0106)		0.5148 (0.0123)
Constant	0.3784 (0.0983)	0.3784 (0.0983)	5.5950 (0.2224)	5.5950 (0.2224)	9.3374 (0.5003)	9.3374 (0.5003)
Country dummies	No	No	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	Yes	Yes
Observasjoner	2,296	2,296	2,296	2,296	2,296	2,296
R <sup>2</sup>	0.963	0.933	0.996	0.993	0.996	0.993

Nynorsk

Oppgåve 1.

Vi har ein enkel lineær økonometrisk modell der  $y$  er avhengig variabel og  $x$  er forklaringsvariabel.

a) Forklar kva som meines med at  $x$  er ein endogen forklaringsvariabel i den økonometriske modellen og gjer *kort* greie for moglege årsaker til dette.

Tabell 1 syner estimerte koeffisientar, avhengig variabel, forklaringsvariable, estimeringsmetode, estimerte standardavvik og ein del andre opplysningar frå ei rekkje estimerte samanhenger basert på datamateriale for variablane  $y$  og  $x$  og  $z1$  og  $z2$ .  $z1$  og  $z2$  er instrumentvariable. Variablane  $e2$  og  $e3$  er residualane basert på estimata i høvesvis kolonne (2) og kolonne (3).

b) Forklar kva for føresetnader som må vere oppfylte for at metodane nytta i høvesvis kolonne (1) og (2) skal gje konsistente estimatorar for effekten av  $x$  på  $y$ .

c) Nytt informasjonen i Tabell 1 til å konstruere og gjennomføre ein F-test for relevansen av dei to instrumentvariablane  $z1$  og  $z2$ .

d) Forklar korleis du kan nytte informasjonen i Tabell 1 til å teste empirisk om  $x$  er ein endogen forklaringsvariabel og utfør testen.

e) Kva meines med at den økonometriske modellen er overidentifisert? Forklar korleis du kan nytte informasjonen i Tabell 1 til å teste for overidentifikasjon og utfør testen.

Tabell 1. Estimeringsresultat. Estimerte standardavvik i parentes under koeffisientane. Avhengig variabel er gjeve øverst i kvar kolonne.

VARIABLE	(1) y	(2) y	(3) x	(4) y	(5) e2
Constant	0.9788 (0.088)	1.1376 (0.116)	0.1947 (0.079)	1.1376 (0.080)	0.0188 (0.106)
x	1.7034 (0.089)	1.0399 (0.194)		1.0399 (0.133)	
z1			0.5700 (0.089)		0.0881 (0.118)
z2			0.2068 (0.077)		0.1817 (0.103)
e3				0.9957 (0.163)	
R <sup>2</sup>	0.785		0.3336	0.845	0.0363
Observasjonar	100	100	100	100	100
Estimeringsmetode	OLS	2SLS	OLS	OLS	OLS

## Oppgave 2.

Tilbudet av ris er modellert som ei lineær regresjonslikning mellom logaritmen til arealet planta med ris ( $la$ ) og logaritmen til prisen på ris ( $lp$ ). Tabell 2 syner ulike versjonar av tilbodsrelasjonen basert på årsdata frå Bangladesh.  $la_1$  og  $lp_1$  er eit års lag i høvesvis  $la$  og  $lp$ .  $dla$  og  $dlp$  er fyrstedifferensen av høvesvis  $la$  og  $lp$ .  $e\hat{a}$  er residualane basert på estimerte koeffisientar i kolonne (1), medan  $e\hat{a}D$  er residualane basert på estimerte koeffisientar i kolonne (3).

- Rekn ut kort og langsiktig priselastisitet for tilbudet av ris i dei estimerte tilbodsrelasjonane i Tabell 2.
- Test ein hypotese om at den langsiktige priselastisiteten er lik 0.
- Kva meiner du med at restleddet er seriekorrelet og peik på årsaker og konsekvensar av seriekorrelasjon
- Forklar korleis du kan teste for seriekorrelasjon i tilbodsrelasjonane og utfør testen når du nyttar informasjonen gjeve i Tabell 2. Bruk desse resultatane saman med anna informasjon til å evaluere truverdet til tilbodsrelasjonane rapportert i Tabell 2.

Tabell 2. Estimeringsresultat. Estimeringsmetode er OLS. Estimerte standardavvik i parentes. Avhengig variabel er gjeve øverst i kvar kolonne.

VARIABLE	(1) la	(2) e $\hat{a}$	(3) la	(4) e $\hat{a}D$	(5) dla
la_1			0.4043 (0.1666)		-0.5957 (0.1666)
lp	0.7761 (0.2775)	-0.0152 (0.2807)	0.7766 (0.2798)	0.0152 (0.2779)	
lp_1			-0.6109 (0.2966)	0.1011 (0.2844)	0.1658 (0.3844)
e $\hat{a}$ _1		0.3986 (0.1673)			
e $\hat{a}D$ _1				0.0656 (0.1868)	
dlp					0.7766 (0.2798)
Constant	3.8933 (0.0613)	0.0133 (0.0608)	2.3662 (0.6557)	-0.0228 (0.0713)	2.3662 (0.6557)
Observasjoner	34	33	33	32	33
R <sup>2</sup>	0.196	0.163	0.304	0.009	0.583

## Oppgave 3.

Tabell 3 syner estimerte koeffisientar i ulike variantar av ein log-linearisert versjon av ein underliggjande Cobb-Douglas produktfunksjon estimert med data frå nasjonalrekneskapstal frå 82 land frå kvart år for perioden 1960 til 1987.  $ly$  er logaritmen til produksjonen,  $ln$  er logaritmen til sysselsetjinga og  $lk$  er logaritmen til kapitalen.

- a) Forklar i) kva for estimerte likningar som tek vare på permanente skilnader i produksjonsnivå mellom landa, og ii) kva for estimerte likningar som kan seiast å ta vare på *både* permanente skilnader i produksjonsnivå mellom landa *og* verknaden av teknologisk forandring felles for alle landa.
- b) Forklar kva for restriksjonar, dersom nokon, som er pålagt koeffisientane på tvers av land og over tid i modellane rapportert i Tabell 3.
- c) Kva er elastisiteten av produksjonen med omsyn på sysselsetjinga og kapitalen i modellane estimert i Tabell 3, og kva for ein av desse modellane vil du føretrekkje? Konstruer 95% konfidensintervall rundt elastisiteten av produksjonen med omsyn på kapital og sysselsetjing i desse modellane.
- d) Bruk resultatane i Tabell 3 til å teste hypotesen om konstant skala-avkastning. Kva er din tolking av denne hypotesen?
- e) Ein kommentator har framlegg om å utvide alle spesifikasjonane i Tabell 3 med ein dummyvariabel DEM, lik 1 dersom landet var eit demokrati i 1960 og 0 elles for å ta vare på verknaden av politiske institusjonar for økonomisk utvikling. Kommenter dette framlegget.
- f) Ein annan kommentator har framlegg om å utvide modellane i Tabell 3 med eit interaksjonsledd mellom  $lk$  og DEM. Kommenter dette framlegget og forklar korleis du ville tolke koeffisientane i modellane i denne utvida versjonen.
- g) Ein tredje kommentator har framlegg om å utvide modellane i Tabell 3 med talet på patentar kvart år i heile verda, WP, for å ta vare på verknaden av teknologisk utvikling. Kommenter dette framlegget.



Tabell 3. Estimeringsresultat. Metode: OLS. Estimerte standardavvik i parentes. Avhengig variabel er gjeve øverst i kvar kolonne.

VARIABLE	(1) ly	(2) ly-ln	(3) ly	(4) ly-ln	(5) ly	(6) ly-ln
ln	0.1373 (0.0068)	-0.0003 (0.0054)	0.2908 (0.0277)	-0.1230 (0.0191)	0.1602 (0.0327)	-0.3250 (0.0315)
lk	0.8624 (0.0049)		0.5862 (0.0106)		0.5148 (0.0123)	
lk-ln		0.8624 (0.0049)		0.5862 (0.0106)		0.5148 (0.0123)
Constant	0.3784 (0.0983)	0.3784 (0.0983)	5.5950 (0.2224)	5.5950 (0.2224)	9.3374 (0.5003)	9.3374 (0.5003)
Country dummies	No	No	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	Yes	Yes
Observasjonar	2,296	2,296	2,296	2,296	2,296	2,296
R <sup>2</sup>	0.963	0.933	0.996	0.993	0.996	0.993

**Engelsk****Question 1.**

We have a simple linear econometric model where  $y$  is dependent variable and  $x$  is explanatory variable

a) Explain what is meant by the statement that  $x$  is an endogeneous explanatory variable and explain *briefly* possible reasons for that.

Table 1 reports estimated coefficients, dependent variable, explanatory variables, estimation method, estimated standard errors and some other information for a number of estimated relationship based on data for the variables  $y$ ,  $x$ ,  $z1$  and  $z2$ .  $z1$  and  $z2$  are instrumental variables. The variables  $e2$  and  $e3$  are residuals based on the estimates obtained in column (2) and column (3), respectively.

b) Explain what assumptions are needed to ensure that the methods used in column (1) and column (3), respectively, give consistent estimators for the effect of  $x$  on  $y$ .

c) Use the information in Table 1 to construct and perform an F-test for the relevance of the instruments  $z1$  and  $z2$ .

d) Explain how you can use the information in Table 1 to test empirically whether  $x$  is an endogeneous variable and perform the test.

e) What is meant by the statement that the econometric model is overidentified? Explain how you can use the information in Table 1 to test for overidentification and perform the test.

Table 1. Estimation results. Estimated standard errors in parentheses. Dependent variable denoted at the top of each column.

Variables	(1) y	(2) y	(3) x	(4) y	(5) e2
constant	0.9788 (0.088)	1.1376 (0.116)	0.1947 (0.079)	1.1376 (0.080)	0.0188 (0.106)
x	1.7034 (0.089)	1.0399 (0.194)		1.0399 (0.133)	
z1			0.5700 (0.089)		0.0881 (0.118)
z2			0.2068 (0.077)		0.1817 (0.103)
e3				0.9957 (0.163)	
R <sup>2</sup>	0.785		0.3336	0.845	0.0363
observations	100	100	100	100	100
Estimation method	OLS	2SLS	OLS	OLS	OLS

**Question 2.**

The supply equation for rice is modeled as a linear regression equation between log of area planted (la) and the log of price of rice (lp). Table 2 presents different versions of the supply relationship estimated on yearly data from Bangladesh. la\_1 and lp\_1 are one year lags in la and lp, respectively. ehat is the residuals based on the estimated coefficients in column (1), while ehatD is the residuals based on estimated coefficients in column (3). ehat\_1 and ehatD\_1 are one year lags of ehat and ehatD, respectively.

- Compute the short and long run price elasticities of rice supply using the estimated supply equations reported in Table 2.
- Test the hypothesis that the long run price elasticity is 0.
- Explain what is meant by serially correlated error term, and explain possible reasons for, and consequences of serial correlation.
- Explain how you can test for serially correlated errors in the supply equations and perform the test using the information given in Table 2. Use these results together with other information based on Table 2 to evaluate the credibility of the supply equations reported in Table 2.

Table 2. Estimation results. All equations estimated by OLS. Estimated standard errors in parenthesis. Dependent variable denoted at the top of each column.

VARIABLES	(1) la	(2) ehat	(3) la	(4) ehatD	(5) dla
la_1			0.4043 (0.1666)		-0.5957 (0.1666)
lp	0.7761 (0.2775)	-0.0152 (0.2807)	0.7766 (0.2798)	0.0152 (0.2779)	
lp_1			-0.6109 (0.2966)	0.1011 (0.2844)	0.1658 (0.3844)
ehat_1		0.3986 (0.1673)			
ehatD_1				0.0656 (0.1868)	
dlp					0.7766 (0.2798)
Constant	3.8933 (0.0613)	0.0133 (0.0608)	2.3662 (0.6557)	-0.0228 (0.0713)	2.3662 (0.6557)
Observations	34	33	33	32	33
R <sup>2</sup>	0.196	0.163	0.304	0.009	0.583

## Question 3.

Table 3 shows estimated coefficients in different versions of a log-linearized version of an underlying Cobb-Douglas production function estimated using aggregate national accounts data from 82 countries for each year for the period 1960 to 1987.  $\ln y$  is the log of output,  $\ln n$  is the log of employment and  $\ln k$  is the log of capital.

- a) Explain i) which of the estimated equations account for permanent differences in output between the countries, and ii) which of the estimated equations can be said to account for *both* permanent differences in output between the countries *and* technological development common to all countries.
- b) Explain what restrictions, if any, are imposed on the coefficients across countries and over time in the models reported in Table 3.
- c) What are the elasticities of output with respect to employment and capital, in the models estimated in Table 3 and which of these models would you prefer? Construct 95 percent confidence interval for the elasticities of output with respect to capital and employment in these models.
- d) Use the results reported in Table 3, to test the hypothesis of constant returns to scale. What is your interpretation of this hypothesis?
- e) A commentator suggests to extend all specifications in Table 3 with a dummy variable DEM equal to 1 if the country was a democracy in 1960 and 0 otherwise to account for the effect of political institutions on economic development. Comment on this proposal.
- f) Another commentator suggest to extend the specifications in Table 3 with an interaction term between  $\ln k$  and DEM. Comment on this proposal and explain how you would interpret the model coefficients in this extended version.
- g) A third commentator suggest that you extend the models in Table 3 by the number of patents each year in the world, WP, to capture the effect of technological development. Comment on this proposal.

Table 3. Estimation results. Method: OLS. Estimated standard errors in parenthesis. Dependent variable denoted at the top of each column

VARIABLES	(1) ly	(2) ly-ln	(3) ly	(4) ly-ln	(5) ly	(6) ly-ln
ln	0.1373 (0.0068)	-0.0003 (0.0054)	0.2908 (0.0277)	-0.1230 (0.0191)	0.1602 (0.0327)	-0.3250 (0.0315)
lk	0.8624 (0.0049)		0.5862 (0.0106)		0.5148 (0.0123)	
lk-ln		0.8624 (0.0049)		0.5862 (0.0106)		0.5148 (0.0123)
Constant	0.3784 (0.0983)	0.3784 (0.0983)	5.5950 (0.2224)	5.5950 (0.2224)	9.3374 (0.5003)	9.3374 (0.5003)
Country dummies	No	No	Yes	Yes	Yes	Yes
Year dummies	No	No	No	No	Yes	Yes
Observations	2,296	2,296	2,296	2,296	2,296	2,296
R <sup>2</sup>	0.963	0.933	0.996	0.993	0.996	0.993