

Institutt for samfunnsøkonomi

## **Eksamensoppgave i SØK3515 / SØK8615** **Mikro- og paneldataøkonometri**

**Faglig kontakt under eksamen: Bjarne Strøm**

**Tlf.: 73 59 19 33**

**Eksamensdato:** 11.12.2014

**Eksamenstid (fra-til):** 6 timer (09.00-15.00)

**Sensurdato:** 12.1.2015

**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.  
Godkjent kalkulator Casio fx-82ES PLUS, Citizen SR-270x, SR-270X College eller HP 30S.

**Annen informasjon:**

**Målform/språk:** Bokmål og nynorsk

**Antall sider (uten forside):** 11

**Antall sider vedlegg:** 5 (tabeller)

Bokmål:

SØK3515/8615 Mikro- og paneldataøkonometri

Oppgave 1.

Tabell 1 viser resultatene fra estimering av etterspørselsfunksjon for flyreiser basert på årlige data fra 1149 amerikanske flyruter for årene 1997-2000.  $l_{passen}$  er logaritmen til antall passasjerer på ruten,  $l_{fare}$  er logaritmen til gjennomsnittlig billettpris på ruten,  $l_{dist}$  er logaritmen til rutelengden målt i km,  $l_{distsq}$  er kvadratet av logaritmen til rutelengden, mens  $concen$  er markedskonsentrasjonen på ruten, målt som andelen av det totale antall flygninger som utføres av det største flyselskapet. I alle estimeringene er dummyvariable for 1998, 1999 og 2000 inkludert

Tabell 1. Regresjonsresultater. Estimerte standardavvik i parentes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metode	OLS	OLS med cluster-korrigerede standardavvik	IV	OLS	Fixed Effects	Random Effects	Fixed effects IV	Fixed effects
Avhengig variabel	$l_{passen}$	$l_{passen}$	$l_{passen}$	$l_{fare}$	$l_{passen}$	$l_{passen}$	$l_{passen}$	$l_{fare}$
Forklaringsvariabel								
$l_{fare}$	-0.565 (0.037)	-0.565 (0.067)	-1.777 (0.236)		-1.155 (0.023)	-1.102 (0.022)	-0.302 (0.277)	
$l_{dist}$	-1.549 (0.327)	-1.549 (0.698)	-2.499 (0.406)	-0.902 (0.128)		-1.971 (0.647)		
$l_{distsq}$	0.123 (0.025)	0.123 (0.052)	0.231 (0.035)	0.103 (0.010)		0.171 (0.049)		
$concen$				0.360 (0.030)				0.169 (0.029)
Instrument variabel			$concen$				$concen$	
Tverrsnittsenhet	flyrute	flyrute	flyrute	flyrute	flyrute	flyrute	flyrute	flyrute
Antall tverrsnittsenheter	1149	1149	1149	1149	1149	1149	1149	1149
Obs	4596	4596	4596	4596	4596	4596	4596	4596

a) Gi en kortfattet tolkning av koeffisientene foran variablene.

b) Forklar framgangsmåten for estimeringen av de ulike variantene av etterspørselsfunksjonen rapportert i tabell 1 og gi en presis beskrivelse av under hvilke forutsetninger de ulike variantene gir konsistente estimater for priselastisiteten for flyreiser. Hvilken av variantene ville du feste mest lit til? Begrunn svaret.

- c) Beregn et 95% konfidensintervall rundt priselastisiteten for flyreiser i de ulike variantene. Test en hypotese om priselastisitet lik -1.
- d) En kommentator foreslår å utvide modellen i kolonne (5) med et lag i variabelen *lpassen* som ekstra forklaringsvariabel. Drøft konsekvensene av dette for egenskapene ved estimeringsmetoden benyttet i kolonne (5) og foreslå eventuelt en alternativ metode.

## Oppgave 2.

Det er estimert en multinomisk logitmodell for unge amerikanske menn, sin yrkesstatus i 1988. Variabelen status tar verdien 1 hvis mannen er i utdanning, 2 hvis mannen verken er i jobb eller under utdanning og 3 hvis mannen er i inntektsgivende arbeid. Variabelen educ angir antall års fullført utdanning, variabelen exper angir antall års jobberfaring, *expersq* er antall år jobberfaring kvadrert, mens variabelen black tar verdien 1 hvis mannen er farget og 0 hvis han ikke er farget. Tabell 2 viser estimerte koeffisienter og en del andre opplysninger oppnådd med estimering av to varianter av modellen med maximum likelihood-metoden.

- a) Gjør rede for hvordan modellen er formulert og gi en kort beskrivelse av estimeringsopplegget.
- b) Resultatene i panel A og B gir estimerte koeffisienter, standardavvik og andre opplysninger fra to ulike formuleringer av modellen. Undersøk om modellen i panel B representerer en tillatelig forenkling av modellen i panel A.
- c) Hvordan vil du tolke og bruke de estimerte koeffisientene?
- d) Vis hvordan du med utgangspunkt i opplysningene i tabell 2 panel A vil gå fram (ikke nødvendig å regne ut eksakt tall) for å beregne forskjellen i sannsynlighet for å være i jobb i 1988 mellom to menn med 12 års utdanning, og 1 års jobberfaring der den ene er farget og den andre er ikke-farget.
- e) Vis hvordan du med utgangspunkt i opplysningene i tabell 2 panel A vil beregne sannsynligheten (ikke nødvendig å regne ut eksakt tall) for å være under utdanning i 1988 for en farget mann med 12 års utdanning og 1 års jobberfaring.
- f) Du får oppgitt at McFaddens pseudo  $R^2$  er beregnet til 0.2433 og 0.1203 i henholdsvis modellen i panel A og panel B. Forklar hvordan disse er beregnet.

Tabell 2. Estimerte koeffisienter. Estimerte standardavvik i parentes.

Arbeidsmarkedsstatus	A		B	
	Status=2	Status=3	Status=2	Status=3
Forklaringsvariable				
EDUC	-0.674 (0.070)	-0.315 (0.065)	-0.759 (0.061)	-0.509 (0.054)
EXPER	-0.106 (0.173)	0.849 (0.157)		
EXPERSQ	-0.013 (0.025)	-0.077 (0.023)		
BLACK	0.813 (0.303)	0.311 (0.282)	1.059 (0.293)	0.203 (0.272)
CONSTANT	10.28 (1.13)	5.54 (1.09)	10.94 (0.912)	9.694 (0.846)
LOG LIKELIHOOD	-907.85723		-1055.339	
Antall observasjoner	1717		1717	

## Oppgave 3.

Du ønsker å studere sammenhengen mellom studenters arbeidsinnsats, initialt kunnskapsnivå og en del andre variable. Du har gjennomført en spørreundersøkelse blant 800 studenter i samfunnsøkonomi der studentene angir hvor mange timer i uken de i gjennomsnitt bruker på studiet. Du har registeropplysninger om hver enkelt students gjennomsnittskarakter fra videregående skole som du bruker som mål på den enkelte students initiale kunnskapsnivå. I tillegg har du opplysninger om alder, antall semester i studiet samt registeropplysninger om foreldrenes inntektsnivå for hver enkelt student. Anta at disse variablene er de relevante variablene som forklarer studentenes arbeidsinnsats. Det viser seg imidlertid at 200 av de 800 studentene ikke besvarer spørreskjemaet. Gjør rede for hvilke utfordringer dette reiser for muligheten til å estimere en troverdig arbeidsinnsatsrelasjon for studentene og hvordan du ville håndtert problemet.

## Oppgave 1.

Tabell 1 viser resultat frå estimering av etterspørselsfunksjon for flyreiser basert på årlege data frå 1149 amerikanske flyruter for årene 1997-2000.  $lpassen$  er logaritmen til talet på passasjerar på ruta,  $lfare$  er logaritmen til gjennomsnittlig billettpris på ruta,  $ldist$  er logaritmen til rutelengda målt i km,  $ldistsq$  er kvadratet av logaritmen til rutelengda, mens  $concen$  er markedskonsentrasjonen på ruta, målt som andelen av det totale talet på flygingar som utføres av det største flyselskapet. I alle estimeringane er dummyvariable for 1998, 1999 og 2000 inkludert.

Tabell 1. Regresjonsresultat. Estimerte standardavvik i parentes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Metode	OLS	OLS med cluster-korrigerede standardavvik	IV	OLS	Fixed Effects	Random Effects	Fixed effects IV	Fixed effects
Avhengig variabel	$lpassen$	$lpassen$	$lpassen$	$lfare$	$lpassen$	$lpassen$	$lpassen$	$lfare$
Forklaringsvariabel								
$lfare$	-0.565 (0.037)	-0.565 (0.067)	-1.777 (0.236)		-1.155 (0.023)	-1.102 (0.022)	-0.302 (0.277)	
$ldist$	-1.549 (0.327)	-1.549 (0.698)	-2.499 (0.406)	-0.902 (0.128)		-1.971 (0.647)		
$ldistsq$	0.123 (0.025)	0.123 (0.052)	0.231 (0.035)	0.103 (0.010)		0.171 (0.049)		
$concen$				0.360 (0.030)				0.169 (0.029)
Instrument variabel			$concen$				$concen$	
Tverrsnitt-Eining	flyrute	flyrute	flyrute	flyrute	flyrute	flyrute	flyrute	flyrute
Talet på tverrsnitteiningar	1149	1149	1149	1149	1149	1149	1149	1149
Obs	4596	4596	4596	4596	4596	4596	4596	4596

a) Gi ei kortfatta tolking av koeffisientane føre variablane.

b) Forklar framgangsmåten for estimeringa av dei ulike variantane av etterspørselsfunksjonen rapportert i tabell 1 og gi ein presis omtale av under kva for føresetnader dei ulike variantane gir konsistente estimat for priselastisiteten for flyreiser. Kva for variantar ville du feste mest lit til? Grunnge svaret.

c) Rekn ut eit 95% konfidensintervall rundt priselastisiteten for flyreiser i dei ulike variantane. Test ein hypotese om priselastisitet lik -1.

d) Ein kommentator kjem med framlegg om å utvide modellen i kolonne (5) med eit lag i variabelen *lpassen* som ekstra forklaringsvariabel. Drøft konsekvensane av dette for eigenskapane ved estimeringsmetoden nytta i kolonne (5) og foreslå eventuelt ein alternativ metode.

## Oppgave 2.

Det er estimert ein multinomisk logitmodell for unge amerikanske menn sin yrkesstatus i 1988. Variabelen status tar verdien 1 dersom mannen er i utdanning, 2 dersom mannen verken er i jobb eller under utdanning og 3 dersom mannen er i inntektsgevande arbeid. Variabelen educ er talet på år fullført utdanning, variabelen exper er talet på år med jobbbrøynsle, *expersq* er talet på år med jobbbrøynsle kvadrert, medan variabelen black tar verdien 1 dersom mannen er farga og 0 dersom han ikkje er farga. Tabell 2 viser estimerte koeffisientar og en del andre opplysningar oppnådd med estimering av to variantar av modellen med maximum likelihood-metoden.

a) Gjer greie for korleis modellen er formulert og gi ein kort omtale av estimeringsopplegget.

b) Resultata i panel A og B gjev estimerte koeffisientar, standardavvik og andre opplysningar frå to ulike formuleringar av modellen. Undersøk om modellen i panel B representerer ei tillateleg forenkling av modellen i panel A.

c) Korleis vil du tolke og bruke dei estimerte koeffisientane?

d) Vis korleis du med utgangspunkt i opplysningane i tabell 2 panel A vil gå fram (ikkje nødvendig å rekne ut eksakt tal) for å rekne ut skilnaden i sannsyn for å være i jobb i 1988 mellom to menn med 12 års utdanning, og 1 års jobbbrøynsle der den eine er farga og den andre er ikkje-farga.

e) Vis korleis du med utgangspunkt i opplysningane i tabell 2 panel A vil rekne ut sannsyn (ikkje nødvendig å regne ut eksakt tal) for å være under utdanning i 1988 for ein farga mann med 12 års utdanning og 1 års jobbbrøynsle.

f) Du får oppgitt at McFaddens pseudo  $R^2$  er rekna ut til 0.2433 og 0.1203 i høvesvis modellen i panel A og panel B. Forklar korleis desse er rekna ut.

Tabell 2. Estimerte koeffisientar. Estimerte standardavvik i parentes.

Arbeidsmarknadsstatus	A		B	
	Status=2	Status=3	Status=2	Status=3
Forklaringsvariable				
EDUC	-0.674 (0.070)	-0.315 (0.065)	-0.759 (0.061)	-0.509 (0.054)
EXPER	-0.106 (0.173)	0.849 (0.157)		
EXPER SQ	-0.013 (0.025)	-0.077 (0.023)		
BLACK	0.813 (0.303)	0.311 (0.282)	1.059 (0.293)	0.203 (0.272)
CONSTANT	10.28 (1.13)	5.54 (1.09)	10.94 (0.912)	9.694 (0.846)
LOG LIKELIHOOD	-907.85723		-1055.339	
Talet på observasjonar	1717		1717	

## Oppgåve 3.

Du ynskjer å studere samanhengen mellom studentar sin arbeidsinnsats, initialt kunnskapsnivå og ein del andre variablar. Du har gjennomført ei spørjeundersøking mellom 800 studentar i samfunnsøkonomi der studentane angir kor mange timar i veka dei i gjennomsnitt brukar på studiet. Du har registeropplysningar om kvar einskild student sin gjennomsnittskarakter frå vidaregåande skule som du brukar som mål på den einskilde student sitt initiale kunnskapsnivå. I tillegg har du opplysningar om alder, talet på semester i studiet samt registeropplysningar om foreldra sitt inntektsnivå for kvar einskild student. Anta at desse variablane er dei relevante variablane som forklarar studentane sin arbeidsinnsats.

Men det viser seg at 200 av dei 800 studentane ikkje svarar på spørjeskjemaet. Gjer greie for kva for utfordringar dette reiser for moglegheita til å estimere ein truverdig arbeidsinnsatsrelasjon for studentane og korleis du ville handtert problemet.