

Evner, fag, indsats i gymnasiet og beståelse på samfundsvidenskab

Karsten Albæk

Økonomisk Institut, Københavns Universitet, E-mail: Karsten.Albaek@econ.ku.dk

SUMMARY: The paper analyses the factors that affect the ability of students to pass the bachelor degree in social science at the University of Copenhagen. Ability in mathematics plays a role for passing in law and economics but not in political science, psychology, anthropology and sociology. Students with low grade point averages and low grades in mathematics have a larger probability to opt out of mathematics in high school. This affects their ability to pass studies with relatively easy access such as law and economics. The paper includes a discussion of the implications for the admission policy to the higher educational system.

1. Indledning

Hvad betyder elevernes valg af fag i gymnasiet for beståelse af et samfundsvidenskabeligt studium? Betyder det noget, om man kommer fra det almene gymnasium, HHX, HTX eller HF? Er det evnerne for matematik, der betyder en forbedring af mulighederne for at bestå, eller er det kurset i matematik, som er afgørende? Hvis man har en gennemsnitlig studentereksamen (et eksamensresultat svarende til gennemsnittet blandt alle studenter), er der så nogle studier, det er godt at søge ind på, og andre studier man skal holde sig fra?

I det følgende søges disse spørgsmål besvaret. Emnet er ikke alene relevant for de elever, som skal vælge fag og studieretninger i gymnasiet. Det er også relevant i den aktuelle uddannelsespolitiske debat.

Resultaterne i det følgende tyder på, at reduktionen af kvote 2 indebærer, at færre studerende består en videregående uddannelse. På fag med meget hård adgangsbe- grænsning er der ingen nævneværdig forskel mellem beståelsen blandt studerende med en gennemsnitlig studentereksamen og studerende med høje karakterer. For-

Revideret version af indlæg udarbejdet til Nationaløkonomisk Forenings årsmøde på Koldingfjord den 6.-7. januar 2006. Tak til Anna Piil Dam, Christian Hjorth-Andersen, Bo Honoré, Troels Østergaard Sørensen og to referees for kommentarer i forbindelse med udarbejdelsen af manuskriptet, til Lene Kjærsgaard for gennemførelse af beregningerne og til Anne Katrine Lefmann og Jesper Krogstrup for konstruktivt samarbejde i forbindelse med oparbejdelse af data. Undersøgelsen er gennemført med finansiel støtte fra Økonomisk Institut, Københavns Universitet, i forlængelse af instituttets målsætning om at bidrage til at forbedre de studerendes studieforløb.

mindsnelsen af kvote 2 indebærer, at der på disse studier optages flere studerende med meget høje gennemsnit, som kunne have klaret sig fint på et alternativt studium. Studerende med karakterer til studentereksamen svarende til den gennemsnitlige student har derimod svært ved at klare sig på studier med svag adgangsbegrænsning som jura og økonomi.

Fra sommeren 2005 er der gennemført en gymnasireform, som indebar betydelige ændringer i undervisningsforløbene i gymnasiet. Gymnasireformen synes imidlertid ikke at have været helt så gennemtænkt som ønskeligt. Der er allerede nu gennemført forskellige justeringer, og der er fremsat ønsker om yderligere ændringer i reformen. Et væsentligt element i forbindelse med gymnasiets udformning er de *valgmuligheder*, det besluttes at stille eleverne over for. Her må det være af relevans at vurdere, i hvilke omfang nogle fag eller kombinationer af fag forbedrer mulighederne for at bestå en videregående uddannelse.

Regeringen har netop gennemført en ny måde at regulere adgangen til de videregående uddannelser. Fremover skal der gives en belønning til eleverne, hvis de yder *en ekstra indsats i gymnasiet*. Den ekstra indsats består i, at de tager ét eller to ekstra fag på højt niveau. Belønningen består i, at eleverne får et tillæg til gennemsnittet ved studentereksamen ved beregningen af den adgangsgivende kvotient til de studier, hvor der er adgangsbegrænsning.¹ Valg af ekstra fag på højt niveau i gymnasiet synes imidlertid ikke at betyde noget for beståelse på samfundsvidenskab på Københavns Universitet. Reformen kan derfor ikke forventes at medføre nogen positiv effekt på elevernes muligheder for at bestå en videregående uddannelse. Tværtimod kan man forvente en negativ effekt, hvis eleverne fravælger svære og arbejdskrævende fag som matematik og fysik for at få tid til at følge mere end to valgfag på højt niveau. Evner for matematik har nemlig stor betydning for de studerendes mulighed for at bestå de to største samfundsvidenskabelige studier på Københavns Universitet, jura og økonomi.

Den empiriske undersøgelse omfatter de *samfundsvidenskabelige studier* på Københavns Universitet. Dette universitet fylder ganske meget rent kvantitativt, hvad angår optaget til de videregående uddannelser i Danmark. Resultaterne har derfor en selvstændig interesse, uanset om tilsvarende analyser for andre universiteter måtte give afvigende resultater.

De samfundsvidenskabelige studier på Københavns Universitet har en væsentligt betydning for uddannelsessystemet i Danmark. Årsagen er, at der er en *meget hård adgangsbegrænsning* til de fleste studier på samfundsvidenskab på Københavns Universitet. Hvis elever i gymnasiet ønsker at have muligheden for at komme ind på disse studier, vil det være naturligt at træffe valg i gymnasiet, så der er en stor chance for at

1. Adgangsbekendtgørelsens § 19, stk. 3. lyder: »Justeringen af eksamensgennemsnittet efter stk. 1 foretages ved at multiplicere det opnåede eksamensgennemsnit med 1,03 for gennemførelse af et ekstra niveau og med 1,06 for gennemførelse af to ekstra niveauer«, jfr. Undervisningsministeriet (2005).

komme ind. Det kan gå ud over mulighederne for at bestå studier, hvor der enten er frit optag, eller hvor der alene er indført adgangsbegrænsning for at sikre de optagne en rimelig chance for at bestå studiet. Det er f.eks. tilfældet på de fleste tekniske og naturvidenskabelige studier.

Dette indlæg ligger i forlængelse af tidligere undersøgelser af beståelse på politstudiet, Albæk (2001) og Albæk (2004a). Der er tale om en udvidelse på to felter i forhold til tidligere: For det første er der medtaget flere studier. For det andet omfatter data elevernes valg af fag i gymnasiet samt karakterer opnået i fagene. Oplysningerne fra gymnasiet gør det muligt at prøve at skelne mellem effekten af et kursus og elevernes evner. Et højt niveau i matematik har en positiv effekt på resultaterne i faget statistik på Handelshøjskolen i Århus ifølge Andersen og Østergaard (2005). Ifølge La Cour and Raimondos-Møller (2005) har høje karakterer i matematik en gavnlige effekt på resultaterne af udvalgte kurser på Handelshøjskolen i København, mens effekten af kurset har et mindre kraftigt gennemslag. Overvejelser af mere teoretisk karakter findes i Albæk (2003b), der præsenterer et optagelsessystem til de videregående uddannelse, som i modsætning til det nuværende optagelsessystem minimerer frafaldet på de videregående uddannelser. Tankegangen anvendes i en diskussion af forslaget til gymnasireform i et indlæg på Nationaløkonomisk Forenings forrige årsmøde, Albæk (2004a).

Indlægget er organiseret som følger. I afsnit 2 præsenteres data og der gives en oversigt over, hvor lang tid de studerende er om at gennemføre de forskellige samfundsvideenskabelige studier. Forskellige forklarende variables effekt på beståelsessandsynligheden på de forskellige studier præsenteres i afsnit 3, som også indeholder en nærmere beskrivelse af betydningen af gennemsnittet ved den adgangsgivende eksamen for sandsynligheden for at bestå. Afsnit 4 indeholder en analyse af elevernes valg af matematik i 3.g på det almene gymnasiums matematiske linje. Dette danner udgangspunkt for analysen i afsnit 5, hvor det forsøges at skelne mellem effekten af evner, indsats og kursus på beståelsen af et samfundsvideenskabeligt studium. Den afsluttende diskussion gennemføres i afsnit 6, hvor forskellige implikationer af resultaterne skitseres.

2. Data og tidsrum for beståelse på samfundsvideenskab

I det følgende præsenteres de data, der anvendes i analysen, og det undersøges, hvor lang tid det tager for en studerende at bestå bachelorgraden på samfundsvideenskab. Den efterfølgende analyse ser alene på determinanterne for beståelse af bachelorgraden inden for 4 år efter immatrikulationen. De studerende kan imidlertid bestå både før og efter disse 4 år. Det har interesse at vurdere, hvor mange det drejer sig om, og hvordan beståelsen afhænger af længden af tid på studiet.

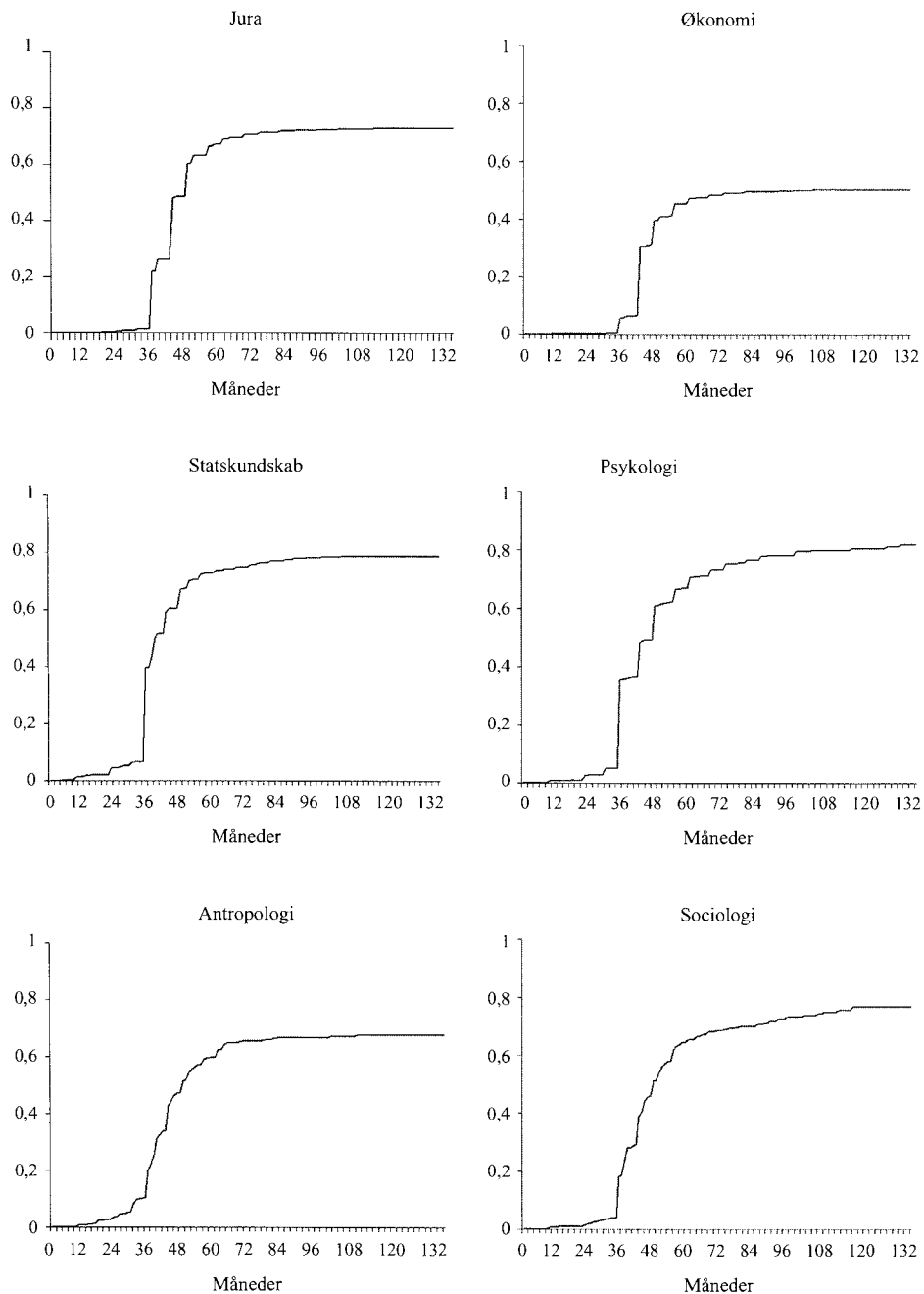
Data omfatter alle, der er immatrikulerede på samfundsvidenskab på Københavns Universitet fra sommeren 1994 og frem til sommeren 2005. Der medtages de studerende, der er indskrevet på universitetet på datoen for semesterstarten. For eleverne i det almene gymnasium suppleres data med valg af fag i gymnasiet og de karakterer, der opnået i de enkelte fag. Det har været muligt for ca. 90 procent af eleverne i det almene gymnasium. Bortfaldet er koncentreret hos de tidlige årgange, hvor der ikke er data fra gymnasiet for de studerende, der har udskudt studiestarten nogle år. For studerende med andre adgangsgivende eksaminer har det desværre ikke været muligt at opnå data om valg af fag og karakterer i fagene; disse oplysninger findes tilsyneladende ikke for HF, HHX og HTX. Data findes heller ikke for det almene gymnasium længere tilbage i tiden.

Det seneste tidspunkt for observation af de studerendes beståelse på Københavns Universitet er resultaterne for eksamen i sommeren 2005. Da succeskriteriet er beståelse af bachelorgraden inden for fire års studier, er den sidste årgang i data de studerende, der blev immatrikulerede sommeren 2001. I perioden sommeren 1994 til sommeren 2001 blev der i alt optaget 11.493 studerende på samfundsvidenskab på Københavns Universitet ifølge oversigten over data i tabel A1 i appendiks. Jura er langt det største studium med 43 procent af optaget i perioden, mens økonomi stod for 24 procent, og statskundskab for 13 procent.

Beståelsen af bachelorgraden afhængigt af tid på studiet fremgår af figur 1, som viser Kaplan-Meier overlevelsessandsynlighederne. Efter 12 års ophold i studiet har lidt over 70 procent af de immatrikulerede jurastuderende bestået bachelorgraden. Overgangssandsynligheden fra ikke-bestået til bestået på det 12. studieår kan alene foretages for den årgang, der blev immatrikuleret i 1994. De efterfølgende årgange er censurerede ved beregning af kurven i takt med, at de ikke bidrager til beregning af overgangssandsynlighederne. Grafen viser derfor resultatet af en beregning, hvor en jurastuderende består bachelorgraden i takt med de overgangssandsynligheder, der har været gældende for årgangene 1994-2001.

Efter fire års studier har 63 procent af de jurastuderende bestået bachelorgraden, som det fremgår af opgørelsen i tabel A1, hvilket også aflæses af figur 1. Hovedparten af dem, der består, er altså bestået inden for fire år, mens ca. 8 procentpoint yderligere kan forventes at bestå efter det fjerde år. Tidspunktet for beståelsen i data er tidspunktet for registrering af beståelse i Københavns Universitets administrative system. Studiestart er i september og en student, der består ved sommereksamen efter fire års studier, består derfor bachelorgraden på lidt mindre end 48 måneder.

På tre af studierne ligger beståelsen til slut noget højere end på jurastudiet: ca. 80 procent af de studerende forventes at bestå på statskundskab, psykologi og sociologi. På de to sidstnævnte studier er de studerende imidlertid noget langsommere til at bestå



Figur 1. Kaplan-Meiers overlevelsesestimater over antal måneder for færdiggørelse af bachelordelen, immatrikulation sommer 1994 – sommer 2001.

end på jurastudiet, idet 59 procent er bestået efter fire år på psykologistudiet, mens andelen på sociologi er nede på 52 procent. Hos antropologerne ligger beståelsen efter mange års studier på knap 70 procent, hvor 54 procent består inden for fire år. På disse studier er det altså mere end 20 procent af de studerende, som består bachelorgraden efter mere end fire års studier. Den laveste beståelse findes på økonomistudiet, hvor ca. halvdelen består, og hvor 40 procent er bestået inden for fire år.

På både statskundskab og psykologi består cirka 40 procent af de studerende bachelorgraden på de normerede tre år. På jura består ca. 25 procent til normeret tid, og på økonomi er det ca. 5 procent, der overholder det normerede studieforløb. På økonomi består den typiske studerende efter 3^{1/2} års studier, og en god del af resten efter fire års studier. Efter 2. årsprøve er bestået, får mange økonomistuderende studierelevant arbejde som studentermedhjælp i ministerier og organisationer, og det bidrager til en vis forsinkelse i studieforløbet. Også på jura er det almindeligt, at de studerende tager studierelevant arbejde, før bachelorgraden er færdiggjort.

Graferne illustrerer det gennemsnitlige forløb på de enkelte studier i observationsperioden. Den relativt lave beståelse på økonomistudiet kan bl.a. tilskrives, at der i en del af perioden har været fri adgang til at begynde studiet uanset eksamensgennemsnit.

3. Adgangsveje, eksamensgennemsnit og beståelse på samfundsvidenskab

Hvordan afhænger chancen for at bestå af typen af adgangsgivende eksamen? Og hvordan varierer beståelsessandsynligheden med de studerendes kvalifikationer som målt ved karaktergennemsnittet ved den adgangsgivende eksamen? I dette afsnit ses der først på, hvordan beståelsen varierer med forskellige karakteristika. Ved hjælp af en grafisk fremstilling ses der herefter nøjere på, hvordan sandsynligheden for at bestå varierer med karaktergennemsnittet.

Det anvendte succeskriterium er som nævnt beståelse af bachelorgraden inden for fire år.² Bachelorgraden er normeret til at tage tre år, men der kan være flere gode grunde til, at nogle studerende bliver lidt forsinket, f.eks. studierelevant erhvervsarbejde. En rimelig effektivitet tilsiger imidlertid, at hovedparten af de studerende, der består bachelorgraden, bør gennemføre inden for fire år. Denne opfattelse deles af regeringen, som har under overvejelse at fjerne taxametertilskuddet for de studerende, der består efter det fjerde år.³

2. Et alternativt succeskriterium kunne være beståelse af første årsprøve inden for ét år som i Albæk (2001). Det er imidlertid uklart, i hvilket omfang første årsprøve hos alle studier omfatter en egentlig prøvning af de studerende, men det må nødvendigvis være tilfældet ved bachelorgraden.

3. Resultaterne er følsomme over for valg af det antal år, der vælges som succeskriterium. På psykologi og sociologi er næsten alle studerende bestået efter ti års studier. Hvis alle studerende består, er der ingen signifikante koefficienter. Det kan imidlertid ikke være nogen succes, at man består efter ti års studier.

Tabel 1. Beståelse af bachelordelen indenfor 4 års studier blandt studerende på samfundsvidenskab, immatrikulation sommer 1994 – sommer 2001. Marginale effekter af probitregression.

Forklarende faktorer:	Jura	Økonomi	Statskundskab	Psykologi	Antropologi	Sociologi
<i>Adgangsgivende eksamen:</i>						
Matematisk student	0,113** (0,017)	0,122* (0,048)	-0,015 (0,030)	0,079* (0,040)	-0,022 (0,059)	0,050 (0,057)
Studenterkursus	-0,291** (0,057)	-0,157 (0,086)	-0,443** (0,087)	-0,126 (0,131)	-0,165 (0,162)	0,347** (0,128)
HF	-0,215** (0,024)	-0,180** (0,051)	-0,353** (0,044)	-0,194** (0,038)	-0,153** (0,056)	-0,122 (0,064)
HHX	-0,010 (0,025)	-0,008 (0,057)	-0,144* (0,063)	0,052 (0,082)	0,131 (0,154)	-0,126 (0,120)
HTX ^(e)	-0,016 (0,140)	-0,020 (0,087)	-0,479** (0,146)	-0,112 (0,358)	-	-
Udland	-0,159** (0,061)	-0,150 (0,077)	-0,483** (0,084)	-0,147* (0,075)	-0,290** (0,106)	-0,079 (0,268)
Andre	-0,115 (0,067)	-0,002 (0,114)	0,046 (0,073)	0,080 (0,113)	-0,126 (0,113)	0,232 (0,233)
<i>Kvotient:</i>	0,073** (0,012)	0,241** (0,013)	0,045* (0,018)	-0,017 (0,022)	0,044 (0,034)	0,018 (0,031)
<i>Eksamensalder:</i>	0,063** (0,020)	0,050 (0,027)	0,083* (0,038)	0,039 (0,049)	0,011 (0,086)	0,061 (0,084)
1 år	0,073** (0,023)	0,083** (0,031)	0,040 (0,043)	0,083 (0,052)	0,094 (0,084)	0,087 (0,086)
2 år	0,033 (0,028)	0,109** (0,041)	0,079 (0,045)	0,121* (0,056)	0,284** (0,073)	0,034 (0,097)
3 år	0,107** (0,031)	0,080 (0,059)	0,085 (0,048)	0,069 (0,059)	0,239** (0,090)	0,012 (0,100)
4 år	0,049 (0,041)	0,137 (0,076)	-0,035 (0,064)	0,060 (0,072)	0,186* (0,092)	0,209* (0,088)
5 år						

fortsættes...

fortsat...

Forklarende faktorer:	Jura	Økonomi	Statskundskab	Psykologi	Antropologi	Sociologi
Mindst 6 år	-0,117** (0,028)	-0,195** (0,046)	-0,121* (0,060)	-0,113* (0,052)	0,241** (0,090)	-0,046 (0,091)
Kvinde	0,095** (0,015)	0,033 (0,023)	0,054* (0,025)	0,029 (0,034)	0,028 (0,055)	0,086 (0,050)
Pseudo R ²	0,068	0,165	0,099	0,058	0,052	0,043
Antal observationer	4,919	2,736	1,531	1,231	538	534
Beståelseandsynlighed:	0,589	0,400	0,677	0,559	0,539	0,522
Referencestudent	0,512	0,359	0,674	0,593	0,400	0,423

Anm.: Robuste standardfejl i parenteser, (*)5 procentns signifikansniveau, (**)1 procentns signifikansniveau. Referencestudenten er en mandlig student med et gennemsnit på 9, som begyndte på studiet lige efter en sproglig studentereksamen på gymnasiet.

Note: (a) Koefficienter for antropologi og sociologi er ikke rapporteret, da variabelen forudsiger beståelse perfekt; se forklaring i tekst.

Effekten af forskellige karakteristika på sandsynligheden for at bestå findes ved probitregression, og præsentationen af resultaterne sker i form af marginale effekter af de forskellige forklarende variable. En matematisk student har 11,3 procentpoint større sandsynlighed for at bestå bachelorgraden på jura end en sproglig student, jævnfør første søjle i tabel 1. Referencestudenten er en mandlig student, har et gennemsnit på 9 ved sin sproglige studentereksamen og er begyndt på studiet umiddelbart efter sin studentereksamen. Nederst i første søjle fremgår det, at en sådan student har 51,2 procent sandsynlighed for at bestå. Matematikeren med samme karakteristika har derfor 62,5 procent sandsynlighed for at bestå.

Tilsvarende har en matematisk student cirka 12 procent større sandsynlighed for at bestå økonomistudiet sammenlignet med en sproglig student, der læser økonomi. For begge studiers vedkommende er der tale om betydelige og signifikante effekter. Man kunne undre sig lidt over, at matematikere skulle have en fordel frem for sproglige på jurastudiet, som slet ikke bruger formler i studiet. Da jurastudiet imidlertid trækker kraftigt på evnen til at tænke logisk, er én mulig fortolkning, at en matematisk studentereksamen har bidraget til at udvikle denne evne. På psykologistudiet er det også en fordel at være matematisk student, det giver 8 procentpoint større sandsynlighed for at bestå sammenholdt med at være sproglig student, og koefficienten er netop signifikant på et 5 procentniveau. For de tre øvrige samfundsvidenskabelige studier er der ingen signifikant forskel mellem beståelsen for matematiske og sproglige studenter.

Studerende med studenterkursus og HF har sværere ved at bestå samfundsvidenskab end sproglige studenter fra det almene gymnasium. Fraset sociologi er alle koefficienterne negative, og flertallet signifikant forskellige fra nul. På statskundskab er det en ulempe at komme fra HHX og HTX sammenlignet med at have taget en sproglig studentereksamen. For de øvrige studier er der ingen forskel i beståelsen.⁴

Det er en ulempe at forsøge at tage en samfundsvidenskabelig uddannelse med en adgangseksamen fra udlandet. Koefficienterne er alle negative, og flertallet er signifikant forskellige fra nul. En del af forskellen kan sandsynligvis tilskrives, at der er betinget på karaktergennemsnittet, som er konverteret fra karaktergennemsnittet opnået ved udenlandske uddannelsesinstitutioner. Indtrykket er, at disse udenlandske gennemsnit ofte overføres til et uforholdsmæssigt højt niveau på den danske karakterskala.

4. Koefficienterne for HTX er ikke rapporterede for antropologi og sociologi. På begge studier er der to studerende i denne kategori, og på antropologi er ingen af dem bestået, mens det modsatte er tilfældet på sociologi. Variablen HTX forklarer derfor venstresiden i regressionen perfekt på de to studier. Maximum likelihood estimatoren itererer sig derfor frem til en så lav koefficient for variablen HTX hos antropologerne, at det er næsten sikkert, at ingen med HTX består. Analogt får sociologer med HTX som baggrund tillagt en værdi, der tilsiger, at alle består. Koefficienterne giver ingen faglig mening, og standardprogrammer plejer derfor at droppe observationer med denne datakonstellation. I arbejdsrapporten Albæk (2006) anvendes en lineær sandsynlighedsmodel, og her er de rapporterede koefficienter selvsagt insignifikante. Bortset fra de koefficienter i probitregressionen, som det ikke giver mening at rapportere, er der ingen nævneværdig forskel mellem koefficienterne i arbejdsrapporten og koefficienterne i denne artikel.

Regressionen indeholder også alder for adgangsgivende eksamen («eksamensalder»), og her er koefficienterne alle positive til og med fem år. Referencestudenten begyndte umiddelbart efter afslutningen af den adgangsgivende eksamen, og en sådan student har derfor i gennemsnit en lavere sandsynlighed for at bestå end de studenter, der udsætter studiestarten. For de to studier med mange observationer, jura og økonomi, er der i flere tilfælde tale om signifikante effekter. Lang tids udsættelse af studiestarten har imidlertid en signifikant negativ effekt på sandsynligheden for at bestå på fire af studierne, som det ses af dummyen for studiestart på mindst seks år efter adgangsgivende eksamen. At studerende med studiestart umiddelbart efter den adgangsgivende eksamen har en lavere sandsynlighed for at bestå end studerende med en højere eksamensalder er velkendt fra litteraturen og diskuteret grundigt i Albæk (2001), hvortil der henvises.

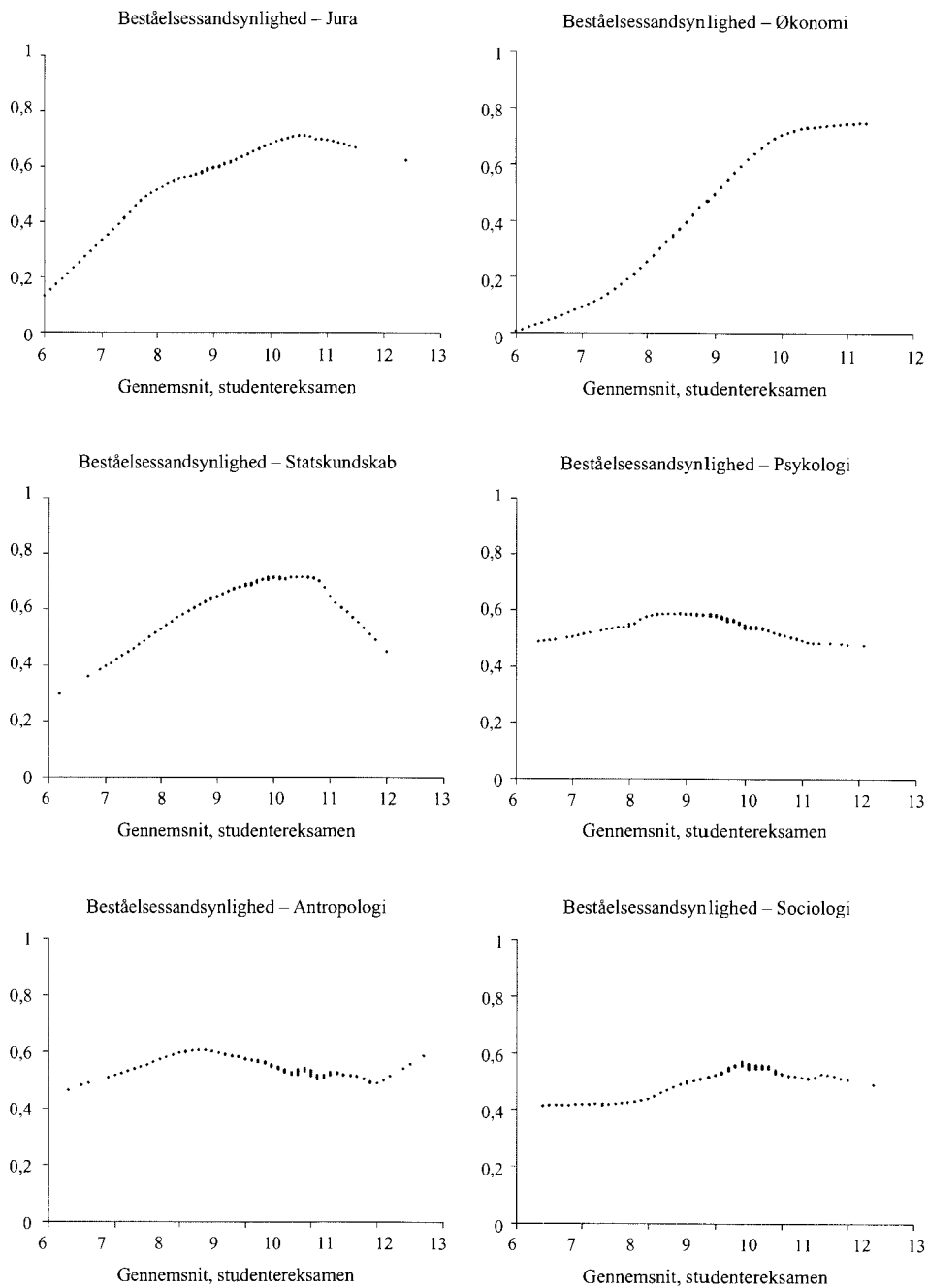
Der er en tendens til, at kvinder klarer sig bedre end mandlige studerende på samfundsvidenskab. Alle koefficienterne til dummyen for kvinde er positive, og to er signifikant forskellige fra nul.

Højere karakterer ved den adgangsgivende eksamen har en signifikant positiv effekt på sandsynligheden for at bestå på jura, økonomi og statskundskab. På de øvrige studier er der ingen signifikant effekt. Variablen »kvotient« er gennemsnittet ved den adgangsgivende eksamen fratrukket 9. I tabel 1 medtages derfor alene den lineære effekt af denne variabel. Karaktergennemsnittet har en betydelig interesse både blandt eleverne i gymnasiet og i forbindelse med diskussionen om tilrettelæggelse af adgangsgivende eksamen. For at få et bedre overblik over sammenhængen mellem beståelse og karaktergennemsnit suppleres med en grafisk fremstilling.

Af figur 2 fremgår, at sandsynligheden for at bestå jurastudiet er pænt, lineært stigende med karakterniveauet. En student med 10 i gennemsnit har cirka 70 procents sandsynlighed for at bestå bachelorgraden inden for 4 år, mens studenter med 9 i gennemsnit har ca. 60 procents chance, og dette falder til ca. 50 procent og 35 procent for studenter med henholdsvis 8 og 7 i gennemsnit. Figuren er konstrueret ved at danne et lokalt gennemsnit og herefter indsætte de predikterede værdier for de enkelte studerende.⁵

På økonomi er resultatet det samme for høje gennemsnit omkring 10, cirka 70 procents sandsynlighed for at bestå. Her er der imidlertid tale om en betydeligt kraftigere vækst i beståelsessandsynligheden med gennemsnittet. En studerende med 9 i gennemsnit har således ca. 50 procents sandsynlighed for at bestå, faldende til ca. 25 pro-

5. Nærmere bestemt er der gennemført en lokal regression af indikatorvariablen for beståelse på karaktergennemsnittet ved den adgangsgivende eksamen. Regressionen gennemføres som en lowess smoother med en lav grad af udglatning, en båndvidde på 0,8. Herefter er dette lokale gennemsnit afbildet i grafen i form af den forudsagte værdi for hver enkelt studerende.



Figur 2. Gennemsnit ved adgangsgivende eksamen og sandsynlighed for at bestå bachelordelen inden for fire år, immatrikulation sommer 1994 – sommer 2001.

hvor g er tæthedsfunktionen. Jo højere forventet matematikkarakter, jo større er sandsynligheden for, at matematik vælges.

Tre centrale antagelser driver beviset. For det første, at eleven foretager valget ved at sammenligne gevinsten eller nytten ved at følge matematik med ét eller andet alternativ. For det andet, at den forventede karakter indgår i gevinsten eller nytten ved valgmulighederne. For det tredje, at der er en spredning i elevernes præferencer for matematik og alternativet. Hvis disse antagelser er opfyldt, vil elever, som forventer højere karakterer i matematik end i andre fag, altså være mere tilbøjelige til at vælge matematik.

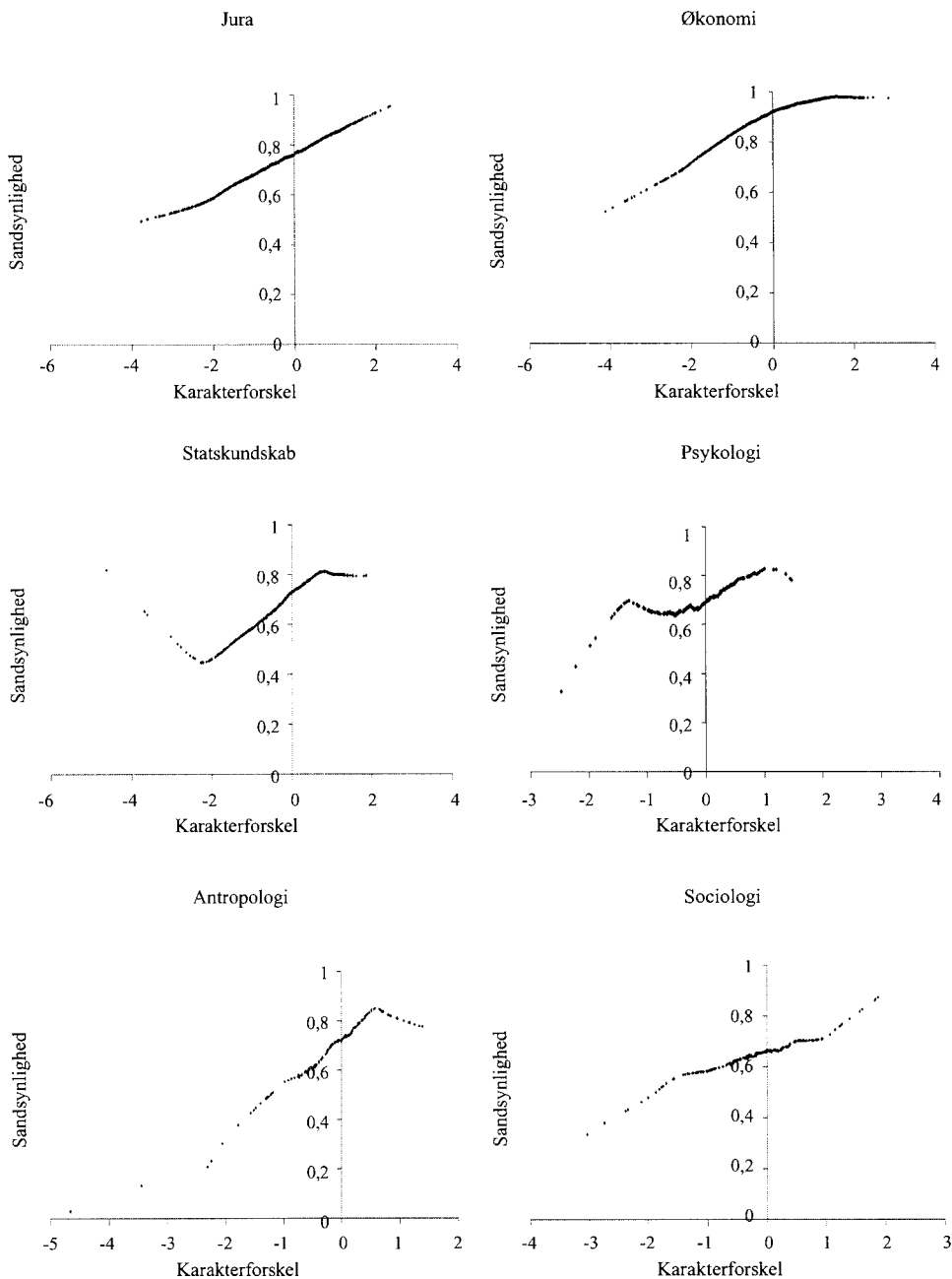
Passer denne forudsigtelse med elevernes faktiske valg af fag i det danske gymnasium? Det kan man udtale sig om på baggrund af data, som omfatter karakteren i matematik efter 2.g, både for dem, der har valgt matematik i 3.g, og for dem, der har fravalgt matematik. Karakteren i matematik efter 2.g kan opfattes som en indikator på karakteren efter 3.g, hvis kurset i matematik vælges.

Figur 3 viser, at der er en klar sammenhæng mellem sandsynligheden for at vælge matematik i 3.g, og den karakter, der er opnået i matematik i 2.g. Variablen på førsteaksen er matematikkarakteren efter 2.g fratrukket gennemsnittet i de øvrige fag i 2.g. Jo højere matematikkarakter sammenlignet med de øvrige fag, jo større er sandsynligheden for valg af matematik. Denne sammenhæng gælder blandt studerende for alle samfundsvidenskabelige studier (undtagelsen er den nedre del i grafen for studerende på statskundskab, men her er observationerne sparsomme). Elevernes valg af matematik stemmer altså overens med resultatet af de teoretiske overvejelser ovenfor.

Højden på 2. aksens i figur 3 ved en karakterforskel på nul angiver sandsynligheden for at vælge matematik, hvis eleven har 9 i gennemsnit i 2.g. Der er en positiv sammenhæng mellem karakterniveau og valg af matematik, men denne variation er fjernet i figur 3 (den afhængige variabel er summen af konstantleddet og residualerne i en lineær regression af valg af matematik på en konstant og gennemsnittet i 2.g fratrukket 9). Det ses altså, at en økonomistuderende med 9 i gennemsnit i 2.g har ca. 90 procent sandsynlighed for at vælge matematik. Studerende på andre studier har en lavere sandsynlighed for at vælge matematik, ca. 75-80 procent.

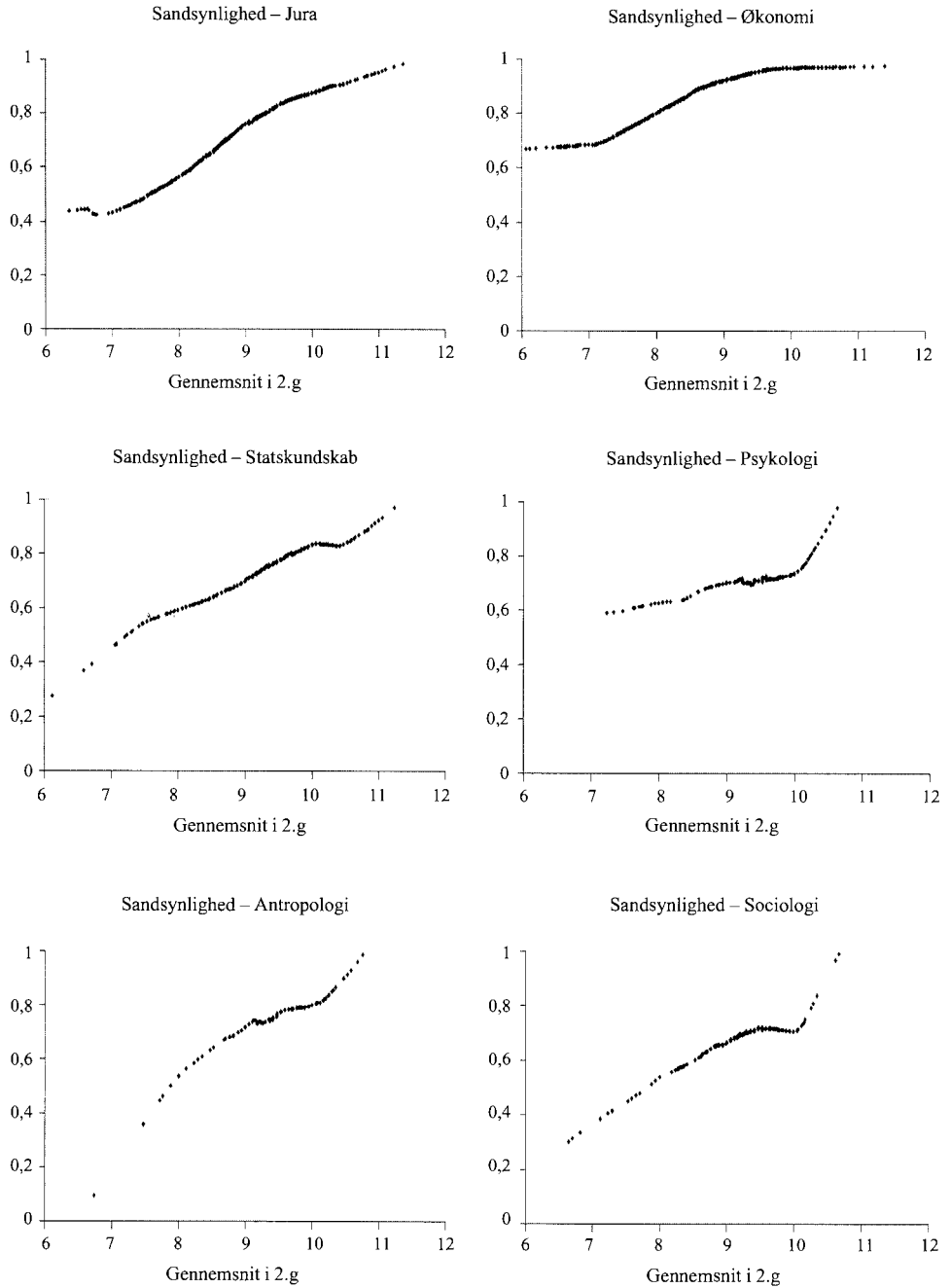
Som supplement til figur 3 er der medtaget en fremstilling af sammenhængen mellem gennemsnitligt karakterniveau i 2.g og valg af matematik. Af figur 4 fremgår, at der er en klar sammenhæng mellem gennemsnit og valg af matematik: jo højere karaktergennemsnit, jo større er sandsynligheden for at vælge matematik i 3.g. Der er imidlertid betydelige forskelle mellem de forskellige studier.

Blandt økonomer med et gennemsnit på 10 har næsten 100 procent valgt matematik i 3.g, mens dette gælder for ca. 70 procent ved lave gennemsnit. For studerende på andre studier er der tale om en stejle sammenhæng: fravalgt ved lave karakterniveauer



Figur 3. Valg af matematik på A-niveau for matematiske studenter sammenholdt med forskellen mellem matematikkarakterer og andre karakterer i 2.g.

Note: Den afhængige variabel er beregnet som konstanten tillagt residualerne fra en lineær regression af matematik på A-niveau på karaktergennemsnittet i 2.g minus ni.



Figur 4. Karaktergennemsnit i 2.g og matematiske studenter sandsynlighed for at vælge matematik på A-niveau.

er betydeligt større. Typisk har kun cirka halvdelen af de studerende valgt matematik ved et gennemsnit på 8. Det er en bemærkelsesværdigt lav andel: i gennemsnit er det ca. 75 procent af alle matematiske studenter, der vælger matematik i 3.g.

Forskellen mellem økonomers valg og valget blandt studerende på andre samfundsvidenskabelige studier er logisk. Hvis en elev allerede i 2.g har bestemt sig til at læse økonomi, kan han eller hun i gymnasiet koncentrere sig om at gøre en indsats for at lære noget, herunder vælge fag der giver en god forberedelse til det fremtidige studium. Kun et fåtal af eleverne på de andre samfundsvidenskabelige studier er i den gunstige situation. De er sat i en ekstrem konkurrencesituation, og en rationel elev må nødvendigvis gøre sig overvejelser om det forventede karakterniveau i valgfagene.

Fravalget af matematik har ingenting at gøre med, om der er behov for faget i det efterfølgende studieforløb, graferne illustrerer netop valget blandt studerende på de samme studier. Når elever med lave karakterer fravælger faget, er det et resultat af den incitamentsstruktur, de er stillet overfor i gymnasiet; i princippet kunne man forestille sig, at det var elever med lave gennemsnit og med lave karakterer i matematik, der havde mest behov for at følge faget.

Det er endvidere forsøgt at teste, om valg af mere end de to obligatoriske fag på A-niveau indebærer en tendens til fravalg af matematik. Umiddelbart synes det ikke at være tilfældet. For alle studerende på samfundsvidenskab er der gennemført en regression med valg af matematik på venstresiden, mens variablene på højresiden er karaktergennemsnit som afbildet i figur 2, forskellen mellem matematikkarakteren og andre karakterer som afbildet i figur 3, dummyer for valg af studium samt forskellige interaktionsled. Inkluderet i regressionen er en dummy for valg af tre eller flere fag på A-niveau, og koefficienten på 0,17 til denne variabel er signifikant forskellig fra nul. Hvis valget af det tredje fag sker tilfældigt, må der imidlertid forventes en positiv koefficient til denne variabel (og tilsvarende for alle andre fag, der måtte stå på venstresiden i en sådan regression). I det omfang, der er tilbøjelighed til at fravælge matematik i forbindelse med valg af tre fag på højt niveau, er fænomenet altså ikke så kraftigt, at koefficienten i en sådan regression kommer ned i nærheden af nul. Den ny belønning ved at vælge flere fag på højt niveau må forventes at få flere til at træffe dette valg, og om resultatet er gyldigt for denne gruppe af studerende er et åbent spørgsmål; det er bestemt ikke givet, at resultatet har ekstern validitet for denne nye gruppe af studenter.

5. Evner, valg og beståelse

Er det evnerne for matematik, eller er det selve kurset, som betyder noget for beståelsen på samfundsvidenskab? Det vil blive forsøgt belyst i det følgende.

Analysen tager udgangspunkt i, at det er muligt at observere karaktererne i matematik i 2.g både for dem, der vælger matematik i 3.g, og dem, der fravælger matematik.

Karaktererne i 2.g i matematik vil blive opfattet som et mål for evnerne til matematik. Det antages altså, at højere karakter i matematik afspejler bedre evner (eller en blanding mellem gode evner og høj indsats). Med denne antagelse bliver det muligt at foretage en sondring mellem evner og effekten af tredje års matematik i det almene gymnasium. Det er altså alene tredje års matematikken, der analyseres. I de to første år i det matematiske gymnasium har eleverne haft obligatorisk matematik, og effekten af dette tages for givet.

Beståelsen på de samfundsvidenskabelige studier antages at kunne beskrives ved følgende model

$$y_i = \beta m_i + \gamma A_i + \delta x_i + e_i.$$

Responsvariablen y_i antager værdien én, hvis student i er bestået og nul, hvis student i ikke er bestået. m_i er en indikator for, om matematik følges i 3.g eller ej, A_i er mål for evner, x_i er andre forklarende variable inklusiv et konstantled, og e_i er fejlløddet.

Den forventede beståelse, betinget af de forklarende variable, bliver

$$E(y_i | m_i, A_i, x_i) = \beta m_i + \gamma A_i + \delta x_i + E(e_i | m_i, A_i, x_i).$$

Hvis målet for evner, A_i , omfatter de forhold, som betinger valget af kurset m_i , kan der opnås et estimat for effekten af kurset, som ikke er behæftet med bias. I modsat fald må sidste led i udtrykket forventes at være forskelligt fra nul og koefficienten til kurset forskelligt fra β , se f.eks. Hayashi (2000), kap. 3.

I regressionerne i tabel 1 indgik gennemsnittet ved den adgangsgivende eksamen, som er et mål for evner generelt. Af forrige afsnit fremgik det imidlertid, at evnerne til matematik, relativt til evnerne for andre fag, også havde betydning for valget af matematik i 3.g. Dette mål inkluderes derfor i analysen, der alene omfatter matematiske studenter, hvor der er data for evner til matematik før og efter valget af matematik-kurset i 3.g.

I appendiks tabel A2 beskrives de data, der indgår i analysen. Studenterne er opdelt efter hvilke kombinationer af fag, de har valgt. Den mest populære kombination af fag er matematik og samfundsfag på A-niveau, som 31 procent af de samfundsvidenskabelige studenter har valgt, efterfulgt af matematik og engelsk på A-niveau med en andel på 23 procent. Matematik og fysik er valgt af 11 procent, matematik på A-niveau og andre fag end samfundsfag, engelsk eller fysik på A-niveau er valgt af 21 procent. Endelig har 19 procent fravalgt matematik i 3.g og har derfor valgt andre fag i forskel-

Tablet 2. Beståelse af bachelordelen inden for fire års studie blandt matematiske studenter på samfundsvidenskaber, immatrikulation sommer 1994 – sommer 2001. Marginale effekter af probitregression.

Forklarende faktorer:	Jura	Økonomi	Statskundskab	Psykologi	Antropologi	Sociologi
Fag:						
Matematik og fysik på A-niveau	-0,016 (0,051)	-0,014 (0,045)	-0,155 (0,087)	0,234** (0,049)	0,026 (0,249)	0,029 (0,285)
Matematik og engelsk på A-niveau	-0,073* (0,037)	-0,071 (0,038)	-0,118 (0,067)	0,190* (0,079)	0,022 (0,167)	-0,098 (0,146)
Matematik på B-niveau	-0,058 (0,046)	-0,148* (0,064)	-0,004 (0,080)	0,326** (0,102)	-0,057 (0,388)	0,003 (0,168)
Matematik på A-niveau og andre fag på A-niveau	-0,038 (0,038)	-0,080* (0,040)	-0,165* (0,070)	0,190* (0,083)	0,223 (0,135)	-0,089 (0,116)
Tre højniveaufag	0,060 (0,046)	-0,055 (0,052)	0,026 (0,070)	0,087 (0,111)	0,191 (0,154)	0,156 (0,184)
Kvotient	0,042 (0,027)	0,212** (0,021)	0,071 (0,042)	-0,134 (0,075)	-0,016 (0,113)	-0,050 (0,088)
Kvotient	0,061** (0,018)	0,132** (0,020)	-0,029 (0,032)	0,032 (0,061)	-0,037 (0,116)	-0,092 (0,084)
Matematikkaraktergennemsnit i 2.g minus kvotient	0,048* (0,019)	0,128** (0,019)	0,033 (0,034)	0,054 (0,053)	-0,045 (0,114)	0,066 (0,071)
Matematikkaraktergennemsnit 3.g minus matematikkaraktergennemsnit i 2.g	0,084 (0,050)	-0,029 (0,067)	-0,020 (0,061)	0,074 (0,132)	0,080 (0,289)	0,068 (0,152)
Matematik på B-niveau x matematikkaraktergennemsnit i 2.g	-0,061 (0,034)	-0,028 (0,051)	-0,027 (0,068)	0,244 (0,163)	-0,163 (0,196)	0,093 (0,151)
Eksamensalder:						
1 år	0,075* (0,030)	0,039 (0,037)	0,104 (0,055)	0,166 (0,097)	0,281 (0,222)	0,112 (0,151)
2 år	0,109** (0,032)	0,107** (0,041)	0,108 (0,056)	0,080 (0,110)	0,438* (0,181)	-0,009 (0,159)
3 år	0,072 (0,041)	0,156** (0,060)	0,077 (0,068)	-0,026 (0,161)	0,474** (0,109)	-0,204 (0,186)

fortsættes...

fortsat...

Forklarende faktorer:	Jura	Økonomi	Statskundskab	Psykologi	Antropologi	Sociologi
4 år	0,167** (0,043)	0,020 (0,097)	0,136* (0,062)	0,045 (0,131)	0,351* (0,171)	0,076 (0,177)
5 år	0,118 (0,064)	-0,032 (0,161)	-0,142 (0,121)	0,084 (0,141)	0,322 (0,186)	0,161 (0,158)
Mindst 6 år ^(a)	0,115 (0,085)	- (0,156)	-0,220 (0,156)	0,082 (0,136)	0,329 (0,175)	-0,004 (0,232)
Køn:	0,056* (0,025)	0,031 (0,032)	0,005 (0,042)	-0,027 (0,073)	-0,047 (0,133)	0,119 (0,087)
Pseudo R ²	0,030	0,156	0,081	0,122	0,107	0,086
Antal observationer	1,385	1,480	511	162	87	159
Beståelses-						
sandsynlighed:	I sample	0,737	0,487	0,747	0,598	0,642
	Referencestudent	0,686	0,566	0,517	0,133	0,595

Anm.: Robuste standardfejli i parenteser, (*) 5 procents signifikansniveau, (**) 1 procents signifikansniveau. Referencestudenten er en mandlig student, som har to valgfag på A-niveau, matematik og samfundsbyg, har 9 i gennemsnit samt 9 i matematik i 2. og 3. g. og som begyndte på studiet lige efter studentereksamen.

Note: (a) Koefficient for økonomi er ikke rapporteret, da variabelen forudsiger beståelse perfekt (ingen er bestået).

lige kombinationer på A-niveau i 3.g. Endvidere ses, at 8 procent af de studerende på samfundsvidenskab har valgt mere end to fag på A-niveau.⁷

Referencestudenten i tabel 2 har valgt matematik på A-niveau og samfundsfag på A-niveau i 3.g. Af første søjle fremgår, at der ikke er nogen signifikant forskel mellem beståelsen for en sådan fagkombination og kombinationen matematik og fysik på A-niveau, når det drejer sig om at bestå *jurastudiet*. Derimod er det en ulempe at have valgt kombinationen matematik og engelsk på A-niveau. Valg af matematik på B-niveau (og hermed to andre fag på A-niveau) giver en negativ, men ikke signifikant, forskel i beståelsen sammenlignet med en student, der har valgt matematik og samfundsfag på A-niveau. Endelig er der heller ikke nogen signifikant forskel mellem valg af andre kombinationer af fag end de allerede nævnte og valg af matematik og samfundsfag på A-niveau.

Der er endvidere medtaget en dummy for, om studenterne har valgt tre fag på A-niveau. Det fremgår, at der ikke er nogen signifikant forskel på beståelsen mellem studerende, der har valgt et ekstra fag på A-niveau, og de elever, der nøjes med det obligatoriske antal fag på A-niveau.

Kvotienten ved den adgangsgivende eksamen har en positiv, men ikke signifikant, koefficient. Så er der medtaget en variabel, som er konstrueret ved at tage matematikkarakteren i 2.g og fratække gennemsnittet til studentereksamen. Denne variabel måler, i hvilket omfang studenten er bedre til matematik end til de øvrige fag i gymnasiet. Der er tale om en signifikant effekt, og gode evner til matematik fremmer altså beståelsen på jurastudiet. Variablen svarer til den variabel, som indgik på førsteaksen i figur 4 i forrige afsnit, men her blev gennemsnittet af de øvrige karakterer i 2.g trukket fra. Der er ingen nævneværdige forskelle mellem, om den ene eller den anden variabel anvendes. Når kvotienten til studentereksamen anvendes i tabel 2 er det af fortolkningsmæssige årsager.

Endvidere indgår en variabel, som angiver forskellen mellem matematikkaraktererne i 3.g og matematikkaraktererne i 2.g. Denne variabel antager værdien nul for de elever, der har fravalgt matematik. Det fremgår, at variabelen har en signifikant positiv effekt på beståelsen på jurastudiet. De elever, som forbedrer præstationen i matematik fra 2.g til 3.g, har altså en signifikant højere sandsynlighed for at bestå. Endelig er der medtaget to

7. I tabel A2 er der endvidere medtaget en opgørelse over data for sproglige studenter. Ifølge opgørelsen er det de færreste sproglige elever, der i gymnasiet har valgt enten ét eller to års matematik. De fleste nøjes med at følge det obligatoriske kursus i naturfag. Suppleret i form af et efterfølgende kursus i matematik er imidlertid ikke indeholdt i data. På for eksempel økonomi skal alle immatrikulerede have matematik på mindst B-niveau, og ca. halvdelen af de sproglige studenter på økonomistudiet må derfor efterfølgende have fulgt et kursus i matematik, uden at det er registreret i data. Tilsvarende kan gøre sig gældende for de øvrige studier, og data kan derfor ikke anvendes til en nøjere analyse af konsekvenserne af de sproglige studenteres forudsætninger i matematik.

interaktionsled med det sigte at undersøge, om effekten af matematikkurset varierer over de studerendes evner. Ingen af interaktionsleddene er imidlertid signifikant forskellige fra nul.

På økonomistudiet er der ingen signifikant forskel mellem at vælge fysik eller samfundsfag som supplement til matematik. Den negative effekt af valg af engelsk er på kanten til signifikans. Samlebetegnelsen »andre fag« som supplement til matematik har en signifikant negativ effekt på beståelsen. Fravalg af matematik i 3.g giver en signifikant mindre beståelse af økonomistudiet. Evnerne til matematik betyder også noget for at bestå økonomistudiet. Høje matematikkarakterer i 2.g relativt til de øvrige karakterer giver signifikant større sandsynlighed for at bestå. En forbedring af præstationen i matematik fra 2.g til 3.g giver også større sandsynlighed for at bestå. Ingen af interaktionsleddene er signifikante, og det kan således ikke afvises, at effekten af matematikkurset er den samme uanset evner, både generelt og for matematik. Der er ikke nogen signifikant effekt af at tage ekstra fag på højt niveau.

Det er værd at notere, at effekten af gode evner til matematik ikke alene har signifikant betydning for beståelse af jura- og økonomistudierne, men at størrelsesordenen af effekten er betydelig. Hvis den studerende har ét karakterpoint højere karakterer i matematik end gennemsnittet til studentereksamen, har studenten altså 13 procentpoint højere sandsynlighed for at bestå på økonomi og 6 procentpoint højere sandsynlighed for at bestå jura.

For *statskundskabs* vedkommende er der et begrænset antal signifikante effekter. Hvad angår valget af fag, er den eneste signifikante effekt, at det tilsyneladende synes at være en dårlig ide at vælge andre fag end fysik, samfundsfag og engelsk sammen med matematik. Antallet af observationer på statskundskab er betydeligt mindre end for jura og økonomi, hvilket sandsynligvis er én af årsagerne til det begrænsede antal signifikante effekter.⁸

På *psykologi* giver kombinationen matematik og samfundsfag tilsyneladende en lav sandsynlighed for at bestå, idet de andre kombinationer af fag har positive og signifikante koefficienter. På *antropologi* og *sociologi* er der endnu færre observationer og ingen signifikante effekter.

Et hovedresultat af analysen er, at valg af tre fag på A-niveau i ingen tilfælde giver signifikant større sandsynlighed for at bestå et samfundsvidenskabeligt studium. Der er tjekket for, om resultatet er robust over for ændringer i specifikationen. Forskellige kombinationer af fag og kvotienter er således forsøgt. Endvidere er konstruktionen af

8. Sandsynligheden for at bestå afhængigt af forudsætninger afhænger selvsagt af studiets indretning. Det kan nævnes, at der tilsyneladende opnås andre resultater for statskundskabsstudiet på Århus Universitet. Det hedder således hos Kristensen (1998, s. 5): »Hvor blot ca. 8% af de studerende med et A-niveau i matematik er faldet fra efter 3 år, er ca. 34% af de studerende med B-niveau og ca. 48% af de studerende med C-niveau faldet fra efter samme periode«.

dummyerne for fag justeret, så fagopdelingen svarer til en klassedeling af eleverne (i tabel 2 medtages en elev, som f.eks. har valgt matematik, fysik og engelsk, både under matematik og fysik, matematik og engelsk og i dummyen for tre højniveaufag). I ingen af tilfældene opnås der signifikante effekter for valg af ekstra fag på højt niveau.

Et andet hovedresultat er, at matematik har signifikant betydning for beståelsen på de to største samfundsvidenskabelige studier, jura og økonomi. For begge studier gælder, at høje karakterer i 2.g i matematik har en selvstændig betydning for beståelsen (udover at bidrage til et højere gennemsnit). Det er værd at fremhæve, at målene for evner i matematik er karaktererne i faget efter to års matematik på gymnasialt niveau. Målet for evner i matematik er altså betinget af, at undervisningen er fulgt de to første år i gymnasiet. Forudsætningen for at kunne besvare spørgsmålene i den skriftlige prøve efter 2.g er, at man har fulgt undervisningen (en elev alene med folkeskolebaggrund ville næppe kunne svare på ét eneste af spørgsmålene). Endvidere gælder for både jura- og økonomistudiet, at en forbedring af præstationen i matematik fra 2.g til 3.g har en selvstændig betydning for beståelsen. Den oplagte fortolkning er, at karaktererne i matematik afspejler en kombination af evner til matematik og indsats i faget. Selve kurset i matematik i 3.g har en selvstændig betydning for beståelsen af økonomi, men ikke på jura. For de øvrige samfundsvidenskabelige fag opnås der ikke nogen positiv effekt på beståelsen, hverken hvad angår evner til matematik eller gennemførelsen af kurset i 3.g.

6. Diskussion

Både kurser i matematik og evner for matematik betyder noget for beståelsen af de største fag på samfundsvidenskab på Københavns Universitet, jura og økonomi. Elever fra det almene gymnasiums matematiske linje er bedre til at bestå disse studier end elever fra det almene gymnasiums sproglige linje. På begge studier har den kombination af evner for matematik og indsats i faget, som måles ved karakteren i matematik, en positiv, selvstændig og væsentlig betydning for beståelsen (udover bidraget til et generelt højere gennemsnit). Selve kurset i matematik i 3.g har en positiv effekt for beståelsen på økonomi, men ikke på jura.

Studenter med en gennemsnitlig studentereksamen har gode muligheder for at bestå de mest attraktive studier på samfundsvidenskab, hvis de ellers kan komme ind. Derimod har de ikke så gode muligheder for bestå de studier, hvor der er en mild adgangsbegrænsning som jura og økonomi. Regeringen har gennemført en forøgelse af optaget på kvote 1 og en tilsvarende *formindskelse af kvote 2*. Det vil sige, at der på nogle af de mest attraktive studier på Københavns Universitet optages flere studenter med høje karakterer ved studentereksamen og færre med karakterer svarende til en gennemsnitlig studentereksamen. Men de studerende med høje gennemsnit til studen-

tereksamen har netop gode muligheder for at bestå studierne med mild adgangsbe- grænsning, jura og økonomi.

Der er den mulighed, at studierne med hård adgangsbe- grænsning har været gode til at finde og optage de studerende, som på trods af en gennemsnitlig studentereksamen er gode til at bestå studierne. Hvis det er tilfældet, er det en usædvanlig dårlig ide at formindske omfanget af optagne på kvote 2. Disse studerende har netop meget svært ved at klare sig på studier med frit optag.

En mulighed for at forbedre det nuværende system kunne være lodtrækning om adgang til studierne på kvote 2. Eventuelt kunne den nuværende kvote 1 medtages i lodtrækningen. De dygtige studenter, der taber lodtrækningen, har gode muligheder for bestå et andet studium. Med en gennemsnitlig adgangsgivende eksamen har de studenter, der vinder lodtrækningen, bedre mulighed for at bestå disse studier end de studier, de ellers har adgang til. En omfordeling af ansøgerne via en sådan kvote med lodtrækning har derfor potentialet til at forøge andelen af studenter, der består en videregående ud- dannelselse.

Det traditionelle argumentet mod lodtrækning er, at det er de bedst egnede, der bør optages på studierne. Men problemstillingen er her, at de bedst egnede klarer sig dårligere end de studerende, der ikke er helt så egnede (når egnethed måles ved gennemsnittet ved den adgangsgivende eksamen). En lodtrækning har den betydelige fordel, at den formindsker den skadelige konkurrence i gymnasiet om at få så høje karakterer, at det bliver muligt at komme ind på attraktive studier med begrænset optag og med meget høje adgangskvotienter. Lodtrækningen vil bevirke, at eleverne i gymnasiet med sindsro kan vælge fag, som forbereder dem til at bestå studier, de kan komme ind på, hvis de taber lodtrækningen.

Regeringen har gennemført en *belønning af elever, der yder en særlig indsats* ved at tage ekstra fag på højt niveau i gymnasiet. Men i sig selv har valg af ekstra fag på højt niveau ingen betydning for beståelsen på samfundsvidenskab ifølge resultaterne i dette indlæg.

Den ændrede incitamentsstruktur vil have konsekvenser for elevernes valg af fag i gymnasiet. For at få plads til tre fag på højt niveau i stedet for to vil det være oplagt at fravælge fag, der er svære og arbejdskrævende. Matematik er svært og arbejdskræven- de, og er derfor en oplagt kandidat til at blive fravalgt, når der skal gøres en ekstra ind- sats. I det omfang matematik fravælges, formindskes kvaliteten af rekrutteringsgrund- laget til de videregående uddannelser. Det samme gør sig gældende, hvis matematik vælges som ét af tre fag på højt niveau, men at indsatsen i matematik formindskes som følge af det ekstra fag.

Et væsentligt element i overvejelserne om *justering af gymnasiereformen* er beslut- ningen om, hvilke valg eleverne skal stilles over for. Dette indlæg har vist, at elevernes

valg af matematik klart er påvirket af incitamentsstrukturen i gymnasiet. Elever med lavt gennemsnit fravælger matematik i større omfang end elever med højt gennemsnit. Hvis eleven opnår lave karakterer i matematik sammenlignet med de øvrige fag fravælges matematik i større omfang sammenlignet med elever, der klarer sig relativt godt i matematik. Dette fravalg af matematik forøger selvfølgelig elevens muligheder for at komme ind på studier med adgangsbegrænsning. Men det indebærer samtidigt en formindskelse af elevens mulighed for at bestå en videregående uddannelse. En ændring af dette forhold kan gennemføres ved enten at justere incitamentsstrukturen eller ved at justere valgmulighederne i gymnasiet.

Litteratur

- Albæk, K. 2001. Hvem består på politstudiet, *Nationaløkonomisk Tidsskrift*, nr. 139, s. 208-22. Optrykt som *Blåt Memo* nr. 204, Økonomisk Institut, Københavns Universitet, januar 2002.
- Albæk, K. 2002. Incitamentsstrukturen i gymnasiet i relation til de videregående uddannelser. *Samfundøkonomen*, nr. 5, s. 20-27.
- Albæk, K. 2003. Optimal adgangregulering til de videregående uddannelser og elevs valg af fag i gymnasiet. *Nationaløkonomisk Tidsskrift* 141, nr. 2, november, s. 206-24.
- Albæk, K. 2004a. Sådan modvirkes gymnasie-reformens virkninger på nationens nationalprodukt, økonomiske vækst og åndsliv. *Blåt Memo* nr. 210, Økonomisk Institut, Københavns Universitet, januar 2004. Revideret version af manuskript udarbejdet til Nationaløkonomisk Forenings årsmøde på Koldingfjord den 9.-10. januar 2004.
- Albæk, K. 2004b. Gymnasie-reform: En illustration af optimal adgangregulering. *Samfundøkonomen*, nr. 5, s. 18-25.
- Albæk, K. 2005. Gymnasie-reformen – ide og realisation. *Samfundøkonomen*, nr. 5, s. 31-34.
- Albæk, K. 2006. Evner, fag, indsats i gymnasiet og beståelse på samfundsvidenskab *Blåt Memo* nr. 213, Økonomisk Institut, Københavns Universitet, april 2006. Revideret version af manuskript udarbejdet til Nationaløkonomisk Forenings årsmøde på Koldingfjord den 6.-7. januar 2006.
- Andersen, S. og P. Østergaard. 2005. Fra elev i gymnasiet til studerende ved en videregående uddannelse – erfaringer med præstationerne i statistik fra uddannelserne ved Handelshøjskolen i Århus. I *Symposium i Anvendt Statistik 2005*, Økonomisk Institut, Syddansk Universitet og Danmarks Statistik, s. 387-421.
- Hayashi, F. 2000. *Econometrics*. Princeton: Princeton University Press.
- Kristensen, I. Pagter. 1998. *Studiestatistisk rapport nr. 6*. Institut for Statskundskab, Aarhus Universitet – kan hentes på <http://www.ps.au.dk/>.
- La Cour, L. and P. Raimondos-Møller. 2005. What affects students' performance? An investigation of the importance of admission characteristics. *Nationaløkonomisk Tidsskrift* 143, nr. 3, december, s. 359-79.
- Undervisningsministeriet. 2005. Bekendtgørelse om adgang m.v. ved bachelor- og kandidatuddannelser ved universiteterne. Bekendtgørelse nr. 365 af 20/05/2005.

Tabel A1. Karakteristika for studerende på samfundsvidenskaber og andel, der har bestået bachelordelen inden for fire års studier, immatrikulation sommer 1994 – sommer 2001.

	Jura		Økonomi		Statskundskab		Psykologi		Antropologi		Sociologi		Total	
	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede
<i>Adgangsgivende eksamen:</i>														
Matematisk student	0,39	0,69	0,68	0,47	0,46	0,73	0,24	0,69	0,22	0,55	0,43	0,58	0,44	0,61
Sproglig student	0,33	0,62	0,04	0,31	0,30	0,76	0,29	0,62	0,39	0,60	0,24	0,54	0,25	0,62
Studenterkursus	0,01	0,31	0,01	0,18	0,32	0,32	0,01	0,53	0,02	0,40	0,01	0,86	0,01	0,33
HF	0,13	0,37	0,08	0,15	0,12	0,40	0,36	0,42	0,27	0,44	0,26	0,40	0,15	0,37
HHX	0,11	0,56	0,13	0,33	0,05	0,59	0,04	0,64	0,02	0,73	0,04	0,40	0,09	0,49
HTX	0,00	0,54	0,02	0,27	0,00	0,29	0,00	0,50	0,00	0,00	0,00	1,00	0,01	0,34
Udland	0,01	0,43	0,02	0,22	0,02	0,25	0,04	0,47	0,03	0,33	0,01	0,50	0,02	0,36
Andre	0,01	0,42	0,02	0,20	0,03	0,72	0,02	0,71	0,04	0,54	0,01	0,67	0,02	0,49
<i>Eksamensalder:</i>														
0	0,20	0,58	0,23	0,39	0,13	0,61	0,18	0,46	0,11	0,36	0,11	0,46	0,19	0,51
1	0,29	0,67	0,34	0,45	0,28	0,76	0,17	0,57	0,21	0,43	0,18	0,54	0,28	0,60
2	0,17	0,65	0,21	0,42	0,18	0,71	0,14	0,64	0,19	0,52	0,18	0,56	0,18	0,59
3	0,10	0,59	0,10	0,39	0,13	0,72	0,11	0,67	0,14	0,70	0,11	0,54	0,10	0,58
4	0,06	0,65	0,04	0,35	0,09	0,71	0,11	0,65	0,08	0,61	0,10	0,50	0,07	0,61
5	0,03	0,57	0,02	0,33	0,07	0,59	0,05	0,65	0,09	0,58	0,10	0,67	0,04	0,57
Mindst 6	0,14	0,40	0,06	0,12	0,11	0,47	0,24	0,47	0,19	0,60	0,23	0,42	0,13	0,41
<i>Kvotient:</i>														
-6,-9	0,01	0,24	0,04	0,09	0,00	0,40	0,01	0,63	0,01	0,20	0,01	0,50	0,02	0,19
7,0-7,4	0,02	0,37	0,05	0,07	0,02	0,35	0,00	0,67	0,02	0,69	0,02	0,36	0,03	0,25
7,5-7,9	0,04	0,48	0,14	0,20	0,03	0,63	0,02	0,33	0,02	0,46	0,06	0,41	0,06	0,33

fortsættes...

fortsat ...

	Jura		Økonomi		Statskundskab		Psykologi		Antropologi		Sociologi		Total	
	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede
8,0-8,4	0,09	0,60	0,21	0,28	0,07	0,45	0,05	0,60	0,03	0,67	0,08	0,46	0,11	0,44
8,5-8,9	0,27	0,58	0,21	0,41	0,08	0,70	0,07	0,65	0,08	0,54	0,07	0,44	0,19	0,54
9,0-9,4	0,32	0,60	0,17	0,55	0,10	0,67	0,09	0,68	0,09	0,78	0,12	0,66	0,21	0,60
9,5-9,9	0,17	0,67	0,10	0,69	0,36	0,71	0,36	0,55	0,21	0,44	0,41	0,55	0,21	0,64
10,0-10,4	0,06	0,70	0,05	0,74	0,25	0,71	0,29	0,53	0,39	0,55	0,18	0,49	0,13	0,63
10,5-	0,02	0,66	0,02	0,67	0,09	0,73	0,11	0,51	0,14	0,47	0,05	0,54	0,05	0,61
<i>Køn:</i>														
Kvinde	0,55	0,63	0,27	0,42	0,45	0,72	0,75	0,57	0,78	0,55	0,71	0,54	0,51	0,59
<i>I alt:</i>	1,00	0,60	1,00	0,40	1,00	0,68	1,00	0,56	1,00	0,54	1,00	0,52	1,00	0,55
<i>Antal observationer:</i>	4.919	2.944	2.736	1.096	1.531	1.036	1.231	688	540	290	536	281	11.493	6.335

Tabel A2. Karakteristika for studerende på samfundsvidenskab fordelt på højniveaufag og andel, der har bestået bachelordelen inden for fire års studier, imatrikulation sommer 1994 – sommer 2001.

	Jura		Økonomi		Statskundskab		Psykologi		Antropologi		Sociologi		Total		
	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel	Andel beståede	Andel beståede	Andel	Andel beståede	Andel beståede	Andel	Andel beståede	Andel	Andel beståede
<i>Matematikere:</i>															
Matematik og fysik på A niveau	0,09	0,78	0,16	0,57	0,09	0,66	0,06	0,90	0,07	0,50	0,03	0,75	0,11	0,65	
Matematik og samfunds-fag på A niveau	0,24	0,75	0,35	0,54	0,44	0,80	0,17	0,56	0,28	0,46	0,22	0,71	0,31	0,65	
Matematik og engelsk på A niveau	0,25	0,71	0,24	0,46	0,16	0,71	0,19	0,77	0,24	0,62	0,15	0,57	0,23	0,60	
Matematik på B niveau	0,23	0,73	0,11	0,29	0,23	0,79	0,25	0,85	0,20	0,59	0,30	0,70	0,19	0,64	
Andre fag på A niveau	0,22	0,75	0,19	0,46	0,14	0,62	0,35	0,72	0,26	0,70	0,32	0,57	0,21	0,62	
Tre højniveau fag	0,07	0,78	0,09	0,53	0,09	0,76	0,07	0,82	0,14	0,67	0,04	0,83	0,08	0,67	
I alt: (e)	1,00	0,73	1,00	0,48	1,00	0,75	1,00	0,75	1,00	0,60	1,00	0,64	1,00	0,63	
Antal observationer:	1.385	1.021	1.487	720	511	384	162	121	87	52	159	102	3.791	2.400	
<i>Sproglige:</i>															
Matematik på B niveau	0,13	0,69	0,45	0,33	0,07	0,84	0,12	0,71	0,10	0,69	0,10	0,33	0,13	0,64	
Matematik på C niveau	0,19	0,70	0,18	0,44	0,31	0,73	0,19	0,62	0,10	0,40	0,23	0,55	0,21	0,68	
Naturfag	0,68	0,63	0,37	0,22	0,63	0,81	0,69	0,73	0,80	0,57	0,67	0,59	0,67	0,65	
Engelsk på A niveau	0,75	0,66	0,70	0,33	0,61	0,79	0,66	0,69	0,66	0,62	0,63	0,62	0,71	0,66	
Samfunds-fag på A niveau	0,37	0,63	0,28	0,25	0,62	0,80	0,23	0,78	0,29	0,59	0,43	0,54	0,39	0,67	
Andre A niveaus fag	0,57	0,65	0,71	0,32	0,44	0,76	0,64	0,69	0,64	0,59	0,58	0,51	0,56	0,64	
Tre højniveau fag	0,04	0,65	0,02	0,50	0,05	0,75	0,02	0,25	0,05	0,71	0,02	0,05	0,04	0,65	
I alt: (e)	1,00	0,65	1,00	0,31	1,00	0,79	1,00	0,70	1,00	0,56	1,00	0,56	1,00	0,66	
Antal observationer:	1.307	854	87	27	354	280	202	142	154	87	95	53	2.199	1.443	

Note: (e) Tallene ovenover adderer til mere end én, da mere end to fag kan vælges på A-niveau, f.eks. matematik, fysik og engelsk.

Price Setting Behaviour in Denmark – A Study of CPI Micro Data 1997-2005

Bo William Hansen

Danmarks Nationalbank, E-mail: bwh@nationalbanken.dk

Niels Lynggård Hansen

Danmarks Nationalbank, E-mail: nlh@nationalbanken.dk

SUMMARY: This paper provides empirical evidence on the degree of price rigidity and price flexibility in Denmark. We use unpublished data from Statistics Denmark on the Danish CPI. The dataset covers the period 1997-2005 and contains around 2.7 million monthly price records. The paper reveals a substantial amount of heterogeneity in the frequency and size of price adjustments across sectors and products. Most price changes are increases, but price decreases are not uncommon. Price changes are generally sizeable compared to aggregate and sectoral inflation rates. We explore how these features are affected by, e.g., seasonality and the level of inflation. Our evidence emphasises the importance of price stickiness and supports the existence of both time and state-dependent pricing strategies.

1. Introduction

Price stickiness is a standard assumption in macroeconomic models and is considered to be a key ingredient for the understanding of the economy's reaction to a wide range of shocks, including the persistent, though not permanent reaction of real output to monetary shocks. In the new Keynesian literature of the 1980s different microeconomic models of price setting behaviour have been proposed under the label of either

This paper is based on Hansen and Hansen (2006) containing a larger set of charts and tables. The paper has benefited from comments given at the biannual meeting of the Danish Economic Society, Koldingfjord, January 2006, as well as at workshops at Statistics Denmark, EPRU (Economic Policy Research Unit at the University of Copenhagen), Norges Bank and Danmarks Nationalbank. We are grateful to Statistics Denmark for providing us with the dataset and for many helpful comments. A special thanks to Mariann Søndergaard for valuable programming assistance. Views expressed are those of the authors, and should not be considered as the official position of Danmarks Nationalbank. Likewise, errors and omissions are the responsibility of the authors.

time-dependent or state-dependent models.¹ Some of these models are now used in the new generation of micro-founded macroeconomic models for policy analysis.

Despite the importance of pricing assumptions for macro models the empirical evidence from micro data has remained relatively scarce. Taylor (1999) gives an overview of the stylised facts of price setting behaviour. More recently, however, Bils and Klenow (2004), followed by Klenow and Kryvtsov (2005), have studied price setting behaviour in the US by examining a large dataset of prices used in the computation of the consumer price index (CPI). Similar comprehensive studies have subsequently been undertaken for the euro area countries within the Inflation Persistence Network (IPN) under the Eurosystem, and also for other countries.²

In this paper we provide descriptive evidence about price setting behaviour in Denmark. To that end we analyse the micro data underlying Danish CPI for the years 1997-2005. Our dataset covers the whole expenditure basket and contains around 2.7 million monthly price records.

We provide evidence on the frequency of price changes, including the duration of price spells and hazard functions, as well as on the size of price changes. Price increases and decreases are treated separately, and special attention is devoted to the investigation of heterogeneity across sectors. As in related contributions we examine how these features are related to the overall evolution of inflation. The analysis is descriptive and merely yields some indicative information regarding the importance of different pricing models. More structural analyses, including formal econometric testing of competing pricing models, are left for future studies.

The structure of the paper is as follows. The next section provides a description of the dataset, including a discussion of data issues, price trajectories, coverage, weighting, etc., as well as a brief overview of the development of inflation over the sample period. Section 3 contains methodology and notation. Empirical results are presented in section 4, and in section 5 the covariation between inflation and the frequency, and size, of price increases and decreases is analysed. Finally, our main results are summarised in section 6.

1. In time-dependent models the timing of the individual firm's pricing decision is independent of the state of the economy. Time-dependent models include Taylor's (1980) staggered pricing model, in which firms keep prices fixed for a certain number of periods, and the Calvo (1983) model, in which firms are allowed to adjust their prices in a given period with a certain probability. In state-dependent models the timing of the firms' pricing decisions is a function of the state of the economy, as in menu-cost models, see e.g. Dotsey, King and Wolman (1999). Time-dependent models are often preferred in macroeconomic models as they are simpler to implement.

2. The Eurosystem comprises the European Central Bank and the national central banks of those countries that have adopted the euro. More information about the IPN, including a large set of articles and working papers, can be found at <http://www.ecb.int/home/html/researcher-ipn.en.html>.

2. The dataset

2.1. *The micro data underlying the Danish CPI*

The dataset contains the micro data collected by Statistics Denmark in order to compute the Danish CPI. It covers the period from January 1997 to December 2005 (108 months). A general introduction to, and overview of, the methodology for compiling the CPI is provided in Statistics Denmark (2005).

The raw database is made up of 2,731,841 monthly price records, which corresponds to around 25,000 records per month on average. Prices are collected between the 7th and 15th of every month. Most individual price quotes refer to a specific item sold in a particular retail outlet at a given point in time. The dataset is subject to statistical confidentiality restrictions and does not enable the name or location of a given outlet to be identified. For each record, we observe the following information:

- the price of the item
- the year and month
- the brand name of the item
- a numeric product code
- a numeric product category code
- the name of the product category
- a numeric code for a given outlet chain
- a numeric outlet code

Together the four numeric codes allow us to identify and track each individual item, i.e. a specific product in a specific outlet. The product category code corresponds to the COICOP³ 5-digit code, also denoted as the elementary product level. From this we obtain a breakdown into COICOP-divisions⁴ (two-digit level), as well as into five special aggregates or components, typically used within the Eurosystem to analyse inflation developments.⁵ In the following we concentrate on the latter breakdown. It has fewer categories, and the groups are fairly homogenous in terms of item characteristics, such as pricing behaviour, and therefore better suited for our analysis. The COICOP-divisions represent a classification of expenditure according to purpose, entailing a

3. The internationally agreed Classification of Individual Consumption according to Purpose.

4. This is the most commonly used decomposition of Danish consumer prices. It entails a breakdown into 12 subgroups: (1) food and non-alcoholic beverages, (2) alcoholic beverages, tobacco and narcotics, (3) clothing and footwear, (4) housing, water, electricity, gas and other fuels, (5) furnishings, household equipment and routine household maintenance, (6) health, (7) transport, (8) communication, (9) recreation and culture, (10) education, (11) restaurants and hotels, (12) miscellaneous goods and services.

5. These are: (1) unprocessed food, (2) processed food including alcohol and tobacco, (3) energy, (4) non-energy industrial goods, (5) services.

large heterogeneity within divisions with respect to, e.g., the production process and pricing behaviour.⁶

2.2. Data issues

The dataset is considered to be of a high quality and well suited for our purpose. However, some important issues need to be mentioned.

The raw dataset covers 100 percent of the expenditure on the Danish CPI basket. To align our study with European standards, we focus on the expenditure basket as defined by the EU Harmonised Index of Consumer Prices (HICP). Owner-occupied dwellings and insurance in connection with owner-occupied dwellings are therefore eliminated, as these are not included in the HICP.

For a number of goods and services prices are not collected monthly, but only quarterly, biannually or annually as they are considered not to change very often. Examples of this include housing rent, postage, license fees, social protection services or financial services. In such situations, the price is carried forward unchanged between the two collection periods.

The dataset has beforehand been subject to statistical editing – plausibility checks – by Statistics Denmark in order to identify possible errors, e.g. extreme price observations or prices that have remained unchanged for a long period of time.⁷ Erroneous price reports resulting in either very large, or very small price changes have thus been removed.

As we are only interested in analysing the market-based price setting behaviour, all administered prices, which to a significant extent are regulated by the government, are eliminated from the dataset. There exists no official delineation of administered prices. The following products, mainly services, have been isolated as products with administered prices: rent, pharmaceuticals, medical services, public transport, radio and television license fees, education, social protection, hunting licences and passport fees.

There are no statistically imputed prices included in the dataset. Imputed prices are used to deal with missing observations, seasonality and quality change in price indices, see Armknecht and Maitland (1999) for statistical imputation techniques.

Not all price quotes are directly linked to an individual item sold in a particular outlet.⁸ In these cases prices are omitted from the dataset.

6. As an example take transport. It includes both the purchase of vehicles, the repair of vehicles, the purchase of fuels for private cars, and fares related to the transport of passengers by rail, air, etc., that is, goods and services that are very different from each other in terms of production and pricing.

7. The statistical editing in the Danish CPI is based on the method developed by Hidiroglou and Berthelot (1986).

8. For pharmaceuticals the price index compiled by the Danish Medicines Agency, which shows the development in the average price of medicine, is used. In the same way, the prices on personal computers and printers – where the model characteristics often change – are averages, calculated on the basis of a large

Another important characteristic of the price records is that the prices are inclusive of all types of sales, rebates and promotions. Compared to databases that exclude such discounts, this can lead to a higher frequency of price adjustment, particularly for the products where sales are common, e.g. clothing and footwear.

2.3. Price trajectories and price spells

From the information in the dataset we construct individual price trajectories, i.e. sequences of price quotes for a specific item sold in a specific outlet. In most cases trajectories are considerably shorter than 108 months. We impose the – rather unrestrictive – rule that every trajectory should cover a period of at least three months, otherwise it is omitted.

Figure 2.1 shows exemplary price trajectories for the five product categories unprocessed food (sirloin of beef), processed food (Danish bitter), energy (district heating, fuel), non-energy industrial goods (armchair) and services (car insurance), selected for their typical pattern. Pricing behaviour is visibly very heterogeneous and the average length of a price spell – defined as the sequence of price quotes with the same price – varies significantly between products. Some prices are adjusted frequently (sirloin of beef, fuel), while other prices (Danish Bitter, district heating, car insurance) appear to be adjusted regularly every six/twelve months. The size of price changes also differs substantially between products. The large price reduction on Gammel Dansk (Danish bitter) in October 2003 reflects a lowering of the excise duties on alcohol, cigarettes and soft drinks in that month (the cross-frontier-trade package).

For quite a number of trajectories, one or more price quotes are missing.⁹ Typically the unfilled interval relates to a single month, but prices are sometimes unavailable for several consecutive months. In such situations, we split up the trajectory concerned into new individual trajectories. For example, the price sequence for the sirloin of beef in Chart 2.1 is treated as three separate trajectories.¹⁰

continued...

sample of prices from the internet (using a so-called matched model approach). Finally, the recorded prices on books have been divided by the natural logarithm of the number of pages in each book. The latter is an attempt to adjust for price movements resulting from differences in the number of pages when calculating the price index for books.

9. Missing prices come about for many reasons e.g. because the product is out-of-stock, the outlet is temporarily closed or the price was not timely collected.

10. In order to deal with missing prices, some of the studies performed within the IPN carry forward the price unchanged from the previous period. We decided not to pursue this method, as we want to strictly rely on observed market prices. Our method reduces the average length of trajectories compared to studies which replace the missing prices. Another option is to simply disregard those trajectories containing missing prices, but in our case that would imply eliminating more than 25 percent of our observations.

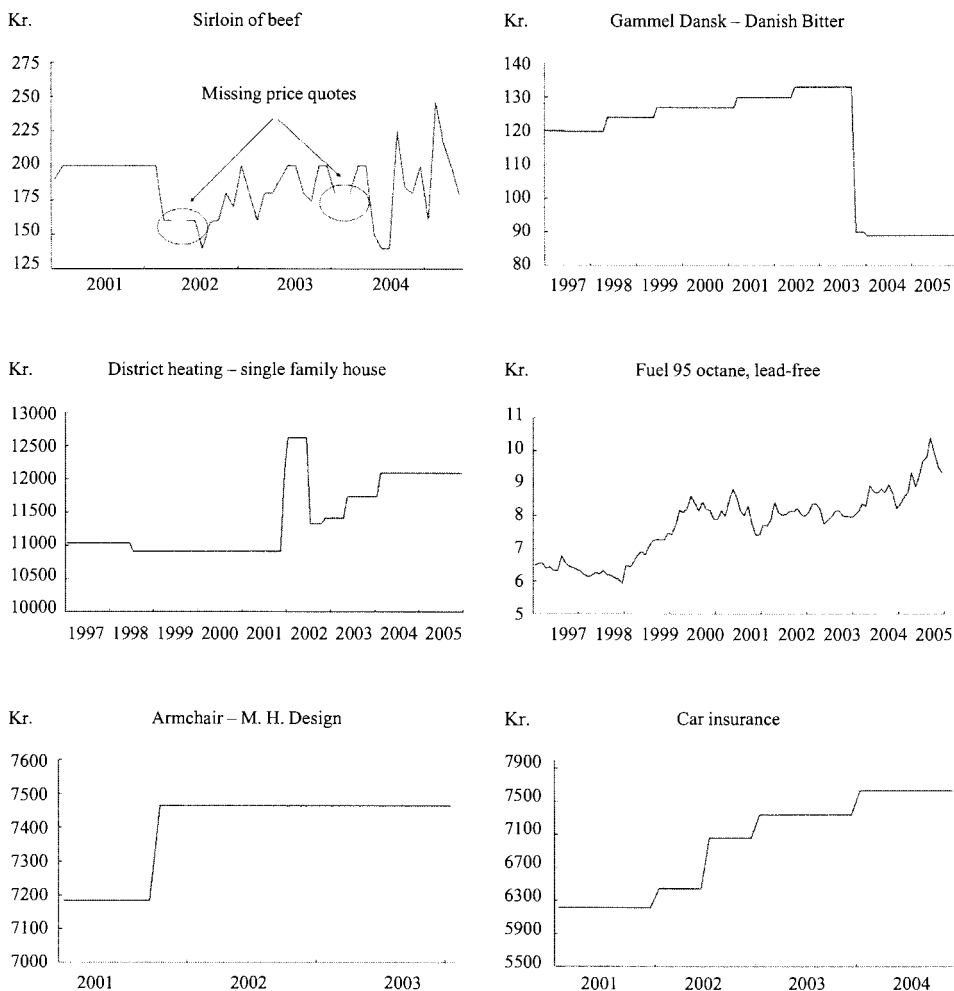


Figure 2.1. Examples of individual price trajectories.

Note: Prices are in levels, denominated in Danish Kroner.

2.4. The coverage of the dataset and weighting

As a result of our filtering of data 108,352 monthly price records are omitted, leaving us with a total of 2,623,489. The number of price records is very unevenly distributed across product groups, cf. Table 2.1.

In order to produce aggregate measures of the statistics described in section 3, we compute weighted averages using the official HICP weights published by Statistics Denmark. As the dataset spans over two goods baskets – 1999 and 2003 – the average

Table 2.1 Coverage of the dataset.

	Official HICP weight	Weight after filtering	# observations after filtering	# items	Average length of trajectories
<i>Main components:</i>					
Unprocessed food	6.01	6.01	661,857	45,368	14.6
Processed food incl. alcohol/tobacco	14.13	14.13	572,576	27,483	20.8
Energy	10.70	10.70	25,290	366	69.1
Non-energy industrial goods	30.69	28.50	1,054,853	53,104	19.9
Services	38.46	25.87	308,913	9,591	32.2
All-items	100.00	85.21	2,623,489	135,912	19.3

Note: The weights are reported as percentages, while the average length of trajectories is in months. The weights have been computed as averages of the two official weighting schemes with reference periods in 1999 and 2003 respectively.

weights of the two weighting schemes are used. The elementary product level is the most detailed level for which the weights are defined. There are around 450 categories/weights at this level. All statistics at the elementary product level are computed as unweighted averages using all the observations of items belonging to that category. Aggregate statistics are then computed by averaging over elementary products using weights.

The filtered dataset constitutes 85 percent of HICP's weight basis, cf. Table 2.1, and comprises 411 elementary product categories. The majority of the eliminated expenditure components are related to administered prices, leading to a significant reduction in the relative weight for services. Unprocessed food is clearly overrepresented in terms of observations, representing 24 percent of the dataset, but only a weight of 7 percent (6.01/85.21).

2.5. Consumer price developments over 1997-2005

For reference purposes it is helpful to describe the evolution of consumer prices throughout the sample period. Figure 2.2 illustrates the year-on-year rate of change of HICP and the main subcomponents.¹¹

The different components display different inflation developments. Whereas aggregate annual inflation was relatively stable over the sample period, the evolution of the indices for unprocessed food, processed food, and in particular energy, was more volatile. The large movements in the energy index were decisively affected by oil price fluctuations. More pronounced shifts in inflation may better facilitate an investigation of the factors driving the inflationary process, see below.

11. The official indices for all-items, non-energy industrial goods and services include some expenditure components which have been omitted from our dataset.

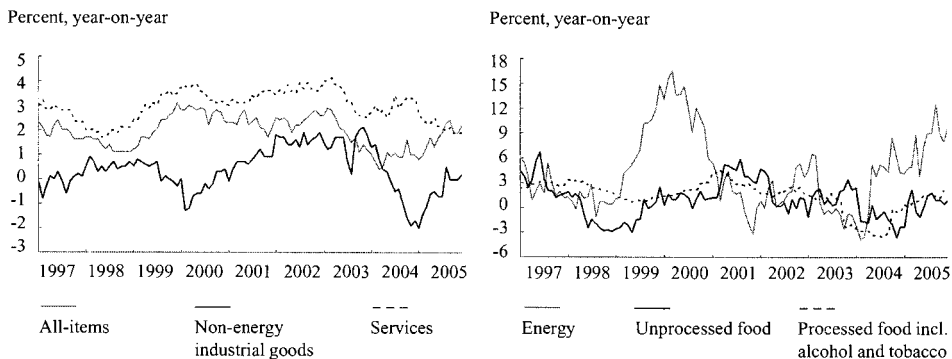


Figure 2.2. Annual inflation rates, HICP and subcomponents.

Source: Eurostat.

3. Methodology

In order to gain an insight into the qualitative nature of the price setting process, we look at the frequency of price adjustments, including the duration of price spells and empirical hazard functions, and size of price adjustments.

This section contains an overview of definitions, notation and formulae used in the computations, as well as an assessment of potential biases and aggregation issues.

3.1. Frequency of price changes and duration

Let $I_{j,t}$ be an indicator for price changes for item j in period t :

$$I_{j,t} = \begin{cases} 1 & \text{if } p_{j,t} \neq p_{j,t-1} \\ 0 & \text{if } p_{j,t} = p_{j,t-1} \end{cases}$$

where $p_{j,t}$ is the price of item j at time t . The first observation in any price trajectory is thereby discarded.

For each elementary product category n the (unweighted) frequency of price changes can be written

$$F_n = \frac{1}{Q_n} \cdot \sum_{j=1}^J \sum_{t=1}^{T_j} I_{j,t}$$

Q_n is the number of observations in product category n , J is the number of items in that category and T_j is the number of observations for item j . Let N be the number of

product categories. The aggregate frequency of price changes is then given by the weighted average of the frequencies at the product category level,

$$F = \sum_{n=1}^N \omega_n F_n$$

where ω_n denotes HICP weights.

The average duration of a price spell is inferred implicitly from the computed frequency (frequency approach). If, under the assumption of continuous time, prices can change at any moment with a constant probability within every time period, the average duration for product category n becomes¹²

$$D_n = \frac{-1}{\ln(1 - F_n)}$$

Alternatively, the duration could be computed directly as the average length of a price spell (duration approach). The duration approach implies a censoring of data as observations before the first price change and after the last price change in any price trajectory must be removed, hence introducing a potential bias as a larger fraction of observations for items with frequent price changes are retained. The frequency approach allows the use of the full dataset and avoids the potential bias from the censoring of the data, but requires on the other hand specific distributional assumptions about the distribution of price changes over time.¹³

The aggregate duration is given by the weighted mean of durations at the product category level,

$$D = \sum_n w_n \left(\frac{-1}{\ln(1 - F_n)} \right) = \sum_n w_n D_n$$

This measure is sensitive to potential outliers, from those cases where frequencies come close to zero, resulting in very high values of the implied durations. Bearing this in mind, we also compute the aggregate median price duration based on the median

12. If instead we assume that price changes happen at most once a month, the mean duration of price spells is asymptotically given by the inverse of the frequency of price changes, $D = 1/F$. However, we consider the chosen method to be more relevant. The difference between the two measures is small when the frequency of price changes is relatively low, as for most product categories in our sample, whereas our measure leads to a shorter duration than the inverse frequency, when the frequency is high.

13. See Dhyne et al. (2005) for a thorough discussion of this issue.

frequencies at the product category level. This measure, which is used by e.g. Bils and Klenow (2004), may be a reasonable alternative indicator of aggregate duration, even if it is not an estimator of the average price duration, as pointed out by Baudry et al. (2004).

The aggregate duration can, in principle, also be derived by simply inverting the aggregate frequency. As demonstrated below, such »pseudo« durations are significantly lower than the weighted durations.¹⁴

3.2. Size of price changes

The size of price changes is computed as logarithmic differences multiplied by 100.¹⁵ If prices are adjusted more than once in a month, the size represents the cumulative change during the month. Similar to the other aggregate statistics, the aggregate size of price changes is computed as a weighted average.

3.3. Empirical hazard functions

The hazard function of item j represents the probability of individual price changes in a given period as a function of the time elapsed since the last price change, k , i.e.

$$\lambda_j(k) = P(I_{t,j} = 1 \mid I_{t-i,j} = 0 \quad \forall i = 1, \dots, k-1)$$

In practice the empirical hazard function is estimated by splitting price trajectories into price spells, taking into account left censoring, i.e. discarding observations before the first price change for any item.

Aggregate hazard functions are often found to be decreasing with the time elapsed since the last price change, see e.g. Dhyne et al. (2005), despite prevailing theories of price setting implying hazard functions to be either constant or increasing with time. However, there need not be any contradiction as non-decreasing individual hazard rates might result in decreasing hazard rates at the aggregate level. The reason for this is oversampling, as products with frequent price changes will result in many short price spells and hence dominate the estimation of hazard rates at short time horizons. Hence this bias can be explained by heterogeneity in price setting either at the store level or at the product level, or both. Álvarez et al. (2005) show theoretically that aggregating

14. The difference reflects Jensen's inequality coupled with a significant variation across product categories regarding the frequency of price changes. Jensen's inequality says that if f is a strictly convex function – like the one relating duration to the frequency of price changes, $D = f(F) = -1/(\ln(1-F))$ – and X is a random variable that takes the value of its expected mean, EX , with probability less than one, then $Ej(X) > f(EX)$.

15. For small changes log-differences are approximately equal to percentage changes. The former has the advantage of preserving additivity, implying that a change in a price from A to B followed by a return to A will be reported as two price changes of the same size, but with opposite signs.

Table 4.1. Frequency of price changes, monthly basis, percent.

	Frequency of price ...			% of price increases
	... changes	... increases	... decreases	
<i>Main components:</i>				
Unprocessed food	39.8	20.6	19.2	51.8
Processed food incl. alcohol/tobacco	10.8	6.8	4.0	63.1
Energy	52.4	28.9	23.5	55.1
Non-energy industrial goods	9.8	5.7	4.2	57.7
Services	8.9	6.7	2.2	75.5
All-items	17.3	10.2	7.1	59.0

across firms applying different time-dependent pricing rules, i.e. with non-decreasing hazard functions, almost always results in aggregate hazard functions that are decreasing.

In this paper we address the problem of an aggregation bias due to heterogeneous price setting by computing the aggregate hazard function as a weighted sum of hazard functions at the elementary product level, where price setting is supposedly more homogenous. Furthermore, we present hazard functions at the disaggregated level. This follows the approach applied by Baumgartner et al. (2005), who find that this generally reduces the negative slope of the aggregate hazard function numerically, and that estimations at a very disaggregate level lead to mostly non-decreasing hazard functions. Similarly, Goette et al. (2005) address the problem of heterogeneity by analysing a small dataset of homogeneous items known to have sticky prices and similar baseline hazards. They find the shape of hazard functions to be increasing, i.e. price-setters become more likely to adjust their prices the longer is the time elapsed since the last price change.

An alternative approach, which will not be pursued in this paper, consists of taking heterogeneity of products into account by estimating the hazard functions conditional on product specific information in panel logit-models with fixed product effects, cf. e.g. Aucremanne and Dhyne (2005) and Gagnon (2005). This generally attenuates the negative slope of the aggregate hazard function, and Baumgartner et al. (2005) even find a positive and significant effect of the duration of a price spell on the probability of a price change.

4. Empirical results

4.1. Frequency of price changes

Table 4.1 presents average monthly frequencies of price changes for the five sectoral components. On average 17.3 percent of prices for all-items are adjusted from

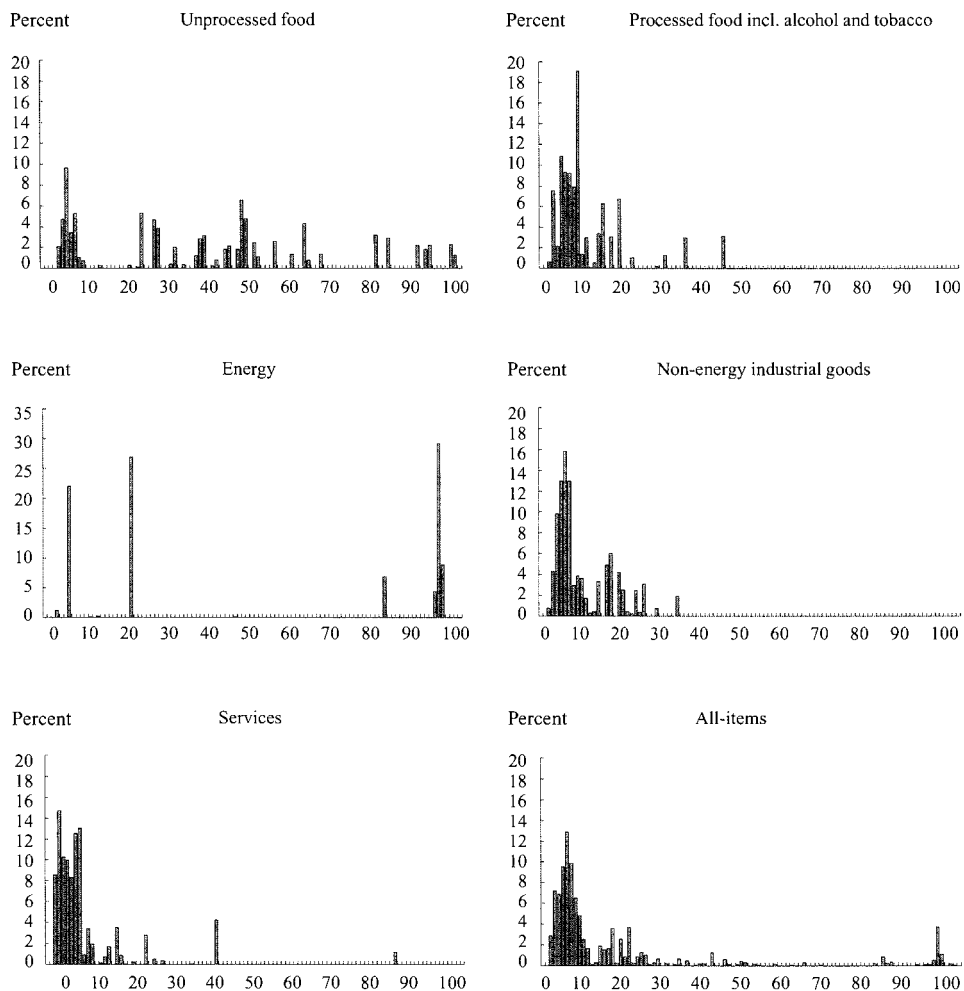


Figure 4.1. Distribution of monthly frequency of price changes.

Note: Frequencies are weighted at the elementary products level.

one month to the next, 10.2 percent are increased and 7.1 percent are decreased, i.e. around 40 percent of all price adjustments are decreases. Prices are therefore lowered nearly as often as they are raised, so prices are not in general more rigid downwards than upwards.¹⁶

16. The large fraction of price decreases could reflect that our dataset includes sales prices. However, there is no identification of sales prices, so any formal analysis on the impact of sales is not possible. Nonetheless, excluding those products typically affected by sales, mainly within industrial goods, does not change our results qualitatively. This is in line with the study by Klenow and Kryvtsov (2005), where the share of price decreases only drops from 45 to 42 percent, when sales prices are excluded.

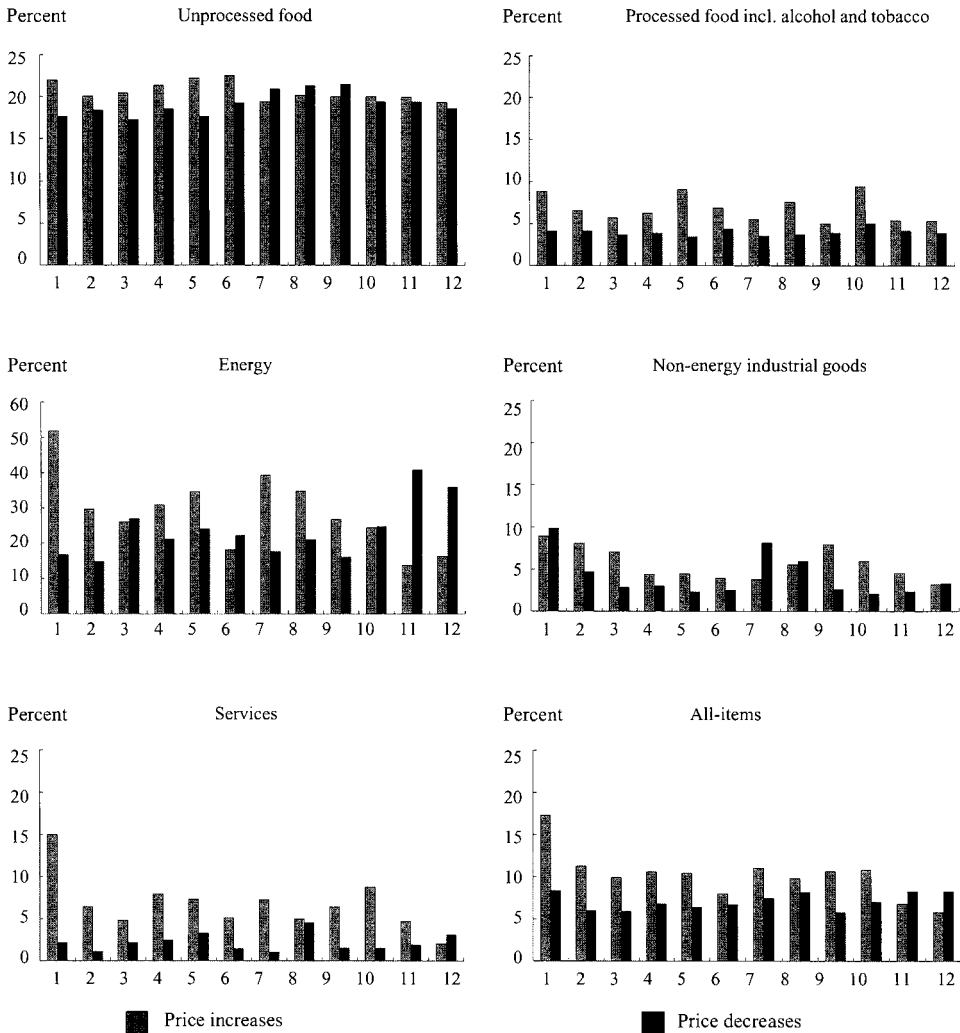


Figure 4.2. Frequency of price changes by month.

Note: Frequencies are weighted at the elementary product level.

Price setting is very heterogeneous across components. Prices for energy and unprocessed food are by far the most flexible, e.g. for energy more than half of the prices are adjusted every month. The frequent price changes for energy and unprocessed food likely reflect that these products are often hit by (supply) shocks, which have a substantial effect on their prices, and not necessarily that the price adjustment costs are lower compared to other items. The highest degree of price stickiness is observed for

services, closely followed by processed food and industrial goods. Price decreases for services occur relatively infrequently, as only 2.2 percent of prices, corresponding to a quarter of all price changes, are lowered every month. In particular prices from restaurants' prices are rarely lowered. This pattern could be related to the high share of labour in the production costs of services, such that stickiness in wage developments translates into infrequent price changes and downward price rigidity.

Not all of the heterogeneity with respect to price changes is captured by our breakdown of data. This becomes evident when looking at the distribution of frequencies at the elementary product level, cf. Figure 4.1.

Energy is composed of items whose prices are adjusted either very frequently, e.g. fuel at the petrol station, and prices that change infrequently, e.g. district heating. None of the other components display a similar dichotomy. Unprocessed food is characterised by evenly distributed frequencies, containing items with high, medium and low frequencies. For processed food, non-energy industrial goods and services, the bulk of the items have frequencies of less than 10 percent.¹⁷

Seasonality seems to be an important characteristic of the frequency of price changes, cf. Figure 4.2.

For unprocessed food price increases are more frequent in the first half of the year, especially in January, April, May and June, than in the second half of the year, whereas price reductions are more frequent in the second half, notably in the third quarter. This pattern is partly driven by prices of fruits and vegetables. For processed food similar spikes are observed, especially regarding the frequency of price increases. The seasonality observed for non-energy industrial goods is to some extent driven by sales. Prices are lowered in January, July and August, where regular end-of-season sales take place, and are raised in the following months. For energy and services most prices are raised in January, reflecting that prices on longer-term contracts such as like electricity, district heating and insurance are often adjusted in January via a simple indexation scheme.

Looking at the frequencies by month and year, cf. Hansen and Hansen (2006), reveals a spike in the frequency of price decreases for processed food in October 2003. This reflects a reduction of excise duties in that month, cf. section 2.3. A number of other spikes are related to changes in excise duties, but some also just reflect randomness. The instant price adjustment by many retailers indicates that simple models of time-dependent pricing are inappropriate for the analysis of this type of policy action. Similar reactions to changes in indirect taxes are found for several euro area countries, see Dhyne et al. (2005). In order to better identify structural features of price setting behaviour, it

17. Frequencies (and sizes) of price changes for each of the 411 elementary products are presented in Hansen and Hansen (2006).

Table 4.2. Duration of price spells, monthly basis.

Implicit average duration based on...	Mean frequencies at the product category level	Median frequencies at the product category level	Inverting the aggregate frequency (pseudo duration)
Unprocessed foods	6.1	5.2	2.0
Processed foods	14.2	12.2	8.8
Energy	5.8	5.4	1.3
Non-energy industrial goods	15.4	10.8	9.7
Services	23.0	17.2	10.7
All-items	15.5	11.8	5.3

could be worthwhile to investigate the price reaction to changes in taxes and duties in more detail.

4.2. Duration of price spells

The aggregate mean duration of all price spells is estimated to be 15.5 months, cf. Table 4.2. As expected, the average price spells of energy and unprocessed food items are the shortest, only lasting around half a year. In contrast, prices of services on average remain unchanged for a period of almost two years.

Besides aggregate mean durations Table 4.2 also presents aggregate median durations as well as aggregate pseudo durations. The median duration for all-items is 11.8 months, nearly 4 months shorter than the mean duration. The gap between the two statistics is largest for components that comprise many items with low frequencies, e.g. non-energy industrial goods and services. This nicely illustrates the problem of outliers, cf. section 3.1. The pseudo durations are significantly lower than both the mean and median durations, e.g. the implied pseudo duration for all-items is just 5.3 months.¹⁸

The aggregate durations mask considerable differences at the elementary products level, cf. Figure 4.3. Almost half of the price spells for unprocessed food and energy lasts less than 2 months, while the prices of some services lasts around 50 months. For all-items the larger part of the distribution relates to durations of less than 30 months.

4.3. Size of price changes

The size of price changes is reported in Table 4.3. Average price changes are relatively large compared to the overall inflation rate observed during the sample period. The size of the typical price increase is 12.3 percent, and that of price reductions is

18. As indicated by Dhyne et al. (2005), a mean duration that is three times higher than the pseudo duration is not unusual for reasonable assumptions on the variation in frequencies across products.

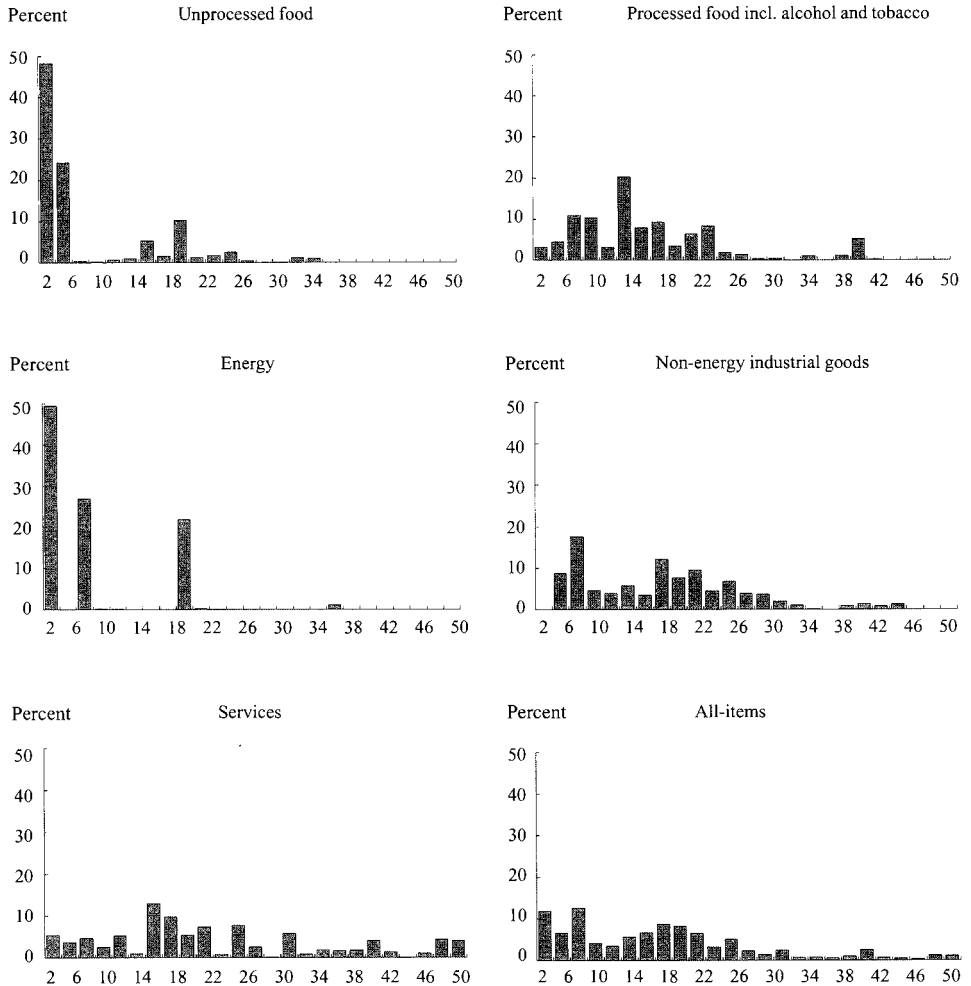


Figure 4.3. Distribution of durations.

Note: Durations are weighted at the elementary products level. The distributions are truncated at the horizon of 50 months.

even larger, namely 15.6 percent.¹⁹ A similar pattern is obtained for other countries, see below.

Price adjustments are largest for unprocessed food. This is perhaps not surprising as these are perishable²⁰ products, which are subject to recurrent supply shocks such as, e.g., changes in weather conditions, diseases affecting livestock production, etc. At the

19. Remember that our data includes all types of sales, rebates, etc. Omitting the products typically affected by sales, however, only has a minor effect on the aggregate size of price changes.

20. In spite of the recent finding of 10-year old meat by the Danish Veterinary and Food Administration, sic!

Table 4.3. Size of price changes, log-differences, monthly basis.

	Size of price ...	
	... increases	... decreases
Unprocessed foods	20.1	21.6
Processed foods	10.5	14.5
Energy	4.4	3.8
Non-energy industrial goods	15.4	19.4
Services	11.1	15.5
All-items	12.3	15.6

same time, the price elasticity of demand is probably moderate. Together these factors are conducive to large price changes. The smallest price changes are found for energy, reflecting both the gradual and moderate adjustment of fuel prices, district heating and electricity. Note, for energy price increases are on average larger than price decreases.

The distribution of size of price changes reveals an apparent heterogeneity between components, cf. Figure 4.4. As prices are raised more often than they are lowered, a greater part of the distribution lies to the right of the origin – in particular for services. The distribution for energy can be characterised as leptokurtic, with both price increases and decreases normally being less than 5 percent. For non-energy industrial goods and services the distributions are also uni-modal, but in these cases large price changes are more frequent. The distributions for unprocessed and processed food are clearly bimodal. For unprocessed food the distribution is almost symmetrical around zero, with almost no absolute price changes of less than 5 percent.

Figure 4.5 plots the size of price changes by month. The size of price changes is not influenced by seasonality to the same extent as the frequency of price changes. For non-energy industrial goods price decreases are largest during the sales in the beginning of the year and over the summer, while prices are raised notably at the beginning of spring and autumn. Prices of services are reduced significantly during the last four months of the year.

4.4. Comparison with international studies of frequency and size of price changes

The frequencies for Denmark are broadly in line with the estimates for the euro area, cf. Table 4.4. The figures for the euro area are based on a sample of 50 products, which is considered to be approximately representative of the full expenditure basket. Restricting our computations to this sample does not change the estimated frequencies for Denmark significantly. The strong increase in the frequencies for energy, when going from the full sample to the 50 product sample, is related to the fact that prices of

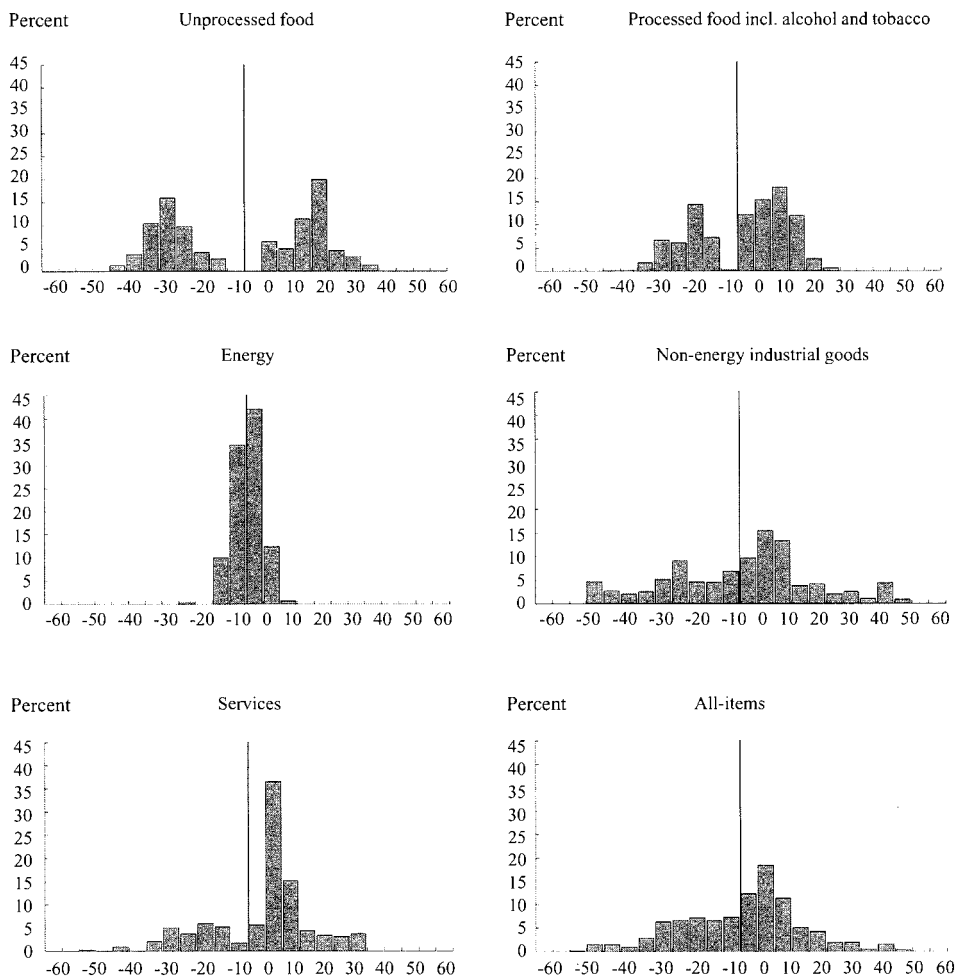


Figure 4.4. Distribution of size of price changes, log-differences.

Note: The size of price changes are weighted at the elementary products level.

district heating, which bear a large weight and change infrequently, are not included in the 50 product sample.

The size of price changes for Denmark is somewhat higher than that reported for the euro area, cf. Table 4.5. However, in the analysis of the euro area the data for some countries include sales prices and others not. Considering this, our results are in accordance with those for the euro area and in both cases the average size of price reductions is larger than the average size of price increases.

Table 4.6 draws a comparison across the Atlantic. According to both the frequency

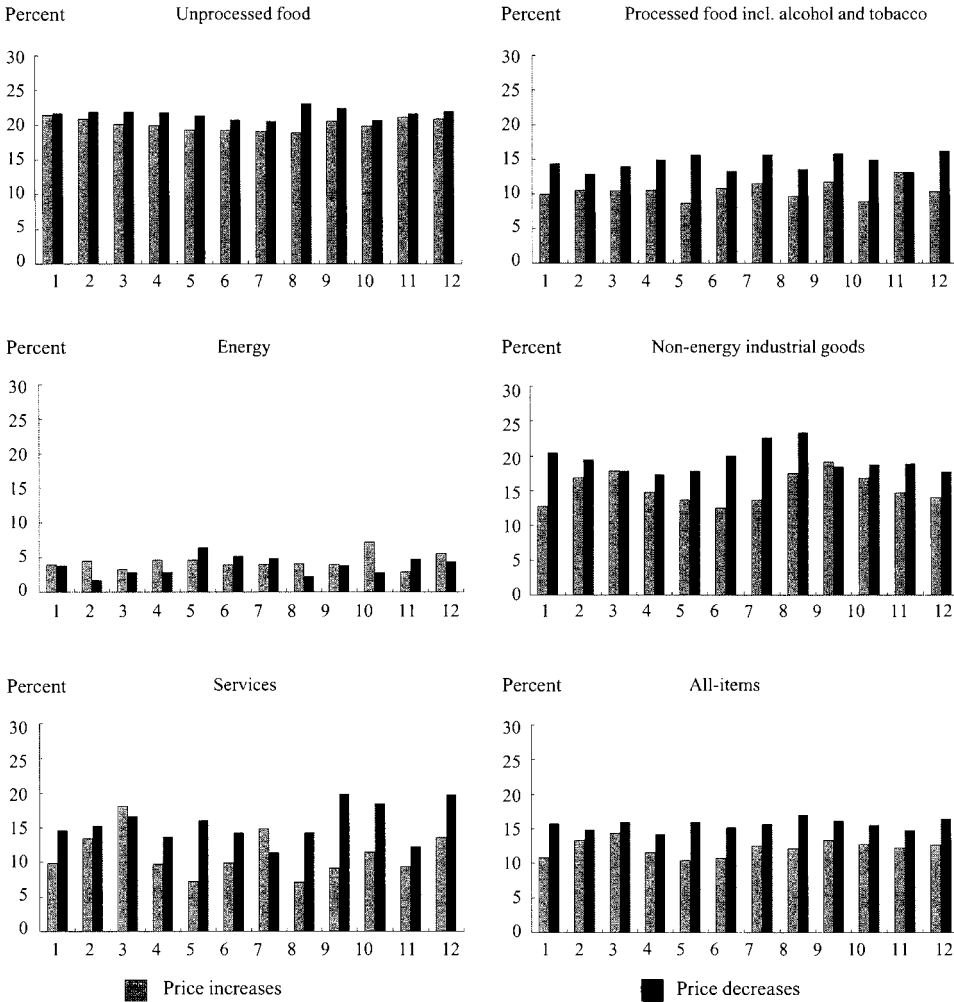


Figure 4.5. Size of price changes by month.

Note: The size of price changes are weighted at the elementary product level.

of price changes and the duration of price spells, prices appear to be more flexible in the US than in Denmark and the euro area, in particular for services. The rigid price setting of the European services sector by comparison with the US may be due to differences in the degree of regulation in the labour and product markets of the two regions.

As for frequencies, the Danish duration figures are close to those for the euro area, whereas the median duration of price spells in the US is more than twice as low as in Denmark and the euro area.

Table 4.4. Frequency of price change, monthly basis, Denmark vs. euro area, log-differences.

Frequency of price...	...increases			...decreases		
	Denmark		euro area	Denmark		euro area
	All products	50 products	50 products	All products	50 products	50 products
Unprocessed food	20.6	29.9	14.8	19.2	27.6	13.3
Processed food incl. alcohol/tobacco	6.8	9.7	7.1	4.0	7.9	5.9
Energy	28.9	51.3	42.0	23.5	43.3	35.8
Non-energy industrial goods	5.7	4.3	4.2	4.2	4.0	3.2
Services	6.7	6.3	4.2	2.2	1.0	1.0
All-items	10.2	10.4	8.3	7.1	7.1	5.9

Note: For the euro area, because of different weighting methods, the frequency of price increases and decreases does not add up to the overall frequency of price changes presented in table 4.5, see Dhyne et al. (2005).

Source: Dhyne et al. (2005) and own calculations.

Table 4.5. Size of price changes, monthly basis, Denmark vs. euro area, log-differences.

Size of price...	...increases			...decreases		
	Denmark		euro area	Denmark		euro area
	All products	50 products	50 products	All products	50 products	50 products
Unprocessed food	20.1	22.3	14.7	21.6	23.5	16.3
Processed food incl. alcohol/tobacco	10.5	13.5	6.9	14.5	16.7	8.1
Energy	4.4	3.4	3.4	3.8	3.1	2.4
Non-energy industrial goods	15.4	18.3	9.4	19.4	22.5	11.4
Services	11.1	9.0	7.3	15.1	16.4	9.7
All-items	12.3	12.6	8.2	15.6	17.6	10.0

Source: Dhyne et al. (2005) and own calculations.

The reference period in the studies for Denmark, the euro area and the US is not exactly the same, but all studies refer to periods with low inflation. Gagnon (2005) studies the price setting behaviour in Mexico over the period 1994-2004, where inflation rose from around 7 percent in 1994 to above 40 percent in 1995 before falling to around 4 percent in 2001. In periods with moderate inflation, price rigidities in Mexico lie between Europe and the US, whereas the frequency increases and becomes closer to that of the US in periods with high inflation.

Table 4.6. Selected aggregate statistics for Denmark, the Euro area and the US.

	Denmark		euro area	US
	All products	50 products	50 products	
<i>Overall frequency of price changes:</i>				
– All-items	17.3	17.5	15.1	26.1
– Services	8.9	7.3	5.6	20.7
<i>Size of price changes:</i>				
– Absolute	13.7	–	–	13.0
– Increases	12.3	12.6	8.2	–
– Decreases	15.6	17.6	10.0	–
<i>Durations:</i>				
– Mean	15.5	15.1	13.0	–
– Median	11.8	11.0	10.6	4.3
– Pseudo	5.3	5.2	6.1	3.3

Note: The average price duration is computed under the assumption of continuous time.

Source: Dhyne et al. (2005), Bils and Klenow (2004), Klenow and Kryvtsov (2005) and own calculations.

Klenow and Kryvtsov (2005) report an average absolute size of price changes of 13.3 percent for the US. Hence, whilst prices are adjusted more often in the US than in Europe, there is no marked difference in the size of price adjustments in the two areas.

4.5. Empirical hazard functions

Empirical studies of price setting behaviour often draw attention to hazard functions of price changes. Different pricing models usually have different implications for the hazard function. In the simple Calvo (1983) model of time-dependent pricing the probability that a firm gets the possibility to change its price in a given period is independent of the time elapsed since its last price adjustment, implying a constant hazard function. In the even simpler Taylor (1980) model of staggered price setting, the hazard function takes the value of zero during the duration of contracts and one at the expiration of the contract. In models of state-dependent pricing things get more complicated, but often the shape of the hazard function will be increasing as the optimal price, which is a function of the state of the economy, is likely to drift away from the current price as time passes, see e.g. the model by Dotsey, King and Wolman (1999). Furthermore, under state-dependent pricing – unlike time-dependent pricing – the hazard rates will be a function of factors such as the inflation rate.

In practice hazard functions are often found to be downward sloping, cf. section 3.3. This divergence between theory and practice stems from heterogeneity in price setting behaviour. As mentioned above, we try to avoid the aggregation bias by estimating weighted average hazard rates.

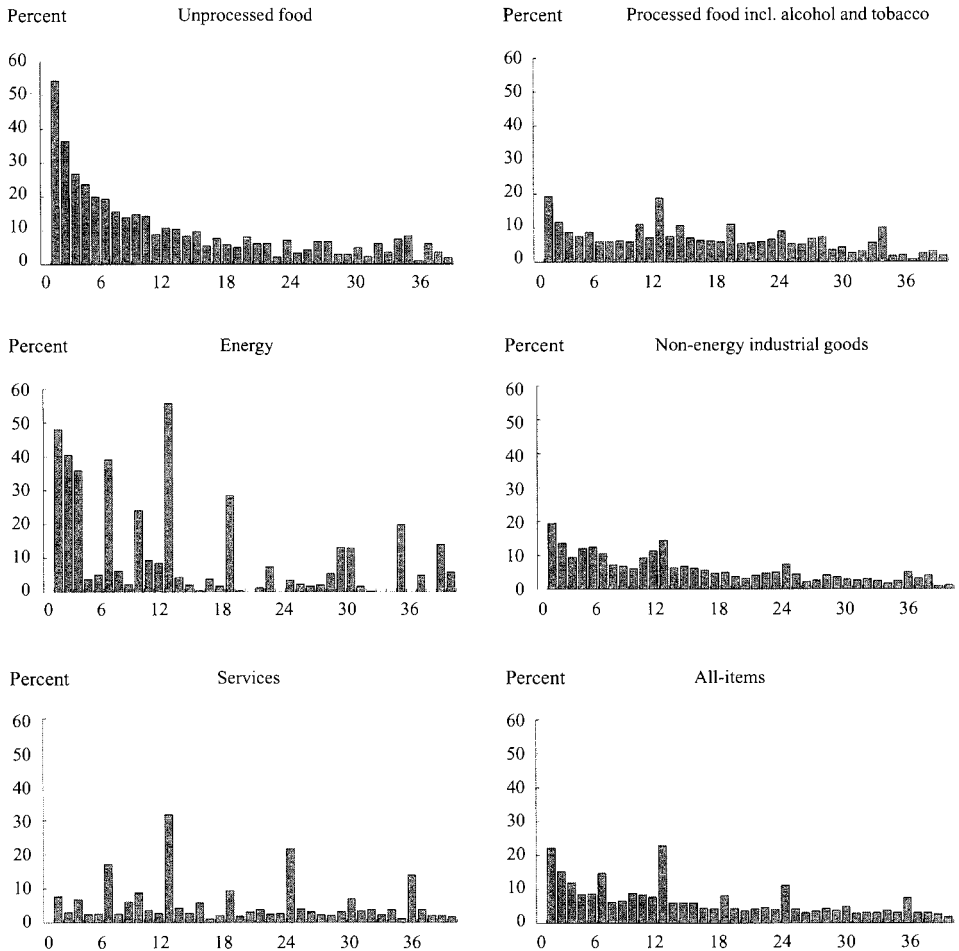


Figure 4.6. Empirical hazard functions.

Note: Hazard rates are weighted at the elementary products level. Empirical hazard rates represent the frequency of price changes as a function of the number of months elapsed since the latest price change. Hazard rates are truncated at the horizon of 40 months. Beyond that horizon the uncertainty of the hazard rates increases substantially, as the number of products where prices have still not changed decreases.

The empirical hazard functions are presented in Figure 4.6. We find that the hazard function for all-items is slightly decreasing. However, the hazard rates for very short durations, in particular one month, are significantly lower than those estimated for other countries cf. Álvarez et al. (2005). This suggests that our attempt to control for over-sampling of short price spells by computing the aggregate hazard function as a weighted sum of hazard functions at the elementary product level was fruitful, hence confirming the results in Baumgartner et al. (2005).

Some bias in favour of a decreasing hazard function might still be left even after controlling for heterogeneity at the product level, as the heterogeneity bias might also stem from aggregating across firms with different price setting rules within the same product categories, as shown by Álvarez et al. (2005). Nonetheless, for all-items the probability of a price change is estimated to be highest 12 months after the last price change. Furthermore, the aggregate hazard function also shows some pronounced spikes at durations of 6, 18, 24 and 36 months, reflecting that many retailers adjust their prices annually or semi-annually in a time-dependent manner. Similar findings are recorded in many other studies cf. e.g. the overview in Dhyne et al. (2005).

Once again the differences across components are substantial. The hazard function for unprocessed food is downward sloping since it is dominated by flexible-price goods. For energy the shape of the hazard function is rather irregular, e.g. the likelihood of a price adjustment occurring 12 months after the last price adjustment exceeding the one-month hazard rate, reflecting e.g. district heating and electricity. The hazard rates for services are generally low and do not decrease as time passes. The spikes at durations of 12, 24 and 36 months reflect that prices on e.g. insurances are adjusted annually via indexation schemes. For processed food and non-energy industrial goods the hazard functions are relatively flat.

5. Inflation against frequency and size

So far we have discovered a great amount of variation in both the frequency and size of price changes over time and across products, hence yielding support to elements of both time- and state-dependent pricing rules. In order to provide further evidence on this, in this section we take a look at the relationship between inflation and the frequency and size of price changes through some simple regressions, leaving out more thorough analyses for future study.

To get a rough impression of the covariation between the overall inflation rate and the frequency and size of price changes, we compare annual inflation rates to 12-month moving averages (MA) of the frequency, and size, of price increases and decreases, cf. Figure 5.1.

The rise in inflation through 1999 from 1 to 3 percent was visibly associated with an increase in the frequency of price increases and a decline in the frequency of price decreases. During the same year, the size of price increases remained almost constant, whereas the size of price decreases went in the »wrong« direction, namely up. From this, it appears that the rise in inflation in 1999 reflects a relatively higher share of price increases. A similar story seems to hold for the evolution of inflation from 2003 to 2005. In this case, the drop in inflation in 2003 and the subsequent rise in inflation was associated with a parallel movement in the frequency of price increases and an opposite

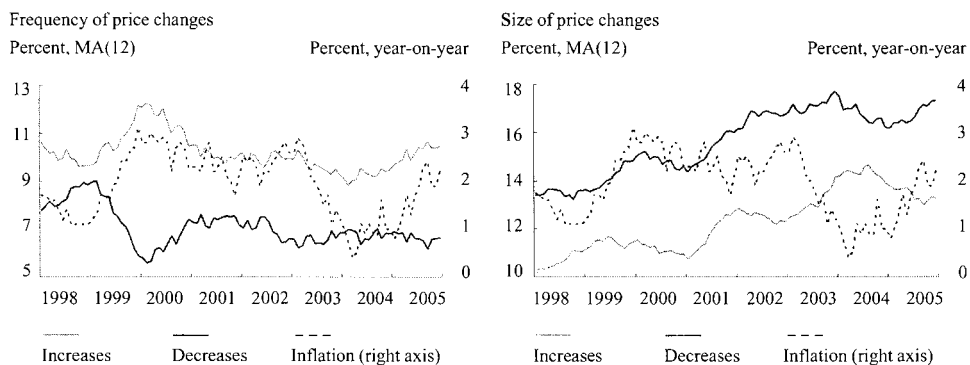


Figure 5.1. Frequency and size of price changes and inflation – all-items.

movement, although less distinct, in the frequency of price decreases. Over the same period, both the size of price increases and decreases moved in the »wrong« direction in relation to the inflation development. These observations also tend to hold at the disaggregate level, cf. Hansen and Hansen (2006).

The previous figures are crude and no more than suggestive. In order to obtain more detailed information about the driving forces of inflation, we look at scatter plots of month-on-month inflation and the frequencies, and size, of price increases and decreases, cf. Figures 5.2-5.5. The matching results from simple linear regressions are reported in Hansen and Hansen (2006).

Looking at Figures 5.2-5.3 it is clear that there is a significant correlation between the month-on-month inflation rates and the frequency of price changes. The relationship is especially strong for unprocessed food, processed food and energy. These are also the three components where fluctuations in inflation have been sizeable. This underlines that variation in the frequency of price changes are important for the evolution of consumer prices. The covariation between inflation and the frequency for services is less visible, although still statistically significant. For non-energy industrial goods there is a strong negative correlation between inflation and price decreases, likely reflecting the influence of sales.

The correlations between inflation and the size of price increases and decreases are low, and sometimes also of the »wrong« sign, cf. Figures 5.4-5.5. The strongest relationship is found for non-energy industrial goods, while for many of the other components the correlations are not statistically significant.

The fact that the frequency of price changes is correlated with inflation is consistent with state-dependent pricing strategies. Gagnon (2005) – using graphical analysis and linear regressions – also finds that inflation is more correlated with the frequency of

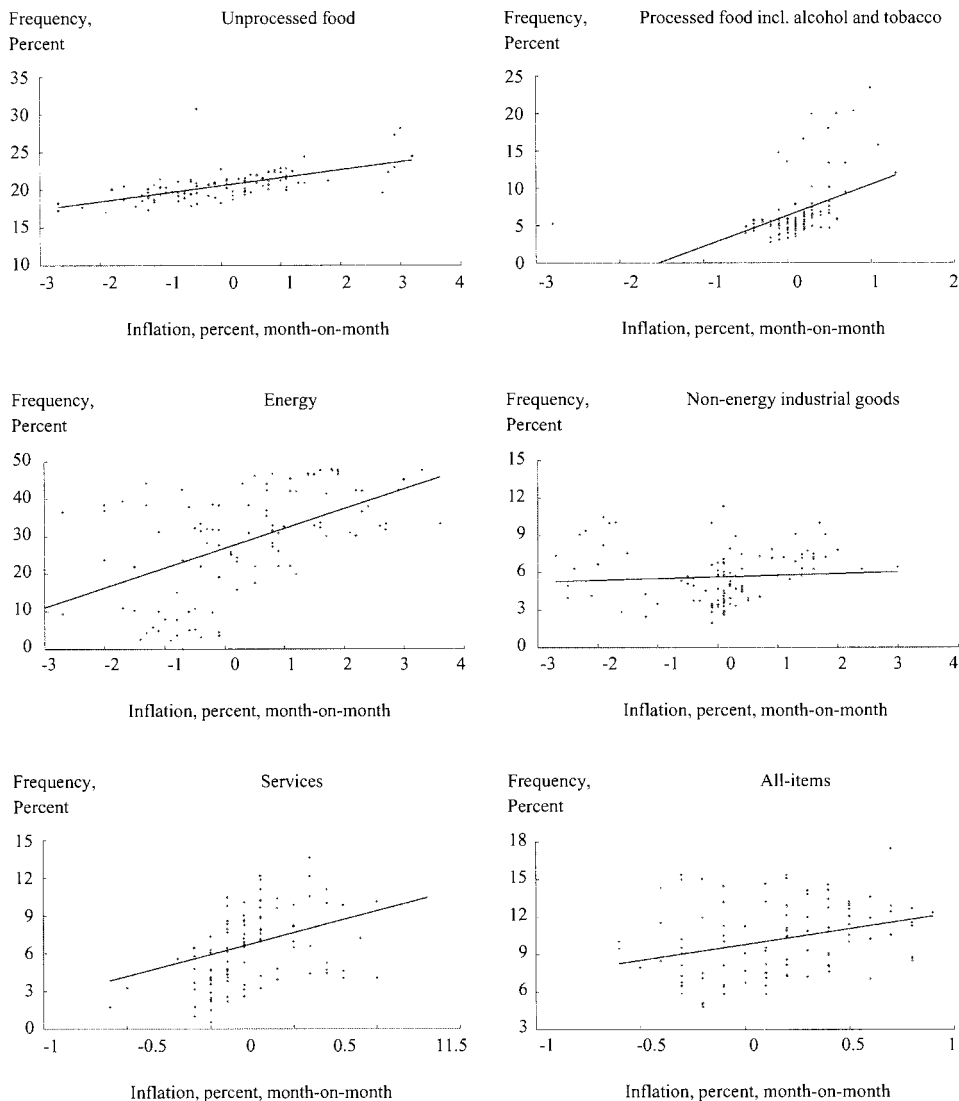


Figure 5.2. Frequency of price increases against inflation.

price increases and decreases than with the size of price increases and decreases. This holds for both high and low levels of inflation in Mexico. However, through a formal inflation variance decomposition Gagnon discovers a much larger role for movements in the size of price changes. For example, when restricted to the low-inflation period, the share of inflation variance represented by the time-dependent term reaches 83.5 percent. Considering the full sample, the time-dependent term only represents 42.5

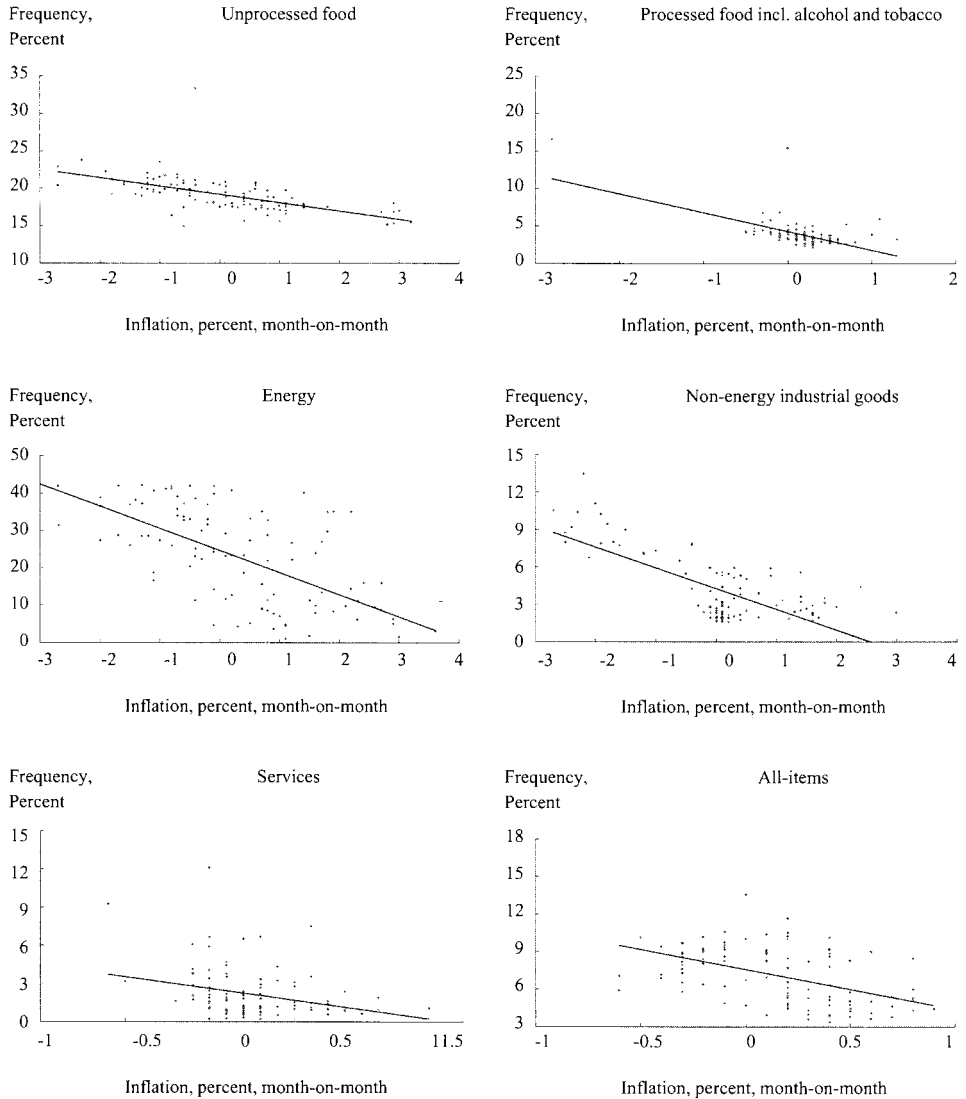


Figure 5.3. Frequency of price decreases against inflation.

percent of the inflation variance. In addition, Gagnon provides evidence of both time- and state-dependence using logit models.

Klenow and Kryvtsov (2005) also quantify the importance of fluctuations in the overall frequency and size of price changes for the variance of inflation. They find that 95 percent of the variation in inflation is accounted for by changes in the size of price changes, leaving almost no role for fluctuations in the frequency.

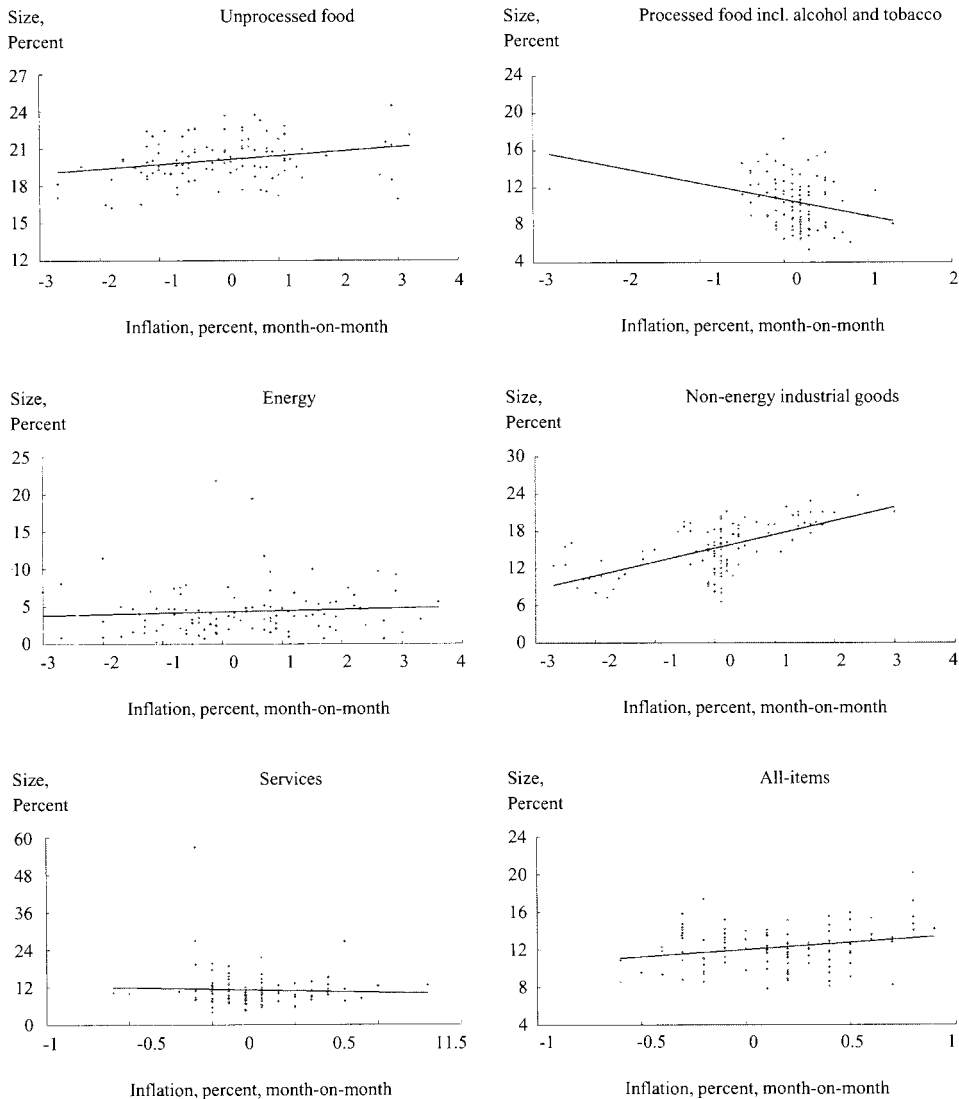


Figure 5.4. Size of price increases against inflation.

On the other hand, many of the country-studies performed within IPN find evidence of state-dependent pricing behaviour. For example, for Austria, Belgium, Germany and Portugal there exists a positive correlation between inflation and the frequency of price changes. As in our case, there seems to be stronger evidence for state-dependence at the sectoral level. Álvarez and Hernando (2004) and Veronese et al. (2005) provide evidence of statedependence using econometric models, and Aucremanne and Dhyne (2005) find evidence of state-dependent pricing for Belgium using logit models.

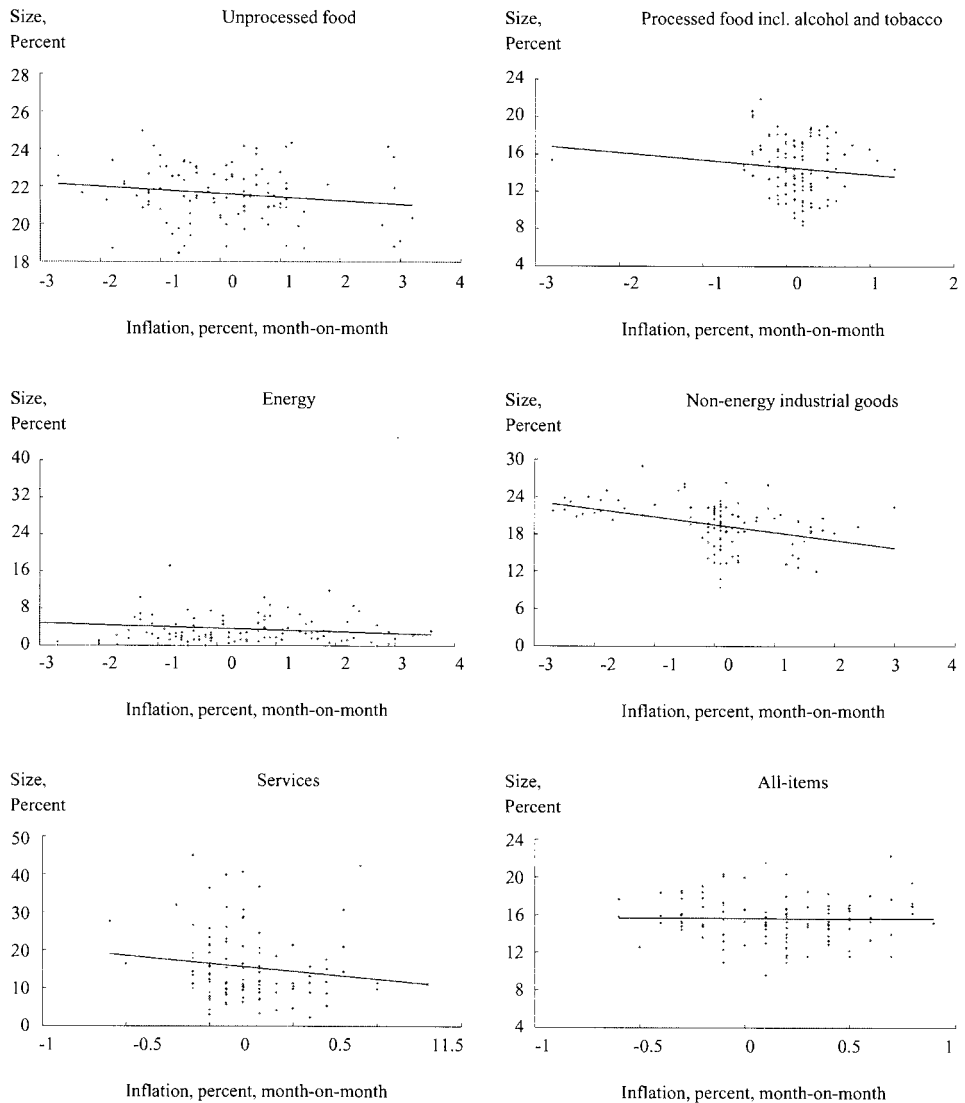


Figure 5.5. Size of price decreases against inflation.

6. Summary

In this paper we have studied the price setting behaviour in Denmark over the period 1997-2005 using CPI micro data. Through a comprehensive descriptive analysis we have derived a set of stylised facts that are broadly in line with recent evidence from euro area countries, though in some cases at odds with our a priori conjectures.

Our main observations are:

- Prices generally change rather infrequently. On average 17.3 percent of prices are adjusted from one month to the next and the average duration of price spells ranges from 4 to 5 quarters.
- There is no sign of downward price rigidity, except for services. On average 40 percent of all price adjustments are decreases.
- Pricing behaviour is very heterogeneous across sectors and products. Prices for energy and unprocessed food are most flexible, while those for services, processed food and non-energy industrial goods are much stickier.
- Price changes are sizeable compared to the overall inflation rate. The magnitude of the average price increase is 12.3 percent, while the average reduction of prices is larger, namely 15.6 percent.
- Whilst prices are adjusted more often in the US than in Denmark and the euro area, there is no marked difference regarding the size of price adjustments.
- Seasonality is an important characteristic of the frequency of price changes, but not so much of the size of price changes.
- Empirical hazard functions are downward sloping for flexible-price products and flatter for sticky-price products. However, controlling for heterogeneity at the product level in our estimations generally reduces the negative slope of the hazard function. For very short durations the hazard rates are significantly lower than the empirical hazard rates for other countries.
- Hazard functions have pronounced spikes at durations of 6, 12, 24 and 36 months, in particular for services but also for other types of products, indicating the use of indexation schemes.
- The frequency of price increases (decreases) is positively (negatively) correlated with inflation.

The results stress the importance of heterogeneity in price setting behaviour. To better understand the dynamics of aggregate inflation, this heterogeneity should be addressed when setting up the theoretical framework for microfounded macroeconomic models.

Our evidence supports the use of both time- and state-dependent pricing strategies. The application of time-dependent strategies is supported by the seasonal pattern in the frequency of price changes and the fact that the hazard functions are characterised by mass points every 6 and/or 12 months. However, the significant relationship between inflation and the frequency of price changes is in line with state-dependent pricing strategies. Likewise, the fact that changes in indirect taxation significantly affect the frequency of price changes supports the use of state-dependence. However, since our analysis is basically descriptive, interpretations should be made with care.

More structural interpretations await further econometric studies that rely on well-founded identification schemes.

Literature

- Álvarez, L. J., P. Burriel and I. Hernando. 2005. Do Decreasing Hazard Functions for Price Changes Make Any Sense? *ECB Working Paper* no. 461.
- Álvarez, L. J., and I. Hernando. 2004. Price Setting Behaviour in Spain: Stylised Facts Using Consumer Price Micro Data, *ECB Working Paper* no. 416.
- Armknrecht, P. A. and F. Maitland-Smith. 1999. Price Imputation and Other Techniques for Dealing with Missing Observations, Seasonality and Quality Change in Price Indices, *IMF Working Paper* no. 99/78, Washington, June.
- Aucremagne, L., and E. Dhyne. 2005. Time-Dependent versus State-Dependent Pricing: A Panel Data Approach to the Determinants of Belgian Consumer Price Changes, *ECB Working Paper* no. 462.
- Baudry, L., H. Le Bihan, P. Sevestre, and S. Tarriau. 2004. Price Rigidity in France: Evidence from Consumer Price Micro Data, *ECB Working Paper* no. 384.
- Baumgartner, J., E. Glatzer, F. Rumler, and A. Stiglbauer. 2005. How Frequently Do Consumer Prices Change in Austria? Evidence from Micro CPI Data, *ECB Working Paper* no. 523.
- Bils, M., and P. J. Klenow. 2004. Some Evidence on the Importance of Sticky Prices, *Journal of Political Economy*, vol. 112, no. 5, pp. 947-85.
- Calvo, G. 1983. Staggered Prices in a Utility Maximizing Framework, *Journal of Monetary Economics*, no. 12, pp. 383-398.
- Dhyne, E., L. J. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünemann, F. Rumler, and J. Vilmunen. 2005. Price setting in the Euro Area: Some Stylised Facts from Individual Consumer Price Data, *ECB Working Paper* no. 524.
- Dotsey, M., R. King, and A. Wolman. 1999. State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output, *Quarterly Journal of Economics*, vol. 114, pp. 655-90.
- Gagnon, E. 2005. Price Setting during Low and high Inflation: Evidence from Mexico, Mimeo, Northwestern University.
- Goette, L., R. Minsch, and J.-R. Tyran. 2005. Micro Evidence on the Adjustment of Sticky-Price Goods: It's How Often, not How Much, *Discussion Paper* 05-20, Department of Economics, University of Copenhagen.
- Hansen, B. W., and N. L. Hansen. 2006. Price Setting Behaviour in Denmark – A Study of CPI Micro Data 1997-2005, *Working Paper* 39, Danmarks Nationalbank.
- Hidiroglou, M.A., and J.-M. Berthelot. 1986. Statistical editing and imputation for periodic business surveys, in *Survey Methodology*, vol. 12, no. 1, pp. 73-83.
- Klenow, P. J., and O. Kryvtsov. 2005. State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation? *NBER Working Paper* no. 11043.
- Statistics Denmark. 2005. The Consumer and Net Price Index: Documentation (only in Danish).
- Taylor, J. B. 1980. Aggregate Dynamics and Staggered Contracts, *Journal of Political Economy*, vol. 88, pp. 1-23.
- Taylor, J. B. 1999. Staggered Price and Wage Setting in Macroeconomics, in: J. B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, vol. 1, Elsevier Science B.V., pp. 1009-50.
- Veronese, G., S. Fabiani, A. Gattulli, and R. Sabbatini. 2005. Consumer Price Behaviour in Italy: Evidence from Micro CPI Data, *ECB Working Paper* no. 449.

Holdninger til immigration – Er danskerne specielle?

Nikolaj Malchow-Møller

CEBR og University of Southern Denmark, E-mail: nmm.cebr@cbs.dk

Jakob Roland Munch

CEBR og University of Copenhagen, E-mail: Jacob.Roland.Munch@econ.ku.dk

Sanne Schroll

CEBR og University of Southern Denmark, E-mail: ss.cebr@cbs.dk

Jan Rose Skaksen

CEBR og Copenhagen Business School, E-mail: jrs.eco@cbs.dk

SUMMARY: This paper analyses attitudes towards immigration in the EU-15 – including the immigration following the recent enlargements of the EU – using data from the two first rounds of the European Social Survey. We explore how socioeconomic factors and perceived consequences of immigration can explain these attitudes, and analyse to which extent Danes deviate from the EU average. While attitudes appear to be driven by the same factors in Denmark and the rest of the EU-15, Danes are relatively more optimistic about the economic consequences of immigration – especially those relating to the labour market – although they are slightly more pessimistic about the fiscal consequences. Controlling for socioeconomic factors and perceived consequences of immigration therefore results in a considerable unexplained difference between the average Danish attitude and the average EU-15 attitude towards immigration.

1. Introduktion

I Danmark har der i de senere år været stor fokus på immigrationspolitikken, hvilket bl.a. har givet sig udslag i lovgivning, der har haft til formål at begrænse tilstrømningen af udlændinge, se f.eks. Velfærdskommissionen (2005a). Ifølge Boeri og Brücker (2005) var Danmark i 2005 således det EU-15 land, der havde den mest restriktive immigrationspolitik. Til sammenligning indtog Danmark i 1990 blot en femteplads.

Dette papir er en del af et fælles projekt mellem CEBR og Rockwool Fondens Forskningsenhed. Forfatterne vil gerne rette en tak til Rockwool Fonden for finansiel støtte, til Vibeke Borchsenius og Jonas Helth Lønberg for fremragende forskningsassistance samt til Torben Tranæs og to anonyme referees for gode kommentarer. Kontaktperson: Jan Rose Skaksen, tlf: +45 3815 2582.

I forbindelse med udvidelsen af EU med 10 nye østeuropæiske lande pr. 1. maj 2004 opstod der i de 15 »gamle« EU-lande en vis frygt for omfanget og konsekvenserne af immigrationen fra de nye og fattigere EU-lande. Det bevirkede, at samtlige EU-15 lande – bortset fra Sverige – indførte overgangsordninger, der havde til formål at begrænse mobiliteten af arbejdskraft fra de nye EU-lande. I denne sammenhæng er det interessant at bemærke, at Danmark, ligesom Storbritannien, Irland og Sverige, valgte en relativ liberal politik.¹

Overgangsordningerne i forbindelse med EU-udvidelsen har til formål at afbøde eventuelle negative økonomiske effekter af en stor immigration fra de nye EU-lande. Danmarks forholdsvis liberale politik i forhold til de nye EU-lande kunne indikere, at danske politikere ikke i særlig høj grad frygter de økonomiske konsekvenser af immigration, men at det i stedet er andre forhold, der forklarer Danmarks generelt restriktive immigrationspolitik i forhold til ikke-EU lande.

Imod dette taler, at Folketinget d. 5. april 2006 besluttede at fastholde overgangsordningerne i forhold til østeuropæisk arbejdskraft – den såkaldte Østaf-tale – i yderligere tre år, mens lande som Finland, Portugal og Spanien har besluttet at afvikle dem. Dette kunne trods alt indikere en vis bekymring for de økonomiske konsekvenser.

Formålet med dette papir er at belyse holdningerne til immigration – herunder den der følger af EU-udvidelsen – blandt borgerne i Danmark og de øvrige EU-15 lande, samt at analysere, hvilke faktorer der ligger til grund for disse holdninger. Hvad er de generelle holdninger til immigration, og hvordan skiller danskerne sig ud? Hvad bestemmer folks holdninger, og er de økonomisk motiverede? Hvordan skiller danskerne sig ud i denne sammenhæng? Forventer og/eller frygter vi andre konsekvenser af immigration?

I den økonomiske litteratur har der været en række studier, der har belyst betydningen af socioøkonomiske forhold for folks holdninger til immigration, se især Scheve og Slaughter (2001) og Mayda (2006). I disse undersøgelser har der været særlig fokus på betydningen af uddannelse, og langt de fleste studier finder, at de lavest uddannede er de mest skeptiske over for immigration. Denne sammenhæng er blevet fortolket som værende i overensstemmelse med, at personlige økonomiske motiver spiller en rolle for folks holdninger, idet det især er de lavest uddannede, der må forventes at blive negativt påvirket af immigration via øget konkurrence på arbejdsmarkedet.

Nyere analyser har imidlertid stillet spørgsmålstejn ved, hvorvidt de lavtuddannedes immigrationsskepsis kan tilskrives økonomiske forhold. Således finder f.eks. Dustmann og Preston (2004), at det i højere grad er racisme, der forklarer, at de lavest uddannede i Storbritannien er mere kritiske over for immigration. I Malchow-Møller m.fl.

1. Ifølge Boeri og Brücker (2005) kan EU-15 landene opdeles i fire kategorier efter hvor restriktive overgangsordninger de valgte. Danmark er her placeret i gruppen med den næstmest liberale politik.

(2006) har vi endvidere argumenteret for, at man ikke kan slutte, at de lavtuddannedes modstand mod immigration skyldes en hensyntagen til deres egen økonomiske situation, eftersom de lavtuddannede på ingen måde synes at være enige om, at øget immigration forringer deres arbejdsmarkedsvilkår. Hvis man ønsker at belyse betydningen af økonomiske forhold, er man i stedet nødt til at kigge direkte på sammenhængen mellem folks opfattelser af konsekvenserne og deres holdninger til immigration.

I denne artikel anvender vi først den traditionelle tilgang fra litteraturen, hvor vi søger at forklare individuelle holdninger til immigration ud fra en række socioøkonomiske karakteristika, herunder folks uddannelsesniveau. I forhold til den eksisterende litteratur, indrager vi en række yderligere karakteristika, som bidrager til bedre at fange de enkelte individers sårbarhed på bl.a. arbejdsmarkedet over for øget immigration. Analyserne baserer sig på data fra European Social Survey (ESS) 2002/3 og 2004/5, hvor folk bl.a. er blevet adspurgt om deres holdninger til immigration.

ESS 2002/3 rummer desuden en række spørgsmål om de opfattede konsekvenser af immigration. Dette giver os mulighed for efterfølgende at undersøge, hvilken betydning folks opfattelser af konsekvenserne har for deres holdninger til immigration og dermed en mere direkte vurdering af, hvilken rolle økonomiske overvejelser spiller for folks holdninger.

Som nævnt er det et særskilt formål med papiret at belyse og forklare eventuelle forskelle mellem danskernes og de øvrige EU-15 borgeres holdninger. Vi undersøger derfor, om forskelle i holdninger kan forklares ud fra anderledes effekter af de forklarende variable i Danmark, eller om de skyldes en anden sammensætning af socioøkonomiske karakteristika og/eller en anden opfattelse af de økonomiske konsekvenser blandt danskerne.

Vi finder, at de mest skeptiske over for immigration er de lavtuddannede, de fattige, de arbejdsløse, de ældre og folk på landet. På dette punkt adskiller danskerne sig ikke fra befolkningerne i de øvrige EU-15 lande. Endvidere er folk, der er bekymret for de økonomiske konsekvenser af immigration generelt mere modstandere af immigration. Heller ikke her er danskerne anderledes. Danskerne adskiller sig derimod ved, at forholdsvis få forbinder øget immigration med negative økonomiske konsekvenser. Kontrollerer vi for dette, kommer Danmark til at fremstå som et væsentligt mere immigrations skeptisk land.

Resten af artiklen er opbygget således: I afsnit 2 beskriver vi kort data fra ESS 2002/3 og 2004/5. I afsnit 3 sammenligner vi danskernes holdninger til immigration med holdningerne i de øvrige EU-15 lande. I afsnit 4 forsøger vi at forklare disse holdninger ud fra socioøkonomiske forklaringsfaktorer, mens vi i afsnit 5 udvider analysen til at inkludere folks opfattelser af konsekvenserne af immigration. I afsnit 6 under-

søger vi nærmere, hvordan danskernes holdninger adskiller sig fra de øvrige EU-15 borgeres, og i afsnit 7 giver vi nogle mulige bud på, hvorfor danskernes holdninger afviger fra de øvrige EU-15 borgeres. Afsnit 8 konkluderer.

2. Data

Vi anvender data fra European Social Survey (ESS) 2002/3 og 2004/5, som er spørgeskemaundersøgelser omkring holdninger, opfattelser og adfærd blandt borgerne i Europa. Endvidere rummer ESS information om en række baggrundskarakteristika som f.eks. uddannelse, arbejdsmarkedstilknytning og familiære forhold. ESS 2002/3 og 2004/5 dækker henholdsvis 22 og 26 europæiske lande.² Vi anvender imidlertid kun data fra de 15 »gamle« EU-lande, idet fokus er på disse landes holdninger til øget immigration, herunder immigration fra de nye EU-lande.

I ESS er der interviewet omkring 2.000 tilfældigt udvalgte personer over 14 år i hvert land.³ 2002/3 runden rummer endvidere et specialmodul med en række spørgsmål omkring immigration. Dette modul er unikt i den forstand, at det udover spørgsmål om holdninger til forskellige typer af immigration rummer en række spørgsmål om, hvilke karakteristika respondenterne vægter højt ved potentielle immigranter samt spørgsmål om respondentens forståelse af (især de økonomiske) konsekvenser af immigration.

Spørgsmålene er endvidere blevet designet med henblik på at sikre en ensartet opfattelse på tværs af lande. Dette er vigtigt, eftersom de europæiske lande er meget forskellige med hensyn til bl.a. kultur, sprog, immigrationshistorie og -lovgivning. For eksempel er ordet »immigrant« ikke anvendt i spørgeskemaet, idet dette ord har forskellige konnotationer i forskellige lande. Spørgeskemaet anvender i stedet formuleringer som »mennesker der flytter hertil«.⁴ For yderligere detaljer omkring design af spørgeskemaet henvises til Card, Dustmann og Preston (2005).

I den empiriske analyse anvender vi en række forskellige variable til at fange respondenternes holdninger til immigration. Respondenterne er blevet adspurgt separat om deres holdninger til immigration fra rige europæiske lande, fattige europæiske lande, rige ikke-europæiske lande og fattige ikke-europæiske lande. Til hvert spørgsmål er der fire mulige svar: »Vi bør tillade mange at flytte hertil«, »En del«, »Nogle få« eller »Ingen«. Dette giver os mulighed for at konstruere fire basale holdningsvariable, AT_{rig_euro} , AT_{fattig_euro} , $AT_{rig_noneuro}$ og $AT_{fattig_noneuro}$, som antager værdier mellem 1 og 4, hvor værdien 1 svarer til kategorien »Ingen«.

2. Yderligere information om ESS kan findes på www.europeansocialsurvey.org.

3. I nogle af landene har der ikke været tale om simpel tilfældig udvælgelse, idet f.eks. personer fra husholdninger af forskellig størrelse ikke har haft samme udvælgelsesandsynlighed. Vi korrigerer for denne ikke-tilfældige udvælgelse i vores estimationer.

4. Vi vil dog for nemheds skyld anvende ordet »immigrant« i den resterende del af artiklen.

Tabel 2.1. Deskriptiv statistik (2002/3 og 2004/5).

Variable	EU-15				Danmark			
	Gns.	Min	Max	# obs	Gns.	Min	Max	# obs
<i>alder</i>	47,231	15	102	56548	48,545	15	102	2645
<i>mand</i>	0,467	0	1	56799	0,503	0	1	2645
<i>by</i>	0,308	0	1	56662	0,370	0	1	2633
<i>immigrant</i>	0,161	0	1	56820	0,097	0	1	2643
<i>primær</i>	0,420	0	1	54980	0,215	0	1	2638
<i>sekundær</i>	0,381	0	1	54980	0,497	0	1	2638
<i>tertiær</i>	0,196	0	1	54980	0,229	0	1	2638
<i>selvstændig</i>	0,085	0	1	56785	0,072	0	1	2635
<i>arbejdsløs</i>	0,028	0	1	56879	0,040	0	1	2649
<i>overførsel</i>	0,119	0	1	56785	0,143	0	1	2635
<i>fattig</i>	0,202	0	1	51452	0,250	0	1	2476
<i>venstre</i>	0,285	0	1	53571	0,225	0	1	2556
<i>højre</i>	0,393	0	1	53571	0,486	0	1	2556

Anm.: Tallene for EU-15 er vægtet med landestørrelser.

Derudover anvender vi svarene til en række spørgsmål om de opfattede konsekvenser af immigration. Respondenterne er således blevet adspurgt om, hvorvidt de er enige i en række udsagn, herunder om gennemsnitslønninger påvirkes negativt af immigration, og om immigration økonomisk rammer de fattige mere end de rige. Ud fra disse spørgsmål har vi defineret en række binære variable, hvor værdien 1 angiver, at respondenterne er »Enig« eller »Meget enig« i et givet udsagn.

Endelig anvender vi en række baggrundsvARIABLE. Der er information om *alder* (i år), *køn* og om individet bor i en *by*, mens *immigrant* angiver om personen er 1. eller 2. generations indvandrere. Desuden er der oplysning om uddannelsesniveaet. *Primær*, *sekundær* og *tertiær* er således dummyvariable, der angiver respondentens uddannelsesniveau. *Selvstændig* er en dummyvariabel, der antager værdien 1, hvis personen er selvstændig, mens *arbejdsløs* angiver om respondenterne (eller partneren) er ramt af arbejdsløshed. Variablen *overførsel* angiver om personen er pensioneret, permanent syg eller handikappet, og derfor sandsynligvis modtager offentlige overførsler. Vi inkluderer også en dummy, *fattig*, der tager værdien 1, hvis husholdningen tilhører den nederste kvartil i indkomstfordelingen. Endelig er der to dummyvariable, *venstre* og *højre*, der angiver respondentens placering i det politiske spektrum.

Deskriptive statistikker for disse variable er angivet i tabel 2.1. Det totale antal observationer for de 15 lande er ca. 30.000, men på grund af manglende observationer for nogle variable vil de efterfølgende analyser være baseret på lidt færre observationer. Det fremgår af tabel 2.1, at i forhold til den gennemsnitlige EU-15 borger, er de adspurgte danskere i gennemsnit lidt ældre, bedre uddannede og mere højreorienterede.

Tabel 3.1. Holdningsvariable, landegennemsnit.

	H1	H2	H3, 2002	H3, 2004	H3, %-ændr.
Belgien	0,3955	0,1702	0,4505	0,5107	0,1337
Danmark	0,4581	0,2674	0,5602	0,6080	0,0853
Finland	0,5837	0,1928	0,6422	0,6839	0,0650
Frankrig	0,4416	0,1492	0,5126	0,5497	0,0724
Grækenland	0,8445	0,2508	0,8620	0,8159	-0,0535
Holland	0,4203	0,1052	0,4493	0,4853	0,0803
Irland	0,3252	0,1386	0,3746	0,3737	-0,0023
Italien	0,3592	0,1856	0,3873	n.a.	n.a.
Luxemborg	0,5121	0,1047	0,5573	0,5386	-0,0335
Portugal	0,6118	0,1189	0,6299	0,7012	0,1133
Spanien	0,5112	0,1267	0,5327	0,4663	-0,1246
Storbritanien	0,4809	0,1678	0,5265	0,5173	-0,0175
Sverige	0,1385	0,0845	0,1599	0,1959	0,2252
Tyskland	0,3899	0,1777	0,4393	0,5686	0,2942
Østrig	0,6461	0,1554	0,6814	0,5283	-0,2246
EU-15	0,4335	0,1646	0,4776	0,5307	0,1111

Anm.: Tallene for EU-15 er vægtet med landestørrelser.

3. Holdninger til immigration i EU

Vi vil i det følgende fokusere på tre forskellige mål for holdninger til immigration i form af binære variable, som vi danner ud fra de fire basale holdningsvariable fra sektion 2. Det første holdningsmål, H_1 , angiver modstand mod immigration fra fattige europæiske lande, som vi vil tillade os at fortolke som de nye EU-lande. H_1 antager værdien 1, hvis respondenterne svarede »Ingen« eller »Nogle få« på spørgsmålet om, hvor mange indvandrere fra fattige europæiske lande vi bør tillade at flytte hertil, dvs. hvis $AT_{fattig_euro} \leq 2$.

Det andet mål, H_2 , angiver en relativ modstand mod immigration fra fattige europæiske lande i forhold til immigration fra rige europæiske lande.

H_2 antager således værdien 1, hvis $AT_{fattig_euro} < AT_{rig_euro}$. Dette mål er medtaget, da det er et mål for, hvad folk mener om immigration fra de nye østeuropæiske EU-lande i forhold til immigration fra de gamle EU-15 lande. Dette relative mål angiver således, i hvor høj grad folk bifalder, at der er restriktioner forbundet med immigration fra de nye EU-lande.

Det tredje mål, H_3 , angiver modstand mod immigration fra fattige ikke-europæiske lande og antager værdien 1, hvis $AT_{fattig_noneuro} \leq 2$. Dette mål kan dannes både for ESS 2002/3 og ESS 2004/5, hvorimod de to øvrige mål kun kan dannes for ESS 2002/3. Dvs. dette mål kan benyttes til at belyse udviklingen over tid i holdningerne til immigration.

Tabel 3.1 angiver de gennemsnitlige værdier for de tre mål i de 15 EU-lande. Det fremgår, at for EU-15 som helhed er ca. 43% af befolkningen imod yderligere immigration fra de nye EU-lande. I forhold til dette gennemsnit for EU-15 landene er en lidt større andel af danskerne skeptiske, nemlig knapt 46%.⁵

Tallene dækker tilsyneladende over en generel modstand mod immigration i EU-15, idet kun 16% er mindre skeptiske over for immigration fra rige europæiske lande end fra fattige, jf. kolonnen for H_2 . Vi kan groft sagt fortolke dette som, at 16% bifalder, at der gælder skrapere regler for indvandring fra de nye relativt fattige EU-lande end fra de gamle »rige« EU-lande. Danmark skiller sig dog ud på dette punkt, idet hele 27% angiver en relativt større modstand mod immigration fra de nye EU-lande. Det skal bemærkes, at dette dækker over, at danskerne er langt mere positive over for immigration fra rige europæiske lande end tilfældet er i de andre EU-lande. Umiddelbart harmonerer dette ikke særlig godt med Danmarks relativt liberale overgangsordninger.

Det er også interessant at bemærke, at for så vidt angår immigration fra fattige ikke-europæiske lande, H_3 , er danskerne noget mere skeptiske over for immigration, end hvad tilfældet er i gennemsnit for EU-15. Dette gælder både i 2002/3 og 2004/5. Desuden er danskerne i 2004/5 mere skeptiske over for immigration fra fattige ikke-europæiske lande end i 2002/3. Stigningen i andelen af immigrations skeptikere svarer dog stort set til stigningen i hele EU-15. Det er også tydeligt, at danskerne langt fra er de mest skeptiske over for immigration, idet f.eks. grækerne, finnerne og portugiserne overgår os på dette punkt. Det er også bemærkelsesværdigt, at svenskere er langt mere positive over for immigration end alle andre.

Samlet set udviser danskerne således ikke meget større immigrations skepsis end de øvrige EU-15 borgere, men vi skelner tilsyneladende skarpere mellem immigration fra fattige og rige lande, hvilket umiddelbart kunne indikere, at vi i højere grad går op i økonomiske faktorer. Der kan dog være mange andre grunde til disse forskelle. I det følgende vil vi undersøge hvilke forklaringsfaktorer, der ligger bag holdningerne til immigration.

4. Socioøkonomiske forklaringer

En række studier forsøger at forklare holdninger til immigration ud fra socioøkonomiske karakteristika, som f.eks. alder, køn, geografi, etnisk oprindelse, politisk orientering og uddannelse. I litteraturen har der specielt været meget fokus på uddannelses betydning. Flere studier finder således, at det typisk er de lavest uddannede, der er de mest skeptiske over for immigration. Denne sammenhæng fortolkes ofte som værende

5. I forhold til en undersøgelse af danskernes, svenskernes, tyskerne og briterne holdninger til immigration foranlediget af Rockwool Fondens Forskningsenhed (RFF) i 2002 placeres Danmark i ESS undersøgelsen som et noget mere skeptisk land, se Nielsen (2004). I RFF undersøgelsen er briterne og tyskerne således mere skeptiske end danskerne.

i overensstemmelse med, at en frygt for egne økonomiske forhold påvirker ens holdninger, idet immigranterne i højere grad vil konkurrere om job med de lavtuddannede end med de højtuddannede, se f.eks. Scheve og Slaughter (2001) og Mayda (2006).

Vi vil i dette afsnit lægge os tæt op ad disse studier og søge at forklare holdninger til immigration ud fra en række socioøkonomiske karakteristika. Udover de variable, der inkluderes i den gængse litteratur, har vi imidlertid tilføjet variable omkring arbejdsløse, selvstændige, fattige og modtagere af overførselsindkomster, idet disse grupper må forventes at blive særskilt påvirket af immigration. De arbejdsløse må således ligesom de lavtuddannede formodes at få deres muligheder på arbejdsmarkedet forringet af øget immigration, hvorimod selvstændige (arbejdsgivere) omvendt kan nyde gavn af den ekstra arbejdskraft. Endelig kan modtagere af overførselsindkomster og fattige husholdninger forventes at være mest udsat for forringelser i den offentlige service, hvis immigranter lægger beslag på offentlige udgifter.

Vi undersøger i det følgende, hvorledes disse forskellige individuelle karakteristika påvirker sandsynligheden for at være imod yderligere immigration ved at estimere probit-modeller for de tre holdningsvariable, H_1 , H_2 og H_3 .

Resultaterne for H_1 fremgår af den første kolonne i tabel 4.1, hvor det er de marginale effekter, der er rapporteret.⁶ Således fremgår det eksempelvis, at individer med tertiær uddannelse har 15 procentpoint lavere sandsynlighed for at være imod immigration fra fattige europæiske lande end individer med en sekundær uddannelse, mens individer med kun en primær uddannelse omvendt har knapt 9 procentpoint større chance for at være imod.⁷ Det fremgår desuden, at de mest skeptiske er de ældre, folk på landet, de lavtuddannede, de arbejdsløse, de fattige og de højreorienterede. De positive er omvendt de højtuddannede, de venstreorienterede og immigranterne.

Desuden viser landedummen for Danmark, at det at være dansker i sig selv øger sandsynligheden for at være imod immigration med ca. 6 procentpoint i forhold til det at være tysker (den udeladte kategori). Vi forsøgte endvidere i denne og de to øvrige regressioner i tabel 4.1 at estimere specifikke danske effekter ved at inkludere krydsled mellem landedummen for Danmark og alle de øvrige variable. Der var imidlertid ingen af disse krydsled, der indgik i modellerne med signifikante fortegn, og de er derfor ikke rapporteret i tabellen. Dvs. der er ikke noget, der tyder på, at det at være f.eks. lavtuddannet i Danmark har en effekt på ens holdning til immigration, der adskiller sig fra effekten i EU-15 som helhed.

6. De marginale effekter angiver ændringen i sandsynligheden for at være imod yderligere immigration ved en marginal ændring i en given forklarende variabel beregnet for en person med gennemsnitlige karakteristika.

7. I ovenstående regressioner er »sekundær« den udeladte uddannelseskategori, og derfor skal effekterne af henholdsvis tertiær og primær uddannelse ses i forhold til personer, der har en sekundær uddannelse.

Tabel 4.1. Probitregressioner, socioøkonomiske faktorer.

	H_1		H_2		H_3	
	Marg.eff.	std.afv.	Marg.eff.	std.afv.	Marg.eff.	std.afv.
alder	0,00346	(0.00127)***	0,00054	(0.00092)	0,00157	(0.00130)
alder^2	-0,00001	(0.00001)	0,00001	(0.00001)	0,00001	(0.00001)
primær	0,08842	(0.00945)***	-0,00311	(0.00680)	0,09683	(0.00957)***
tertiær	-0,14975	(0.00956)***	-0,04734	(0.00668)***	-0,13922	(0.01023)***
venstre	-0,12058	(0.00866)***	-0,04800	(0.00621)***	-0,11125	(0.00906)***
højre	0,00439	(0.00921)	0,02347	(0.00674)***	0,02226	(0.00949)**
selvstændig	-0,00985	(0.01381)	0,01651	(0.01014)	-0,01847	(0.01412)
mand	0,00366	(0.00749)	0,02961	(0.00540)***	-0,00074	(0.00767)
by	-0,01328	(0.00820)	0,01380	(0.00601)**	-0,04021	(0.00842)***
immigrant	-0,10800	(0.01031)***	0,01648	(0.00817)**	-0,09361	(0.01109)***
arbejdsløs	0,10945	(0.01575)***	0,05631	(0.01284)***	0,09464	(0.01538)***
fattig	0,03030	(0.00974)***	-0,02500	(0.00659)***	0,02148	(0.00993)**
overførsel	0,00345	(0.01288)	-0,00437	(0.00908)	0,00707	(0.01325)
Østrig	0,25964	(0.02819)***	-0,00383	(0.02062)	0,23066	(0.02612)***
Belgien	-0,01554	(0.02352)	-0,02014	(0.01555)	-0,01688	(0.02416)
Danmark	0,06157	(0.02990)**	0,06504	(0.02370)***	0,11306	(0.02894)***
Spanien	0,10389	(0.01676)***	-0,02866	(0.01079)***	0,05902	(0.01670)***
Finland	0,19337	(0.02958)***	0,01088	(0.02141)	0,18718	(0.02767)***
Frankrig	0,04935	(0.01231)***	-0,02737	(0.00810)***	0,04538	(0.01230)***
UK	0,05131	(0.01236)***	-0,02561	(0.00812)***	0,03208	(0.01240)***
Grækenland	0,44109	(0.01941)***	0,03844	(0.01927)**	0,39703	(0.01631)***
Irland	-0,11036	(0.03374)***	-0,04687	(0.02183)**	-0,11837	(0.03671)***
Italien	-0,12359	(0.01303)***	-0,01587	(0.00964)*	-0,12763	(0.01396)***
Luxembourg	0,15654	(0.13183)	-0,06004	(0.07400)	0,12853	(0.12702)
Holland	0,02115	(0.01840)	-0,07927	(0.00960)***	-0,01494	(0.01865)
Portugal	0,17612	(0.02683)***	-0,02756	(0.01707)	0,13125	(0.02607)***
Sverige	-0,26713	(0.01685)***	-0,08993	(0.01123)***	-0,32512	(0.01913)***
ESS2					0,03633	(0.04233)
primær x ESS2					-0,06811	(0.01316)***
venstre x ESS2					0,02175	(0.01274)*
højre x ESS2					0,02858	(0.01265)**
arbejdsløs x ESS2					-0,09796	(0.03731)***
Østrig x ESS2					-0,26624	(0.03041)***
Spanien x ESS2					-0,13903	(0.01951)***
Frankrig x ESS2					-0,03616	(0.01597)**
Grækenland x ESS2					-0,17201	(0.03508)***
Irland x ESS2					-0,09286	(0.04866)*
Holland x ESS2					-0,04706	(0.02447)*
# obs	19251		19066		41846	
Pseudo R ²	0,0874		0,0279		0,0787	
Log likelihood	-11890,46		-8246,99		-26723,21	

Anm.: ESS2 er en dummy for ESS 2004/5. ***, ** og * angiver signifikans på henholdsvis 1%-, 5%- og 10%-niveau.

I den tredje kolonne har vi undersøgt holdningen til immigration fra fattige ikke-europæiske lande, H_3 . For at belyse om effekterne ændrer sig over tid, har vi her medtaget observationer fra begge spørgeskemaundersøgelser og inkluderet en indikatorvariabel for, om en observation stammer fra 2004/5 runden samt krydsled mellem denne variabel og alle de oprindelige forklarende variable.⁸ Grundlæggende er resultaterne som for H_1 , men vi ser dog, at de lavtuddannede og de arbejdsløse fra 2002/3 til 2004/5 er blevet relativt mindre skeptiske, mens de højreorienterede omvendt er blevet signifikant mere skeptiske. Det er også værd at bemærke, at danskernes større skepsis over for immigration i forhold til andre er mere udtalt for immigration fra fattige ikke-europæiske lande end fattige europæiske lande.

For de absolutte holdningsmål, H_1 og H_3 , er det således værd at bide mærke i, at højtuddannede og venstreorienterede personer generelt er mere positive over for immigration, mens arbejdsløse, fattige og lavtuddannede er mere skeptiske end andre. Dette synes fuldt i overensstemmelse med, at de grupper, der er særligt truede økonomisk af immigration, er mere skeptiske. Lidt overraskende er det måske, at folk der modtager overførselsindkomster tilsyneladende ikke er mere skeptiske end andre.

Som sagt er det resultat, at uddannelse er negativt korreleret med de absolutte holdningsmål ofte blevet fortolket således, at folks holdninger til immigration er delvist bestemt af personlige økonomiske motiver. I det omfang, immigranter fra fattige europæiske lande har et lavere uddannelsesnivea end immigranter fra rige europæiske lande, vil man dog mere direkte kunne teste dette ved hjælp af det relative mål, H_2 . Det skyldes, at de lavtuddannede i særlig grad burde være imod immigration af lavtuddannet arbejdskraft – se Malchow-Møller mfl. (2006) for en mere detaljeret diskussion heraf. Resultaterne for det relative mål, H_2 , viser dog et mere broget billede, idet individer med en primær uddannelse her ikke adskiller sig signifikant fra individer med en sekundær uddannelse, selvom begge grupper er signifikant mere skeptiske end de højtuddannede.

Resultaterne for det relative holdningsmål afviger også fra de to øvrige mål på en række andre punkter. Som sagt siger det relative holdningsmål noget om, hvorvidt folk mener, at der skal gælde særlige regler for immigration fra de nye EU-lande. Resultaterne viser, at dette især gør sig gældende for de højreorienterede, mændene, folk i byerne og de arbejdsløse, hvorimod de venstreorienterede, de højtuddannede, og – måske noget overraskende – de fattige trækker i den modsatte retning.

Endelig skal det bemærkes, at der også for det relative mål er signifikante forskelle i holdninger på tværs af lande, selv når vi kontrollerer for de nævnte socioøkonomiske karakteristika, jf. de mange koefficienter for ladedummies, der er signifikant forskel-

8. I tabellen er der af pladshensyn kun rapporteret de signifikante koefficientestimer.

lige fra nul. Danmark skiller sig med den største koefficient tydeligt ud på dette punkt. Vi vender tilbage til dette i afsnit 6.

5. Andre forklaringer

Nyere studier har stillet spørgsmålstegn ved, hvorvidt det er økonomiske årsager, der gør, at lavtuddannede er mere skeptiske over for immigration end højtuddannede, se f.eks. Dustmann og Preston (2004) og Hainmueller og Hiscox (2005). I en analyse af danskernes holdning til immigration, finder Andersen (2002) endvidere, at selvom det især er uddannelse, der er afgørende for danskernes holdning til immigration, er der iøvrigt ikke tegn på, at de økonomiske konsekvenser af immigration spiller en væsentlig rolle.

I Malchow-Møller mfl. (2006) har vi argumenteret for, at det at koble uddannelse og holdninger til immigration ikke er den rette vej til at belyse betydningen af personlige økonomiske motiver. Det skyldes, at folk har vidt forskellige opfattelser af de økonomiske konsekvenser af immigration. At især lavtuddannede er modstandere af yderligere immigration kan således ikke tilskrives en frygt for forringelser af privatøkonomien, hvis de lavtuddannede ikke anser øget immigration som en trussel mod deres muligheder på arbejdsmarkedet. Som konsekvens heraf er det svært ud fra ovenstående analyse at drage konklusioner om, hvorvidt økonomiske overvejelser spiller en rolle for folks holdninger.

I Malchow-Møller mfl. (2006) valgte vi at løse dette problem ved at betinge på folks opfattelser af de økonomiske konsekvenser af immigration i de empiriske analyser. I nærværende papir er fokus ikke på betydningen af personlige økonomiske motiver, men derimod mere generelt på om økonomiske overvejelser kan forklare folks holdninger til immigration. I dette tilfælde er det derfor mere oplagt at inkludere opfattelser af konsekvenser direkte blandt de forklarende variable.

ESS data adskiller sig fra andre datasæt ved, at der faktisk findes en række variable for de forventninger, folk har til konsekvenserne af immigration – både i forhold til økonomien og i forhold til kriminalitet og kultur. Mere konkret er folk blevet adspurgt om i hvilken grad de er enige i følgende udsagn:⁹

1. Gennemsnitslønnen bliver generelt trukket ned af de mennesker, der kommer for at bo og arbejde her (*lavere lønninger*)
2. Mener du, at mennesker, der flytter hertil, generelt tager arbejde fra danskere eller generelt bidrager til at skabe nye job? (*færre jobs*)
3. De mennesker, der kommer for at bo og arbejde her, skader generelt de fattiges økonomiske muligheder mere end de riges (*rammer fattige*)

9. Parenteserne efter hvert spørgsmål angiver navnene på de dummyvariable, vi har dannet ud fra spørgsmålene.

Tabel 5.1. *Opfattede konsekvenser af immigration.*

	Gennemsnits- lønnen bliver trukket ned ⁽¹⁾ (lavere løn- ninger)		Immigranter tager job fra de indfødte ⁽²⁾ (færre jobs)		Immigration går økono- misk mere ud over de fattige ⁽¹⁾ (rammer fattige)		Immigranter modtager mere end de yder i skat/ service ⁽²⁾ (belaster staten)		Et lands kul- turliv under- graves af immigra- tion ⁽²⁾ (undergra- ver kultur)		Immigranter forværrer kriminali- tetsproble- merne ⁽²⁾ (forværrer kriminalitet)	
	DK	EU-15	DK	EU-15	DK	EU-15	DK	EU-15	DK	EU-15	DK	EU-15
Enig	20,29	36,68	14,06	36,57	35,4	48,03	52,81	48,33	23,71	24,84	69,8	71,47
Neutral	17,94	23,67	42,9	34,6	16,73	22,55	26,91	28,92	25,94	20,83	21,36	20,09
Uenig	61,76	39,64	43,04	28,83	47,87	29,42	20,29	22,75	50,35	54,33	8,84	8,43
Ialt	100	100	100	100	100	100	100	100	100	100	100	100

(1) Kodet som 1 hvis svaret er »helt enig« eller »enig«.

(2) Kodet som 1 hvis svaret er fra 0-4 på en skala fra 0-10, hvor »5« er den neutrale kategori.

- De fleste mennesker, der flytter hertil, arbejder og betaler skat. De bruger også social- og sundhedssystemet. Tror du, at mennesker, der flytter hertil, alt i alt modtager mere, end de yder? (*belaster staten*)
- Mener du, at Danmarks kultur generelt undergraves (eller beriges) af, at mennesker fra andre lande flytter hertil? (*undergraver kultur*)
- Bliver Danmarks kriminalitet større (eller mindre) af, at mennesker fra andre lande flytter hertil? (*forværrer kriminalitet*)

Det fremgår meget klart af tabel 5.1, at de opfattede konsekvenser af immigration varierer kraftigt blandt de adspurgte individer – både blandt danskerne og i EU-15 som helhed. Eksempelvis er der 37% i hele datamaterialet, der er enige i, at det gennemsnitlige lønniveau falder som følge af immigration, mens ligeså mange er uenige i dette.

Danskerne opfattelser af konsekvenserne af immigration afviger fra det europæiske gennemsnit på en række områder. Hvad angår vurderingerne af de økonomiske konsekvenser af immigration, ser vi, at når det drejer sig om arbejdsmarkedskonsekvenserne af immigration, dvs. de tre første spørgsmål i listen ovenfor, er danskerne nogle af de mindst bekymrede. Det er iøvrigt en vurdering, som danskerne deler med borgerne i de øvrige skandinaviske lande. Der, hvor danskerne er mest bekymrede – også vurderet i forhold til, hvad der gælder i de andre EU-15 lande – er med hensyn til effekten på skatter og/eller offentlig service. Dvs. danskerne frygter, at immigration underminerer velfærdsstaten.

For at vurdere, hvorvidt opfattelsen af potentielle problemer ved immigration på-

Tabel 5.2. Probitregressioner, fuld model.

	H_1		H_2		H_3	
	Marg.eff.	std.afv.	Marg.eff.	std.afv.	Marg.eff.	std.afv.
alder	0,00051	(0.00142)	0,00002	(0.00097)	-0,00152	(0.00147)
alder^2	0,00001	(0.00001)	0,00001	(0.00001)	0,00004	(0.00002)**
primær	0,04511	(0.01035)***	-0,01585	(0.00700)**	0,05193	(0.01060)***
tertiær	-0,08885	(0.01082)***	-0,02490	(0.00731)***	-0,06711	(0.01129)***
venstre	-0,10849	(0.00952)***	-0,04250	(0.00650)***	-0,09390	(0.00993)***
højre	-0,02005	(0.00998)	0,02048	(0.00696)***	-0,00431	(0.01037)
selvstændig	-0,00481	(0.01491)	0,01152	(0.01030)	-0,01599	(0.01518)
mand	-0,01015	(0.00818)	0,02478	(0.00559)***	-0,01115	(0.00840)
by	-0,00761	(0.00895)	0,01870	(0.00625)***	-0,03676	(0.00918)***
immigrant	-0,06196	(0.01167)***	0,03678	(0.00896)***	-0,04480	(0.01216)***
arbejdsløs	0,07486	(0.01694)***	0,05010	(0.01291)***	0,05441	(0.01716)***
fattig	0,02961	(0.01079)***	-0,02291	(0.00691)***	0,01957	(0.01105)*
overførsel	-0,00665	(0.01428)	-0,00583	(0.00945)	-0,00806	(0.01482)
undergraver kultur	0,20063	(0.01016)***	0,05145	(0.00742)***	0,22190	(0.01014)***
forværrer kriminalitet	0,09349	(0.00922)***	0,03401	(0.00639)***	0,10888	(0.00943)***
rammer fattige	0,11531	(0.00893)***	0,04159	(0.00637)***	0,13395	(0.00909)***
sænker lønninger	0,09273	(0.00923)***	0,03098	(0.00651)***	0,09472	(0.00944)***
belaster staten	0,09487	(0.00900)***	0,01623	(0.00634)***	0,09784	(0.00920)***
færre jobs	0,07638	(0.00929)***	0,00024	(0.00637)	0,07478	(0.00950)***
Østrig	0,31451	(0.03047)***	0,01087	(0.02340)	0,29024	(0.02923)***
Belgien	0,02975	(0.02651)	-0,00800	(0.01730)	0,02622	(0.02672)
Danmark	0,15257	(0.03369)***	0,08808	(0.02714)***	0,20727	(0.03126)***
Spanien	0,18534	(0.01863)***	-0,02005	(0.01204)*	0,12401	(0.01868)***
Finland	0,27334	(0.02932)***	0,02792	(0.02302)	0,26484	(0.02782)***
Frankrig	0,06402	(0.01383)***	-0,02818	(0.00856)***	0,05945	(0.01390)***
UK	0,07131	(0.01379)***	-0,02011	(0.00864)**	0,04812	(0.01384)***
Grækenland	0,35268	(0.02806)***	-0,00563	(0.01714)	0,33896	(0.02709)***
Irland	-0,09865	(0.03835)***	-0,03217	(0.02502)	-0,10784	(0.04027)***
Italien	-0,04772	(0.01562)***	0,00863	(0.01115)	-0,04583	(0.01603)***
Luxembourg	0,24080	(0.15413)	-0,04774	(0.09452)	0,21154	(0.15036)
Holland	0,10511	(0.02055)***	-0,06437	(0.01092)***	0,06101	(0.02043)***
Portugal	0,21427	(0.03000)***	-0,01386	(0.01930)	0,17345	(0.02993)***
Sverige	-0,19645	(0.02422)***	-0,06995	(0.01429)***	-0,23810	(0.02473)***
# obs	17630		17502		17591	
Pseudo R^2	0,18380		0,05010		0,18880	
Log likelihood	-9736,38		-7427,02		-9843,86	

***, ** og * angiver signifikans på henholdsvis 1%-, 5%- og 10%-niveau.

virker folks holdning til immigration, har vi foretaget regressioner, hvor svarene til ovenstående spørgsmål inkluderes som forklarende variable. Dvs. vi har estimeret probit-modellerne fra tabel 4.1 udvidet med disse seks variable. Resultaterne fremgår af tabel 5.2.

Ikke særligt overraskende ser vi, at hvis man mener, at der er de ovenstående problemer forbundet med immigration, er man også mere skeptisk over for immigration. Det gælder ikke blot i forhold til niveauet af immigration (H_1 og H_3), men også for folks holdning til immigration fra fattige europæiske lande i forhold til rige europæiske lande (H_2).¹⁰ Det er f.eks. tilfældet, at folk, der mener, at immigranter modtager mere fra det offentlige end de yder i skatter, i højere grad er modstandere af immigration generelt set, og samtidig også er mere stemt for restriktioner på immigration fra de nye EU-lande. Tallene i tabel 5.2 betyder f.eks., at en person, der mener, at indvandrere belaster staten, har 9,5 procentpoint højere sandsynlighed for at være mod immigration fra de nye EU-lande end en person der ikke mener dette, jf. resultaterne for H_1 , mens det kun øger sandsynligheden for at være mere imod immigration fra fattige europæiske lande end fra rige europæiske lande med 1,6 procentpoint. Til sammenligning ses det, at hvis man mener, at kulturen undergraves af indvandrere, så er sandsynligheden for at være modstander af immigration hele 20,0 procentpoint højere.

Igen har vi undersøgt, om effekterne af de forklarende variable er anderledes i Danmarks tilfælde ved at inkludere krydsled mellem dummyvariablen for Danmark og de øvrige forklarende variable. Det var imidlertid igen tilfældet, at ingen af disse krydsled indgik i modellerne med signifikante fortegn, og de er derfor ikke rapporteret. Dvs. der er ikke meget, der tyder på, at f.eks. det at være af den overbevisning, at immigration presser lønningerne ned i Danmark, er relateret til ens holdning til immigration på en måde, der er signifikant forskellig fra EU-15 som helhed.

Det skal bemærkes, at det kan være problematisk at anlægge en kausal fortolkning af koefficienterne til variablene for de opfattede konsekvenser af immigration i tabel 5.2. Der kan således argumenteres for, at disse variable er endogene i forhold til holdningerne til immigration. F.eks. kan man forestille sig, at det er nogle andre uobserverede individuelle faktorer, f.eks. »frygten for det fremmede«, som giver anledning til, at folk både er skeptiske over for immigration og har nogle særlige forventninger til, hvad immigration indebærer. Det kan også tænkes, at respondenterne efterrationaliserer således, at når de har angivet, at de er skeptiske over for immigration, så forsøger de at retfærdiggøre det med, at de også har forventninger om negative konsekvenser af immigration. Denne type endogenitet er kendt fra andre studier, der benytter data fra spørgeskemaundersøgelser og kaldes »justification bias«.

Hvis opfattelsesvariablene er endogene, må vi derfor nøjes med at konstatere, at der er en korrelation mellem holdninger til immigration og opfattelser af konsekvenser af immigration og altså ikke nogen entydig kausal sammenhæng. Vi er dog af den opfattelse, at endogenitet ikke udgør et meget stort problem. Vi har således forsøgt at inklu-

10. Det har dog ikke en signifikant effekt på det relative holdningsmål, at man mener, at immigranter tager job fra indfødte.

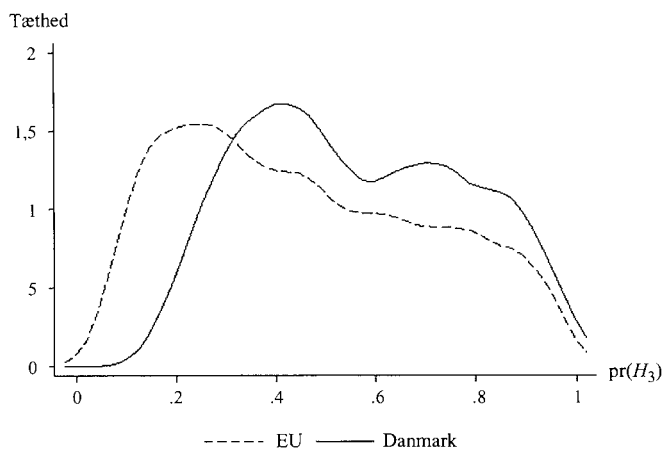
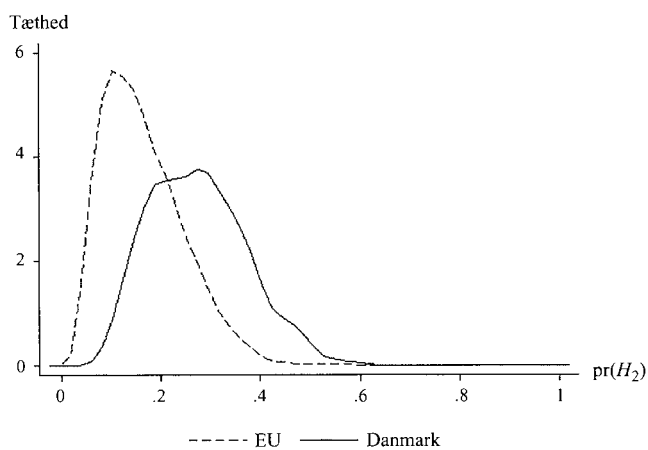
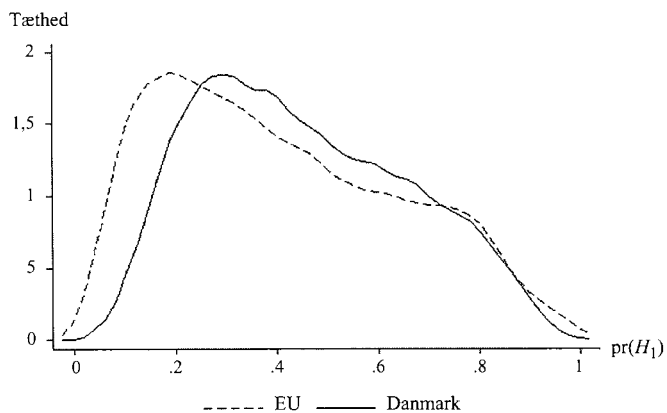
dere to forklarende variable i modellerne i tabel 5.2, der kan siges at fange »frygten for det fremmede«.

De to variable måler i hvor høj grad, respondenterne har modstand mod henholdsvis at have en indvandrer af anden etnisk baggrund som chef, og at en indvandrer af anden etnisk baggrund gifter sig med en person i nærmeste familie (mere om disse variable følger i afsnit 7). Konsekvensen af at inkludere disse variable er for det første, at de har en signifikant positiv effekt på alle tre holdningsmål, og for det andet at koefficientestimerne til opfattelsesvariablene falder en anelse. Eksempelvis falder koefficienten til variabelen *undergraver kultur* i modellen for H_1 fra 0,200 til 0,183, og dette er kendetegnende for størrelsesordenen af ændringerne. Alle opfattelsesvariable vedbliver således med at have en klar positiv og signifikant effekt på holdningerne til immigration, og vi opfatter dette som en indikation af, at der ikke er store problemer med denne type endogenitet i resultaterne i tabel 5.2.

Hvis »justification bias« skulle være et generelt fænomen i stikprøven, burde det være at observere for alle samfundsgrupper. Som tidligere omtalt finder vi imidlertid i Malchow-Møller mfl. (2006), at blandt personer, der forventer negative konsekvenser af immigration, er de, der personligt bliver påvirket af disse negative konsekvenser, endnu mere skeptiske over for immigration. Hvis man f.eks. mener, at mennesker, der flytter hertil, generelt tager arbejde fra danskere, er man mere tilbøjelig til at være modstander af immigration, men personer, der samtidig er arbejdsløse, er endnu mere skeptiske. Dette viser for det første, at privatøkonomiske overvejelser er vigtige i forhold til holdninger til immigration, men for det andet – og dette er relevant i relation til endogenitetsproblematikken – indikerer det også, at opfattelsesvariablene rent faktisk fanger det, der bliver spurgt til. Vi kan imidlertid ikke udelukke eksistensen af justification bias, så i det følgende skal dette forhold tages i betragtning, når resultaterne fortolkes.

6. Er danskerne specielle?

På baggrund af estimationsresultaterne i tabel 5.2 kan det illustreres, hvorledes holdningerne til immigration fordeler sig i Danmark i forhold til EU-15. Dette kan gøres ved først at beregne den predikterede sandsynlighed for at være imod yderligere immigration for hvert enkelt individ i stikprøven. Alle disse sandsynligheder kan dernæst vises i et tætheddiagram som i figur 6.1. Det fremgår, at tætheden for danskerne er forskudt til højre i forhold til tætheden for EU-15. Dvs. danskernes holdninger til immigration er generelt forskudt i retning af at være mere immigrations-skeptiske. Dette gælder især for det relative holdningsmål, H_2 , og for målet for holdninger til immigration fra fattige ikke-europæiske lande, H_3 , hvilket harmonerer med, at det er for disse mål, at danskernes holdninger afviger mest, jf. tabel 3.1. Eksempelvis viser det sidste



Figur 6.1. Kernel tætheder for predikterede sandsynligheder.

Anm.: De predikterede sandsynligheder er beregnet på baggrund af resultaterne fra tabel 5.2. Sandsynlighederne for EU er vægtes med landestørrelser.

diagram for H_3 , at kun ganske få danskere har en predikteret sandsynlighed for at være imod immigration fra ikke-europæiske lande, der er mindre end 0,2, mens der er tale om en noget større andel af europæerne, for hvem dette er tilfældet.

Spørgsmålet er nu, hvor stor en del af disse forskelle, der kan forklares med: (1) at vi har en anden socioøkonomisk sammensætning af befolkningen end EU-15 som helhed; og (2) at vi opfatter konsekvenserne af immigration anderledes. Vi har allerede konstateret, at forskellene ikke skyldes anderledes effekter (koefficienter) af de forklarende faktorer.

Vi kan analysere dette nærmere ved at dekomponere den overordnede forskel i holdninger til immigration mellem Danmark og EU-15 i faktorer der skyldes forskelle i socioøkonomiske forhold, forskelle i opfattelser samt en uforklaret faktor. Den overordnede forskel i holdninger kan beregnes på baggrund af estimationsresultaterne i tabel 5.2 som den gennemsnitlige danskers predikterede sandsynlighed for at være imod yderligere immigration fratrukket den gennemsnitlige EU-15 borgers predikterede sandsynlighed. For det første holdningsmål, H_1 , er disse to gennemsnit henholdsvis 45,2% og 42,0%. Dvs. forskellen, vi søger at forklare, er 3,2 procentpoint.¹¹

Andelen af denne forskel, der kan henføres til forskelle i socioøkonomiske karakteristika, kan nu approksimeres som forskellene i hver socioøkonomisk variabel ganget med den marginale effekt fra tabel 5.2. For holdningsmål, H_1 , viser det sig, at det forhold, at danskerne er forholdsvis højreorienterede, kan forklare næsten 1 procentpoint af forskellen, jf. tabel 6.1. Dette bliver dog modvirket af, at danskerne er relativt højtuddannede, hvilket trækker i modsat retning med cirka samme effekt. Samlet set har de socioøkonomiske forskelle derfor næsten ingen forklaringskraft.

Derimod ser vi, at de seks opfattelsesvariable har noget større betydning. Igen er der dog tale om modsatrettede effekter. Eksempelvis konkluderede vi ovenfor, at danskerne i forhold til europæerne under ét er mere bekymrede for effekten af immigration på skatter og/eller offentlig service. Vi ser, at denne variabel bidrager med at forklare 0,7 procentpoint af de 3,2 procentpoints forskel. I modsat retning trækker de øvrige økonomiske konsekvenser af immigration. Danskerne er forholdsvis lidt bekymrede for disse konsekvenser, hvilket trækker i retning af, at danskerne burde være mindre skeptiske over for immigration. Effekten af de tre øvrige variable summerer faktisk til mere end 5 procentpoint. Alt i alt betyder danskernes afvigende opfattelser af konsekvenserne af immigration, at de burde være mindre skeptiske svarende til 5,1 procentpoint. Det medfører, at den uforklarede forskel i holdningerne stiger til 8,4 procentpoint. Eller sagt på

11. Bemærk, at disse tal svarer meget godt til de rå forskelle i data fra tabel 3.1. I stedet for at beregne predikterede sandsynligheder for henholdsvis en gennemsnitlig dansker og en gennemsnitlig EU-15 borger, kunne man beregne predikterede sandsynligheder for alle danskere (og EU-15 borgere) i stikprøven og dernæst beregne gennemsnittene heraf. Fordi modellen ikke er lineær, giver de to metoder ikke præcis samme resultat.

Tabel 6.1. Dekomponering af forskel.

	H_1	H_2	H_3
<i>Predikteret sandsynlighed:</i>			
Danmark	0,4515	0,2633	0,5820
EU-15	0,4196	0,1516	0,4770
Forskel	0,0318	0,1117	0,1050
<i>Dekomponering af forskel:</i>			
uddannelse	-0,0134	0,0028	-0,0144
politisk overbevisning	0,0096	0,0093	0,0103
andre soc.økon. faktorer	0,0033	-0,0006	0,0006
undergraver kultur	-0,0077	-0,0005	-0,0085
forværrer kriminalitet	0,0020	0,0008	0,0023
rammer fattige	-0,0172	-0,0062	-0,0200
lavere lønninger	-0,0164	-0,0053	-0,0167
belaster staten	0,0073	0,0012	0,0076
færre jobs	-0,0191	-0,0001	-0,0187
I alt	-0,0516	0,0014	-0,0576

en anden måde: Hvis danskerne havde haft samme opfattelse af konsekvenserne af immigration som resten af europæerne i EU-15, ville vi være (langt) mere skeptiske over for immigration, end tilfældet er.

Vender vi os nu mod det relative holdningsmål, H_2 , så er Danmark det land, hvor indbyggerne er mest skeptiske over for immigration fra fattige europæiske lande i forhold til rige europæiske lande. Dette kan aflæses af en forholdsvis stor forskel på 11,2 procentpoint i de gennemsnitlige predikterede sandsynligheder, jf. tabel 6.1. Som det fremgår af dekomponeringen for H_2 , kan denne forskel dog ikke tilskrives forskellige socioøkonomiske karakteristika eller forskellige opfattelser af konsekvenser. De socioøkonomiske forskelle forklarer ca. 1,1 procentpoint af forskellen, mens de anderledes opfattelser trækker i den modsatte retning med cirka samme effekt. Det faktum, at danskerne er mere bekymrede for effekten på skatter og offentlig service kan i sig selv kun mindske forskellen med ca. 0,12 procentpoint. Frygt for de økonomiske konsekvenser kan tilsyneladende ikke forklare, hvorfor danskerne i (langt) højere grad end andre ønsker at skelne mellem immigranter fra fattige og rige europæiske lande.

Angående holdningsmålet for immigration fra fattige lande uden for Europa, H_3 , er forskellen på den gennemsnitlige danskers og den gennemsnitlige EU-15 borgers predikterede sandsynlighed for at være imod 10,5 procentpoint. Når denne forskel dekomponeres, ser vi, at de socioøkonomiske forskelle under ét ikke forklarer noget af forskellen mellem holdningerne i Danmark og i de øvrige EU-15 lande. Forskellene i opfattelser af konsekvenser trækker til gengæld i retning af, at danskerne burde være

mindre skeptiske over for denne type immigration svarende til 6 procentpoint. Den uforklarede forskel i holdninger vokser dermed til 16,5 procentpoint.

Det samlede billede er derfor for alle holdningsmål, at vi står tilbage med en stor uforklaret komponent i forskellene i holdninger til immigration mellem Danmark og EU-15. Forskellene i opfattelser af konsekvenser betyder endda, at den uforklarede del bliver større for de absolutte mål, når der tages højde for disse.

For at få en ide om hvad disse forskelle kan dække over, kan vi kigge på, hvilke kvalifikationer folk vægter hos fremtidige immigranter. Dette er også relevant i relation til, om der skal stilles specifikke krav til de kvalifikationer, eventuelle immigranter skal besidde. Velfærdskommissionen foreslår f.eks., at der skal indføres et pointsystem, der skal sikre, at immigranter besidder de kvalifikationer, der vurderes at være brug for i Danmark, se Velfærdskommissionen (2005a). I ESS 2002/3 er der således stillet en række spørgsmål omkring, hvilke kvalifikationer ved immigranter, som borgerne anser for vigtige, herunder økonomiske kvalifikationer som uddannelse, formue og relevans for arbejdsmarkedet. Danskerne skiller sig imidlertid ikke ud som et land, hvor borgerne ønsker specielt restriktive krav til immigranternes økonomiske kvalifikationer.

Beregningerne ovenfor kan også bruges til at vurdere, hvor stor en del af modstanden mod yderligere immigration, der kan tilskrives bekymringer for økonomiske konsekvenser. Danskerne er som sagt forholdsvist lidt bekymrede, men hvor skeptiske ville vi være, hvis vi slet ikke var bekymrede for de økonomiske konsekvenser? Ved brug af samme metode som ovenfor bliver svaret, at den gennemsnitlige sandsynlighed for at være imod immigration målt ved H_1 ville falde fra 45,2% til 38,0%. For det relative holdningsmål ville modstanden falde fra 26,3% til 24,2%. Endelig ville modstanden målt ved H_3 falde fra 58,2% til 50,3%.¹² I alle tre tilfælde er der således ikke tale om, at de økonomiske bekymringer kan tilskrives en dramatisk rolle.

Bemærk, at dette kun fortæller os, hvorvidt økonomiske argumenter har betydning for holdninger – ikke om det nødvendigvis er en egoistisk økonomisk interesse, der ligger til grund. For at fange dette mere snævre aspekt er man nødt til at interagere de opfattede konsekvenser med mål for, hvorvidt de adspurgte befinder sig i den tilhørende risikogruppe. Se Malchow-Møller mfl. (2006) for flere detaljer.

7. Hvorfor er danskerne specielle?

Analyserne i de foregående afsnit peger i retning af, at danskerne er »specielle« på tre væsentlige områder. Danskerne bekymrer sig relativt lidt om arbejdsmarkedskonsekvenserne af immigration. Til gengæld bekymrer de sig relativt meget om konsekvenserne for velfærdsstaten. Endelig skiller danskerne sig ud ved, at en mindre del af

12. Bemærk, at anvendelsen af de marginale effekter i dette tilfælde er forbundet med større usikkerhed, fordi der analyseres effekter af relativt store ændringer i de forklarende variable.

danskernes modstand mod immigration kan henføres til økonomiske faktorer. Dvs. danskerne er af »uforklarlige« årsager relativt skeptiske over for immigration fra fattige lande. I dette afsnit diskuteres kort forhold, der kan bidrage til at forstå, hvorfor danskerne afviger fra de øvrige EU-borgere på disse områder.

En mulig forklaring på, at danskerne ikke i særlig høj grad bekymrer sig om arbejdsmarkedskonsekvenserne af immigration er, at det fleksible danske arbejdsmarked med en forholdsvis lav arbejdsløshed er bedre indrettet til at modtage en stigning i arbejdsudbuddet end arbejdsmarkedene i de øvrige EU-lande. Udover at være fleksibelt er det danske arbejdsmarked også karakteriseret ved, at der igennem overenskomster er fastlagt minimumslønninger, der er ret høje i forhold til, hvad der gælder i de øvrige EU-lande, se Velfærdskommissionen (2005b). Det kan betyde, at dansk arbejdskraft føler sig relativt godt beskyttet imod konkurrence fra immigreret arbejdskraft. En anden årsag til, at etniske danskere føler sig mindre truede på arbejdsmarkedet af immigration end borgerne i andre EU-lande, kan være, at beskæftigelsesfrekvenserne for immigranter fra ikke-vestlige lande er lave i Danmark i forhold til, hvad der gælder i EU som helhed, se Jensen (2005) og Velfærdskommissionen (2005b). Dvs. danskere har i mindre grad end andre oplevet, at immigranter »tager deres job«.

Lave beskæftigelsesfrekvenser for indvandrere fra ikke-vestlige lande kan også bidrage til at forklare, hvorfor danskerne er relativt bekymrede for, at denne form for immigration har negative konsekvenser for velfærdsstaten. Danmark er et af de lande i EU, der har de mest generøse overførselsindkomster, se Velfærdskommissionen, (2005b), og dette kombineret med, at immigranter fra ikke-vestlige lande i høj grad ikke er i beskæftigelse, bevirker, at immigration fra ikke-vestlige lande har givet anledning til en relativ stor nettoudgift for det offentlige. I Jensen (2005) fremgår det, at i år 2000 var konsekvenserne for de danske offentlige finanser af tilstedeværelsen af ikke-vestlige indvandrere et nettotab på 0,8% af BNP.

Et af hovedresultaterne fra analysen i de tidligere afsnit er, at en mindre del af danskernes modstand mod immigration kan henføres til økonomiske faktorer end tilfældet er for borgerne i de øvrige 15 EU-lande. En mulig forklaring er, at den danske immigrationshistorie afviger fra de øvrige EU-landes, og at dette har en selvstændig indflydelse på danskernes holdninger. Hvis vi ser på tilstedeværelsen af immigranter i de 15 EU-lande, adskiller Danmark sig ikke fra de øvrige EU-lande, idet Danmark ligger i midterfeltet med hensyn til, hvor stor en andel af befolkningen, der udgøres af immigranter fra (fattige) ikke-vestlige lande, se f.eks. Velfærdskommissionen (2005b). Danmark er dog karakteriseret ved, at tilstedeværelsen af immigranter i historisk sammenhæng er et ret nyt fænomen, og der har været en meget stor immigration af personer fra ikke-vestlige lande i 1990'erne, se Velfærdskommissionen (2005b). Velfærdskommissionen (2005b) ser også på sammensætningen af indvandrere fra ikke-vestlige

lande, og her skiller Danmark sig ud ved, at indvandrernes uddannelsesniveau er noget lavere, end hvad tilfældet er i de øvrige EU-lande, og dette er ikke en konsekvens af, hvilke lande immigranterne kommer fra. Med andre ord, så har Danmark for en given landesammensætning tiltrukket de dårligst uddannede immigranter. Med hensyn til landesammensætningen er immigranter i Danmark iøvrigt karakteriseret ved, at i forhold til, hvad der gælder for de øvrige EU-lande, kommer en stor andel fra Mellemøsten og Tyrkiet.

Et åbent spørgsmål er, om danskernes skepsis over for immigration fra fattige lande er et udtryk for »frygt for fremmede« eller måske endda racisme. I ESS 2002/2003 har respondenterne besvaret spørgsmål, der eventuelt kan kaste lys på dette spørgsmål. Svar på følgende spørgsmål forekommer i denne sammenhæng interessante:

1. I hvilket omfang synes du, at Danmark bør tillade mennesker, der tilhører samme etniske gruppe som de fleste danskere, at flytte hertil?
2. I hvilket omfang synes du, at Danmark bør tillade mennesker, der tilhører en anden etnisk gruppe, end de fleste danskere, at flytte hertil?
3. Når det skal afgøres, om en person, der er født, opvokset og bosat i et andet land, skal kunne flytte til Danmark – hvor stor betydning bør det have, at vedkommende er hvid?
4. Når det skal afgøres, om en person, der er født, opvokset og bosat i et andet land, skal kunne flytte til Danmark – hvor stor betydning bør det have, at vedkommende kommer fra en kristen baggrund?
5. Når det skal afgøres, om en person, der er født, opvokset og bosat i et andet land, skal kunne flytte til Danmark – hvor stor betydning bør det have, at vedkommende følger dansk levevis?
6. Hvis du tænker på de mennesker, der er flyttet til Danmark fra et andet land, og som tilhører en anden etnisk gruppe end de fleste danskere – hvad ville du sige til, at en sådan person fik stillingen som din chef?
7. Hvis du tænker på de mennesker, der er flyttet til Danmark fra et andet land, og som tilhører en anden etnisk gruppe end de fleste danskere – hvad ville du sige til, at en sådan person giftede sig med en i din nærmeste familie?

Angående de to første spørgsmål, skiller danskerne sig markant ud fra de øvrige EU-borgere. I Danmark mener 75%, at man skal tillade flere immigranter, der tilhører samme etniske gruppe som de fleste danskere, mens kun 48% mener, at man skal tillade flere immigranter med en anden etnisk baggrund. Blandt EU-15 landenes borgere som helhed er det 66%, der mener, at man skal tillade flere med samme etniske baggrund, og 54%, der mener, at man skal tillade flere med en anden etnisk baggrund.

Dvs. danskerne er langt mere positive over for immigranter med samme etniske baggrund end de øvrige EU-borgere, men er til gengæld mere kritiske, når det gælder immigranter med en anden etnisk baggrund. Denne skepsis over for immigranter med en anden etnisk baggrund kan naturligvis have rod i flere forhold. Den kan f.eks. skyldes racisme, frygt for fremmede, eller simpelthen, at danskerne mener, at der er for store problemer med at integrere personer med en anden etnisk baggrund i det danske samfund.

Et spørgsmål, der mere direkte går på betydningen af racemæssig baggrund, er spørgsmål 3. Her er det 11% af danskerne, der mener, at det er vigtigt, at immigranter er hvide, og det er færre end blandt EU-borgerne som helhed, hvor det er 13%. Angående betydningen af en kristen baggrund skiller danskerne sig heller ikke markant ud, idet 26% mener, at det er en vigtig egenskab, mens 24% blandt EU-borgerne som helhed mener, at det er en vigtig egenskab. Det er også interessant, at blandt danskerne er der 70%, der mener, at det er vigtigt, at immigranter følger dansk levevis, mens det tilsvarende tal for EU-borgerne som helhed er 80%. Dvs. danskerne lægger faktisk relativt lidt vægt på, at immigranter følger landets levevis. Det eneste af de 15 EU-lande, hvor borgerne lægger mindre vægt på dette, er Irland.

Angående de to sidste spørgsmål, er der 21% af danskerne, der ville have noget imod, at deres chef havde en anden etnisk baggrund, hvilket er præcis det samme, som det der gælder blandt EU-borgerne som helhed. Til gengæld er der 32% af danskerne, der ville have noget imod, at et medlem af den nærmeste familie blev gift med en person af anden etnisk herkomst, mens det kun gælder for 26% af EU-borgerne som helhed.

En anden måde hvorved man kan afklare, om danskerne skulle være specielt racistiske eller fremmedfjendske er at se på, hvor mange der oplever diskrimination som følge af race eller anden etnisk baggrund. I Velfærdskommissionen (2005b) refereres en undersøgelse fra Europa Kommissionen, hvor det fremgår, at Danmark er blandt de lande, hvor relativt få oplever diskrimination.

Baseret på ovenstående diskussion må man sige, at indretningen af det danske arbejdsmarked – kombineret med den danske immigrationshistorie – kan give en del af forklaringen på, hvorfor danskerne ikke er bekymrede for arbejdsmarkedskonsekvenserne ved immigration, men til gengæld er bekymrede for, hvorvidt det vil forringe velfærdsstaten. Hvorfor danskerne af »uforklarlige« årsager er mere skeptiske over for immigranter fra fattige lande, end tilfældet er blandt EU-borgerne som helhed, er det langt vanskeligere at give et godt bud på. Der er dog en del, der peger i retning af, at danskerne grundlæggende set er mere skeptiske over for mennesker med en anden etnisk baggrund end tilfældet er blandt EU-borgerne som helhed. Det er dog vigtigt at understrege, at der ikke er tale om et entydigt billede.

8. Konklusion

Danskere er mere skeptiske over for immigration, end hvad der gælder i EU-15 som helhed. Det gælder specielt, når vi betragter den relative modstand mod immigration fra fattige europæiske lande i forhold til immigration fra rigere europæiske lande. Tallene tyder således på, at en markant større andel af danskerne bifalder, at der gælder skrappe regler for indvandring fra de nye relativt fattige EU-lande, end der gør fra de gamle »rige« EU-lande. Umiddelbart kunne dette tyde på en vis frygt for især de økonomiske konsekvenser af immigration i Danmark.

Vi har forsøgt at forklare holdninger til immigration ud fra socioøkonomiske karakteristika samt de opfattede (især økonomiske) konsekvenser af immigration. De mest skeptiske over for immigration er generelt set de lavtuddannede, de fattige, de arbejdsløse, de ældre og folk på landet. På dette punkt adskiller danskere sig ikke fra befolkningerne i de øvrige EU-15 lande. Endvidere er folk, der er bekymrede for de økonomiske konsekvenser af immigration generelt mere modstandere af immigration fra fattige lande.

Danskerne adskiller sig derimod ved at forholdsvis få forbinder øget immigration med negative økonomiske konsekvenser. Danskere er således betydeligt mindre bekymrede for arbejdsmarkedskonsekvenserne af immigration, mens de er en anelse mere bekymrede for finansieringen af velfærdsstaten. Kontrollerer vi for de økonomiske opfattelser, står vi faktisk tilbage med en betydelig uforklaret forskel mellem danskernes og de øvrige europæeres holdninger til immigration – både de absolutte og de relative. Det indledende billede af en større frygt for de økonomiske konsekvenser i Danmark holder således ikke.

Alt i alt tyder vores resultater derfor på, at danskerne hverken skiller sig ud mht. hvordan socioøkonomiske faktorer eller frygten for økonomiske konsekvenser af immigration påvirker holdningerne. Til gengæld er forholdsvis få danskerne bekymrede for de økonomiske konsekvenser af indvandring, og dette trækker i retning af, at danskerne alt andet lige burde være relativt positive over for indvandring. Sammenholdes dette med, at danskerne i gennemsnit er mere skeptiske over for indvandring end EU-15 landene som helhed, kommer Danmark til at fremstå som et væsentligt mere immigrations-skeptisk land.

Litteratur

- Andersen, J. G. 2002. Fremmedhad eller tolerance? Danskernes holdninger til indvandrere. En forskningsoversigt, *Arbejds-papir*, Aalborg Universitet.
- Boeri, T. og H. Brücker. 2005. Why are Europeans so Tough on Migrants?, *Economic Policy*, 20, 629-703.
- Card, D., C. Dustmann og I. Preston. 2005. Understanding Attitudes to Immigration: the Migration and Minority Module of the First European Social Survey, *CREAM Discussion Paper* nr. 03/05.
- Dustmann, C. og I. Preston. 2004. Racial and Economic Factors in Attitudes to Immi-

- gration, *CREAM Discussion Paper* nr. 01/04.
- Hainmueller, J. og M. J. Hiscox. 2005. Educated Preferences: Explaining Attitudes Toward Immigration in Europe, Working paper.
- Jensen, B. 2005. *Indvandringen til Europa. Velfærdsstat og integration*. Gyldendal.
- Malchow-Møller, N., J. R. Munch, S. Schroll og J. R. Skaksen. 2006. Attitudes Towards Immigration – Does Economic Self-Interest Matter?, *IZA Discussion paper* nr. 2283.
- Mayda, A. M. 2006. Who is Against Immigration? A Cross-Country Investigation of Individual Attitudes Towards Immigrants, *Review of Economics and Statistics*, 88, 510-30.
- Nielsen, H. J. 2004. *Er danskerne fremmed-fjendske? Udlandets syn på debatten om indvandrere 2000-2002*. Rockwool Fondens Forskningsenhed, Aarhus Universitetsforlag.
- Scheve, K. F. og M. J. Slaughter. 2001. Labor Market Competition and Individual Preferences over Immigration Policy, *Review of Economics and Statistics*, 83, 133-45.
- Velfærdskommissionen. 2005a. *Fremtidens velfærd og globaliseringen*, Analyserapport, Marts 2005.
- Velfærdskommissionen. 2005b. *Fremtidens Velfærd – sådan gør andre lande*, Analyserapport, Marts 2005.

A Dynamic Random Effects Multinomial Logit Model of Household Car Ownership

Thomas Bue Bjørner

AKF, Danish Institute of Governmental Research, E-mail: tbb@akf.dk

Søren Leth-Petersen

Department of Economics, University of Copenhagen, E-mail: soren. leth.petersen@econ.ku.dk

SUMMARY: Using a large household panel we estimate demand for car ownership by means of a dynamic multinomial model with correlated random effects. Results suggest that the persistence in car ownership observed in the data should be attributed to both true state dependence and to unobserved heterogeneity (random effects). It also appears that random effects related to single and multiple car ownership are correlated, suggesting that the IIA assumption employed in simple multinomial models of car ownership is invalid. Relatively small elasticities with respect to income and car costs are estimated. It should, however, be noted that the level of state dependence is considerably larger for households with single car ownership as compared with multiple car ownership. This suggests that the holding of a second car will be more affected by changes in the socioeconomic conditions of the household and by economic policy shocks.

1 Introduction

The stock of cars held by individual households is observed to be persistent when the household is observed repeatedly. If a given household has one car in one year then it is likely to be observed holding one car also in the following year. This pattern will have important implications for assessing the impact of policy shocks on the probability of changing the car stock in the short and long run. In this paper we present an empirical analysis of the dynamics of car holdings that allow us to characterize the persistence in the car stock at the household level.

Persistence in the car stock at the household level can arise for different reasons.

We thank Kenneth Train for providing the mixed logit estimation software developed by Kenneth Train, David Revelt, and Paul Ruud. The research was supported by the Danish Environmental Research Programme.

One possibility is that persistence is caused by unobserved household specific preferences for car holdings that are constant across time. This is known as »spurious« state dependence, Heckman (1981). Alternatively, it can be due to, for example, the presence of transaction costs associated with adjusting the size of the car stock or with habit formation. Transaction costs are unobserved but not fixed across time, and will result in »true« state dependence. These two sources of persistence nevertheless have very different implications in terms of policy analysis. For example, consider a policy introducing a new tax on car ownership. If persistence is caused by unobserved fixed differences in preferences for cars then this policy will have an immediate effect that is identical in both the short and the long run. On the other hand, if persistence is generated by the presence of transaction costs associated with adjusting the car stock then the policy will have an effect that is different in the short and the long run.

An econometric model should be capable of handling both sources of persistence. To distinguish between these two types of persistence we estimate a dynamic discrete choice model with unobserved time-invariant heterogeneity employing the approach suggested by Wooldridge (2005). Estimating this model will permit us to evaluate the importance of time invariant unobserved heterogeneity versus state dependence. The analysis is based on a panel data set allowing us to follow the car stock of a large number of individual households over a period of ten years. The dataset is based on merged public administrative register data that give us information about the car stock, income, family composition, age etc. for a large number of Danish households for the period 1992 to 2001.

Most previous studies are based on either aggregate data or micro cross-section data. Some studies based on repeated cross-section data employ pseudo panel methods; see for example Dargay and Vythoulkas (1999), and Dargay (2001). Both types of data lack the idiosyncratic aspect, and such methods cannot be used to distinguish between persistence due to unobserved heterogeneity and persistence due to state dependence. The few previous studies on car ownership using micro panel data have either made arbitrary assumptions regarding the nature of the persistence in data or used a less general specification of the empirical model than the one applied here.¹ As an example, Meurs (1993) assumed that the persistence in car ownership should be attributed to unobserved heterogeneity. Kitamura and Bunch (1990) estimated models which included both state dependence and unobserved heterogeneity. However, they applied an ordered probit model, which is restrictive in the sense that the same index function is used to determine both single and multiple car ownership (with the exception of a threshold parameter). That is, changes in socioeconomic variables have the same ef-

1. An exception is the work reported in Hensher et al. (1992).

fect (on the index) for single and multiple car ownership. The multinomial logit model is more flexible as it allows for different parameters for single and multiple car ownership. This is likely to be important if a first car does not serve the same purpose as a second car.

In a previous paper based on the same data source as applied here Bjørner and Leth-Petersen (2005) assessed the dynamic properties of car ownership of single people and of households consisting of two adults («couples»). For couples the decision to hold two cars relative to one was modelled as a separate decision from the decision to hold one car relative to no car. This assumption simplifies estimations, but the assumption is questionable, because households that for some unobserved reason prefer to hold one car more than no car are also likely to have unobserved preferences for holding two cars. Such unobserved correlated heterogeneity in demand for car ownership can arise, for example, because households have different access to public transport that can be used as a substitute for car transportation. In this paper a more complete analysis is presented for the decision for couples to hold 0, 1 or 2 cars. Here the choice of car stock is modelled as a dynamic multinomial choice model that allows for unobserved correlated heterogeneity across car stock categories.

A number of previous micro (cross section) studies on car ownership have modelled the demand for car ownership and choice of car type as interrelated choices, see among others Train (1986) or Hensher (1985) for some early contributions and Bento et al. (2005) for a recent example, which also includes modelling of car use. The intuition is that the utility of different car types (represented by their different observed characteristics) influences the utility of car ownership, which is estimated using nested logit models. With the noteworthy exception of Hensher et al. (1992) the previous micro panel studies appear to have focussed on the ownership decision without including the car type decision.² In this article we also focus on the car ownership decision without looking at the choice of car type. However, the dynamic discrete models applied here may potentially be merged with the line of models, where car ownership and type choice are interrelated.

Results show that both unobserved time invariant heterogeneity and state dependence are important factors to include in the analysis of household car ownership. Households are shown to respond very little in the short run to changing income and user cost levels. This has important implications for understanding the effects of policy measures in the short run.

In the next section we present the dataset. In section 3 the econometric framework is established and results are presented in section 4. Section 5 assesses the economic importance of different policy shocks. Section 6 concludes.

2. A fairly recent methodological survey of this literature can be found in Bunch (2000).

2 Data

The analysis is based on a household level panel data set with information about 10,565 Danish households that we are able to follow in all years in the period 1992-2001. The dataset is constructed by merging different public administrative registers at the individual level (using the civil registration number) allowing us to observe changes in car ownership, income, age, family composition, location of residence and labour market participation status. To be more specific, the information on car ownership was obtained from the Danish Central Register for Motor Vehicles, which is used to collect annual ownership taxes and is therefore considered to be very accurate. Based on information from this register we calculate the degree of car ownership during the year (share of car ownership days) and subsequently define a discrete car ownership variable (0 if the degree of car ownership during the year was less than 0.5, 1 if the degree was between 0.5 and 1.5 etc.).

Company cars available to private households, but owned by a company cannot be linked with households based on the information from the Danish Central Register for Motor Vehicles. However, information about the presence of a company car in a household was obtained from a tax register (as individuals with a company car in Denmark pay income tax on the benefits of having a car at their disposal).

Socioeconomic variables related to the household were extracted from the tax register and other sources. We have information about income (before and after tax), social transfers, demographic information, labour market status and location at municipal level. The municipality of the workplace was also obtained and used to calculate a measure of commuting distance. An index for the user cost of car ownership was calculated from aggregate information about fuel prices, ownership tax, repair costs, insurance costs, price of new cars and net rate of return (alternative cost). There is mainly variation in the car cost index across time.³

For the analysis we consider a selected sample of households. First of all we consider only households consisting of couples in a balanced panel, and we consider the

3. The measure of user cost reflects the costs of an »average car«. To be more specific cost weights were calculated for the year 1997 using information from Statistics Denmark and FDM. Apart from the insurance cost, where we were able to separate between the three different regions in the model, the cost components had equal size across individuals (geographic differences in car cost that are fixed over time will, however, be captured by the geographic dummies). To calculate the change over time in the user cost consumer price indexes for the different cost components were applied. For the capital cost also the net rate of return was included in the calculations. As pointed out by one of the referees the user cost for an average car may not necessarily apply to a potential new car owner as e.g. insurance costs may be different. This is not taken into account in the analysis, but as the effect of the user cost (mainly) is identified due to variation over time, we expect this to have limited impact in the current setting)

Table 1. Number of changes in car ownership status across households in the period 1992-2001.

	Number of ownership changes in period				
	0	1	2	3	4+
Share of households % (10,565 households)	56	25	13	4	2

Table 2. Distribution of car ownership status for households with stable ownership status.

	Car ownership level		
	0 Car	1 Car	2 Cars
Share of households % (5,939 households)	14	82	4

choice between holding 0, 1 or 2 cars.⁴ For single households the choice of car stock is in practice binary. This analysis is presented in Bjørner and Leth-Petersen (2005). Moreover, households where one person is self-employed are deselected. This is because self-employed individuals have highly unstable incomes when measured by the tax assessed income (which may not reflect their real consumption possibilities). We consider only households where the oldest person is aged 18 years or more. Finally, we do not consider households that have a company car. We prefer to focus on privately owned cars, because it seems likely that the decision-making process to own a private car is different from the process of obtaining a company car.

Descriptive statistics

We have postulated that car ownership status is persistent across time. This claim is backed by the evidence in table 1. The table shows the number of changes in car ownership status across households in the period 1992-2001. The table shows that more than half of the households in the sample never change ownership status in the observation period, and that including up to two shifts in ownership status accounts for roughly 95% of the sample.

For those never changing ownership status in the observation period car ownership status is distributed as shown in table 2. The overwhelming majority of households with stable ownership status have held one car.

4. Only 1-2% of the singles have multiple private car ownership, while less than 1% of the couples own 3 cars (calculated for households without self-employed and without company cars). In general it is worth noting, that car ownership ratios are low in Denmark (even by European standard) presumably due to a combination of a high sales tax on cars and a well developed public transport sector.

Table 3. Distribution of changing pattern for households changing ownership status one time in the observation period.

Share of households % (2,652 households)	Change in car ownership level			
	0 → 1 Car	1 → 0 Car	1 → 2 Cars	2 → 1 Car
	35	15	28	22

The distribution for those households changing status one time in the observation period is given in table 3. Most shifts are from either no car to one car or from one car to 2 cars, i.e. there is an indication of a general tendency for accumulating cars within the households in the sample.

The included explanatory variables are briefly described below. As the income measure we use the log of household income after tax measured in 1997 price level (*linc*). This measure includes wage, pensions, net capital income as well as the most important non-taxed public transfers such as child support (paid in Denmark independent of income), subsidies for housing rents and social benefits. Age (of the oldest person in the couple) is included both in linear and squared forms. A number of dummies indicate labour market status for males and females distinguishing between status as employed (*work*) and unemployed (*unemp*). The reference is individuals outside the labour market (as described, households with self-employed individuals are excluded). For respondents employed we calculate a measure of commuting distance based on the mean distance between municipalities of residence and place of work. For individuals living and working in the same municipality the expected commuting distance was calculated based on the size of the municipality. The square root of commuting distance was included in the models (denoted *distm_sr* and *distf_sr* for males and females, respectively).⁵ Dummy variables are included to indicate the presence of children under 18 years of age (*dchild_m*) and adult children living with their parents (*dchild_a*). The variable (*lusc*) is the log of car user costs, giving the development in car cost (purchase, ownership and use) relative to consumer prices (normalized to 1 in 1997). A *trend* variable normalized at 0 in 1993 is included to account for time effects (annual dummies cannot be included along with *lusc*). Finally, two dummy variables are included to indicate degree of urbanization. One dummy (*cph*) indicates the household resides in Copen-

5. Municipality of workplace was not recorded in 2% to 5% of the cases (for persons working). To accommodate this in our econometric models we include dummy variables taking the value one if this information is missing.

hagen and another dummy (*town*) indicates the household lives in an urban area outside Copenhagen. The reference is households living in rural areas.⁶

Summary statistics for all the variables included in the analysis is given in the appendix.

3 Econometric model

The purpose of the paper is to estimate probability models of car ownership status that can fall into one of three categories: 0, 1 or 2 cars. We estimate three versions of the multinomial logit model with an increasing degree of sophistication. The reference model is the (static) pooled multinomial logit model.

In the second model we expand the pooled multinomial logit model by introducing unobserved heterogeneity. It is assumed that the unobserved heterogeneity of each category is uncorrelated with the explanatory variables in the model, i.e. a random effect. However, we allow the random effects of the categories to be correlated to take into account that households that for some unobserved reason prefer to own one car instead of no car may also be likely to prefer two cars.

In the third model the random effects model is extended by also allowing for state dependence, i.e. inclusion of lagged car ownership status. By introducing state dependence we will be able to assess the quantitative importance of the two sources of persistence that appears in the raw data series.

In terms of estimation, the model with random effects and state dependence is the most complicated. The econometric setup is therefore outlined in terms of this model. The less complicated models are special cases of this model. In the next section we present how to estimate a multinomial logit model with random effects for panel data, and finally we describe how the initial conditions problem associated with dynamic panel data models is handled. Here we follow the approach proposed by Wooldridge (2005).

3.1 Estimation of the dynamic multinomial logit model with random effects

Consider the discrete choice model, where a given individual i chooses the car stock j . The model is

6. It may be argued that some of these variables could be endogenous. For example, couples with high preferences for car ownership may prefer to reside in rural areas where parking space is not restricted. Likewise, it can be argued that income and perhaps especially commuting distance may be endogenous, as (multiple) car ownership is likely to allow the couple to search for job in a larger geographical area. However, as we condition on initial car ownership (see section 3.2) and interpret results as short run (see section 4) the potential endogeneity is likely to be less severe (as compared to a cross section analysis).

$$\text{Prob}(y_{it} = j \mid y_{it-1} = k, y_{i0} = l, x_{it}, \mu_{ji}) = \frac{\exp(\gamma_j y_{it-1} + \lambda_j y_{i0} + \beta_j x_{it} + \mu_{ji})}{\sum_{m=0}^J \exp(\gamma_m y_{it-1} + \lambda_m y_{i0} + \beta_m x_{it} + \mu_{mi})} \quad (3.1)$$

where $j, k, l, m = 0, 1, 2$ and $i = 1, \dots, N$ and $t = 1, \dots, T$. y_{it-1} is past observed holdings. x_{it} is a vector of K observed exogenous variables including age, labour market status, commuting distance, children, user costs, urbanization and a trend. y_{i0} is the initial stock. Thus, in the current period individual i holds $y_{it} = j$ cars, in the previous period he held $y_{it-1} = k$ cars, and in the initial period he held $y_{i0} = l$ cars. j, k, l are equal if the individual has the same car stock in all three periods, but j, k, l can also be different from each other if the car stock varies over the observation period. The inclusion of the initial stock y_{i0} in the model is related to the initial condition problem and will be motivated further in section 3.2. Finally, μ_{ji} is an unobserved household effect specific to the car stock so that a given household is allowed to have an idiosyncratic time invariant preference for a particular stock of cars. We wish to allow for idiosyncratic effects to be correlated across alternatives. To make this setup operational for estimation we assume that μ_i follows a J -dimensional multivariate normally distribution, where J equal 2. Write $\mu_{ji} = C\xi_i$ where μ_i is a $J \times 1$ dimensional vector of multivariate normal distributed (conditional on (y_{i0}, x_{it})) idiosyncratic effects. The fact that μ_i is allowed to be multivariate implies that we do not impose the IIA assumption. ξ_i is a $J \times 1$ vector of independent normally distributed variables, and CC' is the $J \times J$ covariance matrix of μ_i and C is the lower triangular Cholesky factorization of it, containing the unknown parameters of the multivariate normal distribution of time constant idiosyncratic effects. C is given by

$$C = \begin{bmatrix} c_{00} & & \\ c_{10} & c_{11} & \\ c_{12} & c_{21} & c_{22} \end{bmatrix} \quad (3.2)$$

Errors are assumed to be independently distributed across alternatives time and individuals conditional on $y_{it-1}, y_{i0}, x_{it}, \xi_i$, and extreme values distributed.

Conditional on ξ_i (and y_{it-1}, y_{i0}, x_{it}) the probability for a particular household choosing car stock j at time t is then

$$\text{Prob}(y_{it} = j \mid \xi_i) = \frac{e^{\gamma_j y_{it-1} + \lambda_j y_{i0} + \beta_j x_{it} + C_j \xi_i}}{\sum_{k=1}^J e^{\gamma_k y_{it-1} + \lambda_k y_{i0} + \beta_k x_{it} + C_k \xi_i}} \quad (3.3)$$

where C_j is the j^{th} row of C . The probability that household i is observed with a sequence of stocks y_{i0}, \dots, y_{iT} is

$$\text{Prob}(y_i | \xi_i) = \prod_t \prod_j \text{Prob}(y_{it} = j | \xi_i)^{1[y_{it} = j]} \quad (3.4)$$

The unconditional choice probability is

$$\text{Prob}(y_i) = \int_{\xi_i} \text{Prob}(y_i | \xi_i) f(\xi_i) d\xi_i \quad (3.5)$$

where $f(\xi_i)$ is the multivariate distribution for ξ_i . The log likelihood function is

$$\log L = \sum_{i=1}^N \ln [\text{Prob}(y_i)] \quad (3.6)$$

We evaluate the integral in (3.5) by drawing ξ_i^d from the distribution of ξ_i , calculating $\text{Prob}(y_i | \xi_i^d)$, and repeating this $D = 1000$ times, to obtain an average hereof.

$$\text{Pr\hat{o}b}(y_i) = \frac{1}{D} \sum_{d=1}^D \text{Prob}(y_i | \xi_i^d) \quad (3.7)$$

Instead of using pseudo random draws we use Halton draws. For details we refer to Train (2003).

Setting the location

The model is estimable in difference form where we evaluate one alternative relative to a reference alternative. This is because the probabilities of choosing either of the stocks 0, 1, and 2 must sum to unity. Therefore a reference stock must be chosen for which the probability is given by 1 minus the sum of the probabilities of choosing the other stocks. In the example where stock $j = 0$ is the reference and the probabilities for the three stocks are denoted (p_0, p_1, p_2) then $p_0 = 1 - p_1 - p_2$. Thus, estimates of the parameter sets (γ_1, β_1, C_1) and (γ_2, β_2, C_2) are always relative to the base category, $J = 0$. Therefore, C takes the form

$$C = \begin{bmatrix} 0 & & & \\ 0 & c_{11} & & \\ 0 & c_{21} & c_{22} & \\ & & & \end{bmatrix} \quad (3.8)$$

So that the covariance matrix of the random effects term in difference form becomes

$$CC' = \begin{bmatrix} c_{11}^2 & & & \\ c_{11}c_{21} & c_{21}^2 + c_{22}^2 & & \\ & & & \\ & & & \end{bmatrix} \quad (3.9)$$

The term $(c_{11} \ c_{21})$ allows for correlation across alternatives thereby avoiding imposing the assumption of independence across alternative. For example, consider a household with a high unobserved preference for holding two cars and a positive value of the term $(c_{11} \ c_{21})$. This household is then likely also to have a high preference for holding one car as opposed to not holding any cars. Allowing for this sort of unobserved preference correlation is potentially important, because without such correlation the preference of a given household for holding two cars is completely independent of its preference for holding one car. The sign and size of $(c_{11} \ c_{21})$ are to be estimated and are of course not restricted a priori.

3.2 The initial conditions problem

Estimation of dynamic panel data models with unobserved effects is a nontrivial problem. The difficulty arises because it is implausible to assume that the initial observation, y_{i0} , is independent of the unobserved effect, μ_i . In our application this amounts to assuming that the initial stock (i.e. the one firstly observed) is independent of the unobserved preference for holding cars. This is clearly untenable. Here we follow the approach of Wooldridge (2005) to handling the initial conditions problem by modelling the distribution of the unobserved effect, assumed to be normal, conditional on the initial value of the dependent variable, y_{i0} .

Apart from being fully parametric the main assumption underlying Wooldridge's approach is that we specify a parametric model for the density of μ_i conditional on the initial observation of the dependent variable y_{i0} . In practice this amounts to including y_{i0} as an additional regressor.⁷ This modelling approach has previously been applied by Erdem and Sun (2001), albeit without allowing for correlation of the unobserved effects across alternatives.

7. We use a simplified version of the approach suggested by Wooldridge (2005). He also conditioned on the observed history of the exogenous explanatory variables. However, in our case inclusion of variables for the history of some explanatory variables in preliminary regressions indicated that these generally were insignificant and could be excluded without affecting the remaining parameters.

4 Results

Estimation results are presented in table 4. The table shows results from estimating three models: The pooled multinomial logit, the random effects multinomial logit, and the random effects multinomial logit model with state dependence. The order of the presentation of the models represents the increasing level of sophistication. The pooled model gives estimates from the standard multinomial logit model. The random effects model allows for unobserved time invariant heterogeneity that can be correlated across alternatives, but is assumed to be uncorrelated with the explanatory variables. The final model conditions on the lagged level of the dependent variable and allows for unobserved time invariant heterogeneity. The lagged dependent variable will capture the persistence in car ownership status. As we include the lagged dependent variable in the model, the interpretation of the model parameters becomes shorter run. This is because the parameter set then governs responses that are specific to the exact car holding in the previous period. We therefore think of the obtained estimates as having relevance for understanding short run responses.

Considering first the importance of introducing more conditioning elements into the model it is seen that the log-likelihood value becomes numerically smaller as the model becomes richer. This indicates that both unobserved heterogeneity and state dependence are relevant aspects to include in the model from a statistical point of view. It is also evident that the parameters describing the covariance structure become much smaller when state dependence is introduced. This is an indication that (true) state dependence absorbs most of the persistence in the data. It is noticeable that in both models where the covariance structure is estimated there is evidence of correlation across alternatives. This is evidence that the independence of irrelevant alternatives (IIA) invoked in the pooled multinomial logit model is restrictive. The positive parameter on the covariance term indicates that individuals having a preference for holding one car relative to no car also have a preference for holding two cars.

In all the models the parameters on income are positive in the categories holding one car and holding two cars. The parameters on the income variable in the category holding two cars are larger than for the category holding one car. It is tempting at this stage to conclude that this indicates that holding two cars relative to one is more income elastic. It is, however, premature to conclude on the quantitative importance based on the parameter estimates, because the model is nonlinear. In section 5 we shall return to the quantitative importance of income. Both parameters on age and agesquared are significant in all three models in the categories holding one car and holding two cars and indicate a concave relationship.

In the static models the parameter on small children is positive in the category holding one car. This is reversed in the dynamic model. This could indicate that child

expenditures crowd out car expenditures. In all three models the parameter on adult children is positive in the category holding two cars.⁸

The parameters on the variables indicating degree of urbanization appear reasonable. There are negative parameters on the dummy variables for Copenhagen and other urban areas (base case is rural areas) in the categories holding one car and holding two cars. Moreover, the parameter on the user cost variable is »in all cases« negative (both for one and two cars). It should be recalled that we mainly have time variation in the car cost index and the size and significance of the parameter to the car cost index are sensitive to the inclusion/omission of the trend variable, so the impact of changes in car cost should be interpreted cautiously.

Finally, it should be noted that the size of the parameters in the different models cannot directly be compared because the variance of the error term plus random effects are different, see e.g. Wooldridge (2002). The impact of changes in income (after tax) and car costs will therefore be described further in the next section. The t -values of the pooled multinomial logit model are generally considerably larger than in the other models. This is because the t -values of the pooled multinomial logit are calculated subject to the (incorrect) assumption that the errors of each household are uncorrelated. This is clearly not the case, so the t -values of the pooled multinomial logit are strongly upwards biased.

5 Quantitative importance of state dependence, income and user costs

To assess the quantitative importance of state dependence and the economic importance of changes in variables such as income and user costs we need to calculate average predicted probabilities. There are no simple estimators for the average probabilities available for the multinomial logit with normally distributed random effects, see e.g. Wooldridge (2002). Therefore, we calculate the average probabilities over the observed characteristics of the individuals using a simulation approach, where the probability for each household is calculated many times adding draws from the estimated normal distribution to the index function. Let ξ_{ijs} be random draws from the standard normal distribution, where s indexes the draws (1, ..., S), where we set $S = 1000$. The simulated average probability is then:

8. This result could potentially be related to speculation in lower insurance premiums. Young people face very high insurance premiums. Living with their parents they can save money if their car is registered as belonging to one of their parents (given that the parents already have a car and have earned discounts in insurance premiums from collision free years).

Table 5. *The quantitative importance of state dependence: Estimated probability of car ownership in 2001 conditional on different ownership levels in 2000.*

	Assumed car ownership in 2000		
	0	1	2
0 car	0.739	0.020	0.000
1 car	0.261	0.939	0.457
2 cars	0.000	0.041	0.543

$$\text{Average Prob } (y_{it} = j \mid \xi_i) = \frac{1}{N} \sum_{i=1}^N \frac{1}{S} \sum_{S=1}^S \frac{e^{\hat{\gamma}_j y_{it-1} + \hat{\lambda}_j y_{i0} + \hat{\beta}_j x_{it} + \hat{C}_j \xi_i}}{\sum_{k=1}^J e^{\hat{\gamma}_k y_{it-1} + \hat{\lambda}_k y_{i0} + \hat{\beta}_k x_{it} + \hat{C}_k \xi_i}}$$

Table 5 gives the calculated probabilities of holding 0, 1 or 2 cars under different assumptions about lagged car holding. The table clearly shows that probabilities centre on the category of the lagged value. This is most pronounced for the category holding 1 car, and less pronounced for the category holding 2 cars. This indicates that households are quicker to adjust to a policy shock if they hold two cars than if they hold 1 car.

In table 6 partial effects from a one percentage change in income and user costs are presented. The partial effects are calculated as the average probability after the change less the average probability before the change times 100, and they are calculated based on the estimates from the pooled multinomial logit, the random effects multinomial logit (RE) and the random effects multinomial logit with state dependence (RD SD). The bottom row of table 6 gives the elasticities of the total car stock with respect to income and user costs. These elasticities can be compared directly with »macro« elasticities. The numbers in table 6 are most appropriately thought of as characterizing the short run responses.⁹

The general picture appearing from the income effects in table 6 is that the partial effects get smaller as the level of sophistication of the models increases. The largest effects are found in the pooled multinomial logit model that condition on neither unobserved heterogeneity nor lagged levels of car ownership. The smallest response effects are found in the random effects model with state dependence. The estimates from the static random effects model generally lie between those of the other models, as would

9. In dynamic linear models typically employed when using time series methods, it is straightforward to calculate the long run response using the estimated parameter on the lagged dependent variable, see for example Dargay (2001). This is not possible in a nonlinear model.

Table 6. Effect of changes in income and user costs in 2001.

	Income			User costs		
	Pooled	RE	RE SD	Pooled	RE	RE SD
0 car	-0.156	-0.043	-0.018	0.202	0.015	0.118
1 car	0.029	-0.028	-0.027	-0.037	0.082	0.060
2 cars	0.127	0.071	0.046	-0.165	-0.098	-0.178
Car stock elasticity ⁽¹⁾	0.283	0.114	0.065	-0.366	-0.114	-0.295

Note: 'Pooled' is pooled multinomial logit without random effects and state dependence. 'RE' is logit with random effects, but without state dependence. 'RE SD' is logit with random effects and state dependence. The upper part of the table presents partial effects multiplied by 100. ⁽¹⁾ The elasticity of the total car stock is calculated as the percentage change in the predicted car stock (derived from the changes in probabilities also reported in table 6) following a one percentage change in income/user costs.

be expected. Generally, for the models including random effects, the largest income responses are found for the category holding two cars. This is consistent with the interpretation that the second car has less of a necessity nature (in the jargon of demand analysis). The general conclusion, though, is that income changes have little impact on car ownership in the short run.

The partial effects with respect to user costs do not generally become smaller as the number of conditioning factors increase. The largest responses are found in the pooled model and in the model with random effects and state dependence. The most important effect of including random effects and state dependence is that the partial effect for the category holding no car becomes smaller than in the pooled model. The pooled model thus exaggerates the extent of downsizing following an increase in user costs.

The short-run income effects in the random effects model with state dependence are considerably smaller than typically found in other studies, e.g. in studies using synthetic panel data, i.e. panel data constructed from repeated cross sections, like Dargay (2001) and Dargay and Vythoulkas (1999). They found short-run income elasticities (based on macro time series methods) ranging from 0.18 to 0.48. In the same studies, long-run income elasticities range from 0.28 to 0.80. This difference is likely to arise because it is not possible to take into account idiosyncratic effects using synthetic panel data. Also other studies based on micro cross-section data, e.g. de Jong (1990) and Ramjerdi and Rand (1992), have found income elasticities at 0.33 and 0.15, respectively. Previous studies based on Danish data (micro cross section) yielded income elasticities at 0.41, Bjørner (1999) and from 0.39 to 0.55, Fosgerau and Nielsen (2002). As expected these income elasticities are closer to the ones we have found in the pooled multinomial logit model.

The car cost (purchase, ownership and variables costs) responses are largest for the category holding two cars, but the general conclusion is that changes in costs have little impact on car ownership in the short run. The estimated responses appear to be in the same range as those found in other studies. A recent Danish study, Fosgerau et al. (2004) using macro time series found a short run car purchase cost elasticity at -0.19 and a short run variable cost elasticity at -0.22.¹⁰ Dargay (2001) finds a car purchase cost elasticity at -0.13, while Dargay and Vythoulkas (1999) find long-run elasticities with respect to purchase and variable costs at -0.33 and -0.51 (for »middle« levels of income and car ownership). However, it should be recalled that the car cost responses estimated in this study are based on changes in car cost over time for a relatively short period. The estimates could therefore reflect too little variation in the data rather than a genuine behavioural effect, and the estimates should therefore be interpreted with caution.

6 Conclusion

Using a panel data set with information about car ownership for 10,565 households observed over the period 1992-2001 we have demonstrated that car ownership status is very persistent. This shows very clearly at the descriptive level, as 56% of the households did not change ownership status during the ten-year period they were followed.

We estimate models describing ownership status as a function of income, user costs, demographic and geographic characteristics. In the reference model we condition only on the observed characteristics. Next, we allow for unobserved fixed heterogeneity, and finally for both unobserved heterogeneity and lagged ownership status. The two latter models capture idiosyncratic effects, and this can only be done using panel data. The reference model yields estimates that are comparable to those of other studies not based on panel data. The random effects model and the random effects model with state dependence indicate that both unobserved heterogeneity and state dependence are important factors in explaining car ownership in the short run.

The results from these models indicate that responses to changes in income and user costs are much smaller than what was thought based on previous studies. One interesting feature of the dynamic model is that ownership of two cars is more responsive in the short run than ownership of one or no car. This suggests that car holdings of multiple car ownership households respond stronger to changes in incentives. Overall the general conclusion remains, however, that income and user cost changes have little im-

10. It should be noted that the car purchase cost and the variable cost elasticities cannot directly be compared with the user cost elasticity estimated here, as the latter include variable, ownership and purchase cost. However, if it is assumed that variable and purchase cost each accounts for half of the user cost, the corresponding short run user cost elasticity from Fosgerau et al. (2004) would be the sum of the two elasticities (about -0.4).

pact on car ownership in the short run, and that car ownership adjusts very slowly so that policy instruments aimed at reducing car ownership are not likely to be very effective in the short run.

References

- Bento, A. M.; L. H. Goulder, E. Henry, M. R. Jacobsen and R. H. von Haefen. 2005. Distributional and Efficiency Impacts of Gasoline Taxes: An Econometrically Based Multi-market Study. *American Economic Review* (papers and proceedings) 95(2), 282-87.
- Bjørner, T. B. 1999. Demand for Car Ownership and Car Use in Denmark: a micro econometric model. *International Journal of Transport Economics*, 24(3), 377-95.
- Bjørner, T. B. and Leth-Petersen, S. 2005. Dynamic Models of Car Ownership at the Household Level. *International Journal of Transport Economics*, 32(1), 57-75.
- Bunch, D. S. 2000. Automobile Demand and Type Choice. In D. A. Hensher and K. J. Button (eds.): *Handbook of Transport Modelling*. Elsevier.
- Dargay, J. 2001. The effect of income on car ownership: evidence of asymmetry. *Transportation Research Part A*, 35, 807-21.
- Dargay, J. and Vythoulkas, P. C. 1999. Estimation of a Dynamic Car Ownership Model: a Pseudo-Panel Approach. *Journal of Transport Economics and Policy*, 33(3), 287-302.
- de Jong, G. C. 1990. An Indirect Utility Model of Car Ownership and Use. *European Economic Review*, 34, 971-85.
- Erdem, T. and Sun, B. 2001. Testing for Choice Dynamics in Panel Data. *Journal of Business and Economic Statistics*, 19(2), 142-52.
- Fosgerau, M. and Nielsen, J. E. 2002. *Wealth as a Source of Bias in Cross-Sectional Evaluation of Car Ownership*. Discussion Paper (04/07/2002). Danish Transport Research Institute.
- Fosgerau, M.; M. Holmblad and N. Pilegaard. 2004. ART – En aggregeret prognosemodel for dansk vejtrafik. *Notat 5*. Danish Transport Research Institute.
- Heckman J. J. 1981. The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In C. F. Manski and D. McFadden (eds): *Structural Analysis of Discrete Data with Econometric Applications*, pp. 179-95. MIT Press, Cambridge, MA.
- Hensher, D. A. 1985. Empirical Vehicle Choice and Usage Models in the Household Sector: A Review. *International Journal of Transport Economics*, 12(3), 231-59.
- Hensher, D. A.; N. C. Smith, F. W. Milthorpe and P. O. Barnard. 1992. *Dimensions of Automobile Demand – A Longitudinal Study of Household Automobile Ownership and Use*. North-Holland, Amsterdam.
- Kitamura, R. and Bunch, D. S. 1990. Heterogeneity and state dependence in household car ownership: a panel analysis using ordered response probit models with error components. In M. Koshi (ed). *Transportation and Traffic Theory*, pp. 477-96. Elsevier, New York.
- Meurs, H. 1993. A Panel Data Switching Regression Model of Mobility and Car Ownership. *Transportation Research A*, 27A, 461-76.
- Ramjerdi, F. and Rand, L. 1992. *The National Model System for Private Travel*. Institute of Transport Economics report 150/1992, Oslo.
- Train, K. E. 1986. *Qualitative Choice Analysis – Theory, Econometrics, and an Application to Automobile Demand*. The MIT Press, Cambridge MA.
- Train, K. E. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.
- Wooldridge, J. M. 2005. Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics*, 20(1), 39-54.
- Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT press, Cambridge, MA.

Appendix with summary statistics

	All years				1992				2001			
	Mean	std	min	max	Mean	std	min	max	Mean	std	min	max
Pear = 0	0.1573	0.3641	0	1	0.1993	0.3995	0	1	0.1313	0.3377	0	1
Pear = 1	0.7221	0.4480	0	1	0.6980	0.4592	0	1	0.7357	0.4410	0	1
Pear = 2	0.1206	0.3257	0	1	0.1027	0.3036	0	1	0.1330	0.3396	0	1
Inc (DKK)	279,896	116,063	5,959	7,491,460	276,328	136,640	5,959	7,491,460	277,591	117,692	21,207	2,775,631
age	50.6183	14.1114	18	95	46.1183	13.8166	18	86	55.1183	13.8166	27	95
Dehild_m	0.3923	0.4883	0	1	0.4307	0.4952	0	1	0.3370	0.4727	0	1
Dehild_a	0.1341	0.3408	0	1	0.1357	0.3425	0	1	0.1230	0.3285	0	1
Workmale	0.6168	0.4862	0	1	0.6536	0.4759	0	1	0.5715	0.4949	0	1
Workfem	0.5571	0.4967	0	1	0.5715	0.4949	0	1	0.5374	0.4986	0	1
Unempmale	0.0817	0.2739	0	1	0.1217	0.3270	0	1	0.0526	0.2233	0	1
Unempfem	0.1209	0.3260	0	1	0.1818	0.3857	0	1	0.0710	0.2568	0	1
Distm_sr	2.4870	2.6712	0	22.4722	2.6749	2.6277	0	22.2036	2.3343	2.7455	0	22.4722
Distf_sr	1.7918	2.0084	0	22.5389	1.9137	1.9780	0	21.2132	1.7230	2.0496	0	20.0499
Distm_miss	0.02526	0.1569	0	1	0.0219	0.1462	0	1	0.0256	0.1578	0	1
Distf_miss	0.05931	0.2362	0	1	0.0522	0.2223	0	1	0.0544	0.2269	0	1
usc	33.1300	0.8925	32.0190	35.2510	33.1029	0.3634	32.7830	33.6710	34.6288	0.3854	34.2990	35.2500
Cph	0.2701	0.4440	0	1	0.2754	0.4468	0	1	0.2656	0.4417	0	1
Town	0.3692	0.4826	0	1	0.3733	0.4837	0	1	0.3685	0.4824	0	1
Observations	105,650				10,565				10,565			

Debat og kommentarer

A Fable on Wealth and Capital

Niels Blomgren-Hansen

Department of Economics, Copenhagen Business School, *E-mail: nbh.eco@cbs.dk*

On a small sunny island there once lived a community of coconut growers. The islanders had the same simple life as their forefathers and were happy about it. When they were young they climbed the trees and picked coconuts. They consumed half of the coconuts they picked and stored the other half for consumption when they grew old and could no longer climb the trees.

The coconuts they stored for later consumption gradually decayed. In fact, half rotted away, so the number of coconuts for consumption as pensioners was only half of the number of coconuts they had stored as young people. But they didn't complain. They used to tell their children that God had created the island, themselves and the coconut trees. Even though they didn't like to see people grow old and die and the coconuts rot they stuck to the belief that what God had created was perfect.

The island was headed by a wise man called a hogan. The hogan settled conflicts among the villagers, gave good advice, taught the children what was right and what was wrong, and talked to God on behalf of ordinary people who felt a bit uneasy about how to address God. Because of that he was considered a holy man and treated with much respect. He didn't have to work; to show their respect the villagers brought him the coconuts he needed. He lived in a hut somewhat larger than the ordinary villagers. Nobody but the hogan was allowed to enter the hut, because it was considered a holy place.

One day when the hogan sat outside his house thinking about God and the villagers' good life it struck him that the young people stored fresh coconuts and the old people ate old half-rotten coconuts. The decaying of the stored coconuts was clearly a loss to society. If somehow that loss could be avoided everybody would be better off.

God liked that the hogan cared for ordinary people and said to him in a dream, »Go and tell the young people to bring the coconuts to your hut instead of storing them in their own huts. For each coconut you receive you shall give them an IOU and tell the young men that for each IOU they return when they get old they will get a fresh coconut«.

So he did, and the young people found it was a bright idea and looked forward to doubling their standard of living when they got old compared to what used to be the

case. They didn't quite understand how the hogon could keep their coconuts fresh, but the hogon was a holy man and they didn't doubt his word.

When the time came he delivered as he had promised. Every day a flow of young people delivered fresh coconuts to the hogon's hut and got IOUs in return, and an equally large flow of old people came to the hut and got a fresh coconut in return for each one of the IOUs they had collected as young people.

So life continued for many generations. But as time passed people's belief in God and the hogon's holiness weakened and one day a young man said, »How can I be sure that the hogon can and will fulfil my IOUs when I grow old? I will go and see if he really has a pile of fresh coconuts in his hut«. And so he did. He forced his way into the hogon's hut and found that there wasn't a single coconut.

He got angry and called all the villagers – young and old – to a general meeting. Here he told them that the hogon had cheated them, that their claims were just pieces of paper not backed by a single coconut, and that, consequently, they could never be fulfilled. The fraud had made the village poorer. In the good old days they owned a large stock of coconuts. Now they had nothing but a pile of paper notes.

The hogon was called to the meeting and asked for an explanation. He had none. When he took over the post as hogon there were no coconuts in the hut and, to his knowledge there had not been for generations. What actually happened to the originally collected coconuts he couldn't tell. Some villagers suggested that the old hogon might have eaten the coconuts himself, other that the coconuts had simply rotted away, and those villagers who respected the office of the hogon the most thought that the old hogon was a good man and had given the coconuts to the poor.

The angry young man requested that the fraud should be stopped, the hogon driven out of the village and the villagers return to individual coconut storing and self-support.

But an old sober-minded man stepped in and said, »Listen, let us cool down. Sure, it was also a shock to me to hear that our IOUs have no backing and haven't had for generations. But, nevertheless, the practice has worked fine. To give it up will be a disaster. The old people will have nothing to live on and the young will find their standard of living as old halved. So let us continue saving for our retirement by selling coconuts to the hogon for IOUs as we use to do. After all the hogon has always fulfilled the IOUs we have returned, so why shouldn't he do so in the future?«

The villagers were convinced and returned to their daily work.

After the meeting the hogon asked the angry young man to come to his hut for a private talk. »You are a smart young man and I understand your logic when you say that we have all become poorer«, he said. »The villagers own IOUs but these IOUs correspond exactly to my debt or rather to the debt of my office, so on balance we own

nothing. But, nevertheless, in my mind it makes little sense to say that we have all become poorer, when in fact we all feel wealthier and have a better life. I think that the appropriate notion of wealth should reflect our standard of living and include not just the value of physical coconuts but also the value of virtual coconuts – i.e. the IOUs – that the villagers have collected. Or – to put it differently – when accounting the village's wealth we should not subtract the debt of my office«.

»No«, answered the young angry man, »that would be wrong. One day – maybe in a distant future – the debt must be repaid, and as you own nothing, you or your successor will have to tax the villagers. So, even though the villagers don't feel so, your debt is in fact the villagers' debt. They might feel wealthier, but that is just an illusion. Illusions – virtual coconuts – can never count for real wealth.«

The hogon smiled. »Listen, the debt collected by the hogons in the past corresponds to one half of one generation's total production. Assume that we decide that it should be repaid in this period. How could that be achieved? We might renege on the outstanding debt. That would rob the old generation of their means of living. We might instead tax the young generation, i.e. force them to continue to deliver coconuts to my hut without getting anything in return. As the young will have to care for their retirement by stockpiling coconuts and half of the stock will rot, the young generation's consumption over their entire lifetime will be just $\frac{3}{8}$ of what it is today and $\frac{1}{2}$ lower than the level of consumption that prevailed before the old hogon started issuing IOUs long time ago. No matter what, it would be a disaster for the villagers and to nobody any good.«

The young man nodded; the hogon clearly had a point. »But«, the hogon continued, »if we must conclude that it is a bad idea to repay the debt now, why should it be a good idea to repay the debt in the future? If future generations are as wise as we are they will just go on rolling over the debt as we have done and never repay. Is a debt that never has to be repaid really a debt? In my view it is a gift from God just as this island and the coconut trees are gifts from God. I don't think we should subtract any of God's gifts when we measure our wealth«.

The young man still didn't like the idea of counting virtual coconuts – pieces of paper – as wealth. But he didn't know what to answer and left the hogon's hut.

Some time after this talk the young man had a dream himself. Whether it came from God or the Devil he couldn't tell. He saw himself easily climbing a coconut tree along something he had never seen before – he later called it a ladder. He thought that climbing a tree by a ladder looked smart and would impress the women, so he decided to use part of his time building a ladder – to be precise, one quarter of his time. Consequently, he picked only three quarters of the coconuts he used to pick and as he consumed as many as before, he only delivered half as many – one quarter of his harvest –

to the hogon's hut. But he wasn't worried. He reasoned that when he got old he could sell his ladder to a young man in return for a constant delivery of coconuts for the rest of his lifetime, thereby making up for the fact that he had delivered fewer coconuts to the hogon and consequently had fewer IOUs. And he was right. He had no difficulty in selling his ladder and his standard of living as an old man was unaffected. As time passed more and more young men copied his idea and eventually all had ladders.

But the hogon got in deep troubles. When the young men started to deliver fewer coconuts the hogon could no longer fulfil his obligations towards the old people. »Sorry«, he said, »but I can only pay you a fraction of a coconut for each IOU you return to me«. They got angry, but could do nothing about it. And the fraction of a coconut they got in return for each IOU continued to fall until all the young men had got a ladder. At that time it had fallen to one half.

The event caused a lot of palaver at the general meeting. Naturally, the old people felt cheated. But they were sophisticated and instead of complaining over their own fate they argued that building ladders had reduced the production of coconuts and, consequently, had been bad for the island as a whole.

However, they were silenced by a young man who was so sharp and eloquent, that the villagers called him Diamond.

»First«, Diamond argued, »total production has not fallen. It is true that we have produced fewer coconuts, but instead of coconuts we produced ladders, i.e. we invested in capital goods. The value of the ladders measured in coconuts corresponds exactly to the coconuts we could have produced, so total production remained constant. Production that is not consumed is saved. The more we save, the more we invest, and the more we invest the larger our stock of capital will grow«.

»Second«, he went on, »it is true that the purchasing power of your IOUs has shrunk. But IOUs are not part of real wealth. They are just one hand's outstanding debts to the other. It is only real capital that counts, so it is only through investment that we can get rich. Today the village has a stock of real capital corresponding to one quarter of one generation's production. Before we started investing in ladders, we had none. So you must admit that the village as such has become richer.«

The hogon didn't argue against Diamond. The villagers's respect of the hogon had dwindled with the value of the IOUs and they wouldn't listen to him. He just shook his head. But back in his hut he talked to God and said, »They are simple minded people. They don't understand what they don't see. They don't realize that you created their wealth by a miracle by telling them to issue IOUs. They have not grasped that your gift is a grace that cannot be enlarged by building tangible symbols of wealth. The more tangible capital they collect the more of your undeserved intangible gifts they destroy. Forgive them their ignorance and be patient. One day they will understand and turn their back on Diamond and the idolatry, he preaches«.

It so befell as he had predicted. After some time the young men changed their mood. Using ladders became regarded as old-fashioned and lax. Climbing the trees without ladders exposed the young men's muscles and strength and both the young men and the young women liked that. So the young men didn't any longer want to buy ladders from the old people, and the price of ladders in terms of coconuts fell to zero. Instead they brought twice as many coconuts to the hogon's hut. For each coconut they got two IOUs as this had become the current price«.

The hogon now got twice as many coconuts as he needed to meet the old people's demand. He soon considered reducing the price of coconuts to one IOU. That would make it possible for the old people to maintain their standard of living and would not hurt the young men. But he skipped the idea; it would be too difficult to explain to the villagers. So he kept the prevailing price of coconuts unchanged – two IOUs for one coconut. Instead he proclaimed a welfare program: Due to the mercy of God the hogon would grant compensation to the old people for the loss in wealth they had suffered because the failing demand for ladders had made their capital worthless. And all the villagers praised God and the hogon for his charity.

The old and now useless and worthless ladders lay idle on the island. Most islanders considered them junk and a disfiguration of their beautiful island. So the hogon told them to collect the ladders and place them in a pile on the beach for the children to play with. Here they gradually disintegrated. No capital good is so durable that it can stand the wear of playing children.

Only two ladders were kept for the future. They were exhibited in front of the hogon's hut in what he called the »Museum of Capital Idolatry« as a warning to future generations and in memory of the people who had suffered from Diamond's false teaching. – And here they can still be seen, if nobody has removed them.

hvori det antydedes, at man i visse udenlandske undersøgelser havde fundet en positiv effekt af etniske ghettoer, se Christiansen & Schmidt (2002).

Kapitel 1 af alle tre forfattere er en bred oversigt over udviklingen i Danmark af de relevante grupper og en historisk oversigt over brugen af ghettobegrebet. Kapitel 2 og 3 af Marie Louise Schultz-Nielsen er en analyse af bosætnings- og flyttemønstre i Danmark 1985-2004. Her etableres et helt nyt data-materiale, hvor Danmark opdeles i boligområder på 150-600 familier; problemstillingen er, at en traditionel inddeling i kommuner er for grovmasket et net til at identificere f.eks. etniske enklaver (forfatterne foretrækker betegnelsen »etniske enklaver« i stedet for »ghettoer«, da ghetto-begrebet har en anden mening i amerikansk litteratur). Grundlaget er en opdeling af Danmark i kvadrater på 100 gange 100 m. Sådanne kvadrater indeholder imidlertid for manges vedkommende så få familier, at det ville stride mod Danmarks Statistiks diskretionskrav at offentliggøre tal for sådanne kvadrater; de er derfor slået sammen, så man får så homogene boligområder som muligt med over 150 familier. Udviklingen måles dernæst bl.a. ved et dissimulationsindeks D ,

$$D = \frac{1}{2} \sum_i |m_i - f_i|$$

hvor m_i er den andel af et mindretal, der bor i område i , og f_i er den andel af flertallet, der bor i dette område. Hvis mindretallet er fordelt på samme måde som flertalsbefolkningen, bliver indekset 0, hvis mindretallet og flertallet er helt separeret vil indekset blive 1. Undersøgelsen viser, at de etniske grupper er meget mere koncentreret og separeret fra flertallet end mindretallene af overførselsindkomstmodtagere, lavt uddannede og lavtlønnede. Det analyseres også, hvilke karakteristika, der er kendetegnende for beboerne i de forskellige boligtyper, og hvad der får folk til at flytte.

Kapitel 4 og 5 af Anna Piil Damm analyserer, hvad det betyder for ens chancer i livet, at bo i et boligkvarter med mange lavt uddanne-

de eller indvandrere. Hovedproblemstillingen er, om en etnisk enklave er befordrende for indbyggernes succes i samfundet eller det modsatte. Her er der i teorien modgående tendenser – indvandrere i en enklave har et tættere netværk af landsmænd, der kan støtte og hjælpe dem, men der er også en risiko for, at de derved isolerer sig fra det øvrige samfund. Problemet med empirisk at undersøge, hvilken af disse effekter der er stærkest, er, at der er potentielt stærke selektionsskævheder. Hvis beboerne i en etnisk enklave klarer sig dårligere end indvandrere med samme egenskaber (uddannelse, alder osv.) uden for enklaverne, så kan det være, fordi enklaverne hæmmer beboernes integration i det danske samfund, men det kan også være, fordi indvandrere med ikke-observerbare svagheder (usikkerhed, skepsis over for det danske arbejdsmarked o.l.) vælger at slå sig ned i enklaverne, mens dem med mere gå-på-mod vælger at bo uden for enklaverne.

I kapitel 6 af Anna Piil Damm og Torben Tranæs analyseres dette problem på kommunal for flygtninge. Flygtninge er nemlig fordelt og kan ikke selv vælge bopælskommune. Selektionsskævheden eksisterer derfor ikke for denne gruppe. Resultaterne indikerer, at beboerne i enklaverne klarer sig dårligere end beboere med samme observerbare karakteristika uden for enklaverne, når man ser på beskæftigelse og løn. Men renser man for selektionsskævheden, synes enklaverne og deres netværk at bidrage positivt til beboernes succes på arbejdsmarkedet, vel at mærke hvis der er mange landsmænd, mens mange af andre indvandrergupper trækker i negativ retning. Når de etniske enklavers beboere klarer sig dårligere på arbejdsmarkedet end dem, der bor andre steder, synes det således at skyldes beboernes personlige egenskaber, mens selve det at have et nærmiljø med mange landsmænd bidrager positivt til, at man klarer sig godt.

Kapitel 7 af alle tre forfattere er en kort sammenfatning. Til sidst i bogen er der tre kommentarer til undersøgelserne: syv sider fra Rikke Hvilshøj, et tilsvarende antal sider fra Henrik Sass Larsen og endelig 12 sider fra den norske indvandrerforsker Svein Blom.

Hvor der altså ikke er nogen tvivl om, at bogen bringer nyt interessant stof, så kan forment meget vel diskuteres. Målgruppen for en bog som denne er lidt uklar. Det kan godt diskuteres, om resultaterne endnu er så gennemdiskuterede blandt fagfolk, at de er modne til formidling til beslutningstagere og offentlighed. Jeg nærer selv en betydelig tvivl om dissimilationsindeksets egenskaber. Udgangspunktet i form af numeriske forskelle er, hvad man har brugt i den statistiske teori vedrørende »robust« estimation for at undgå, at store afvigelser fra gennemsnittet får for stor vægt, se f.eks. Kærgård (1989). Men når man ser på dissimilation, er man vel netop interesseret i de mest udsatte områder? Det valgte indeks er kun afhængigt af, hvor stor en andel af indvandrere, der bor i de områder, hvis andel af indvandrere er større end deres andel af danskerne. Indekset ændrer sig ikke, hvis indvandrere flytter rundt mellem områder med overgennemsnitlig indvanderandel, blot de bliver ved med at være overgennemsnitlige (det svarer helt til, at medien er uændret, selv om alle de største observationer nedsættes til 51%-fraktilen). Der findes næppe mange, om nogen, danskere, der har så megen erfaring med sådanne indeks, at de for alvor kan tolke dem helt driftssikkert. Man må tænke på, hvor meget erfaring det kræver at bruge korrelationskoefficienter på en vederhæftig måde, når stikprøver af forskellig størrelse og beskaffenhed skal sammenlignes. Hvorfor prøver man i øvrigt ikke med den simple korrelationskoefficient mellem m_i og f_i ? Mon den vil give samme resultater? Herved ville man få lagt større vægt på de særligt ekstreme områder, fordi der, når det er kvadrerede afvigelser og ikke numeriske, der indgår, lægges langt større vægt på de mest ekstreme observationer.

Et tilsvarende specifikt problem peger Svein Blom på. Spørgsmålet er, om den positive effekt af etniske enklaver måske kun er en korttidseffekt. Måske giver netværket i enklaven bedre mulighed for beskæftigelse på kort sigt f.eks. ved kontakt til typiske indvanderjobs, men det er måske på bekostning af dansk kundskaber, kendskab til det danske

samfund og kontakt til det almindelige arbejdsmarked, som på langt sigt er central for succes.

Det skal ikke skjules, at jeg nok havde foretrukket, at forfatterne først havde publiceret deres forskningsresultater i fagtidsskrifter, og så, når de var nogenlunde gennemdiskuterede, havde skrevet en mere afklaret formidlende debatbog med henblik på politiske beslutningstagere. De i Svein Bloms indlæg givne sammenligninger med Norge og hans faglige kommentarer er meget givende; de viser potentialet for de positive effekter, man kunne få ud af en international debat i artikelform om de rejste – vigtige – spørgsmål.

Nu prøver man at kortslutte processen ved at gå direkte til »debatbogen«, og så sikre sig en politisk reaktion med det samme ved at påhæfte kommentarer fra en politiker fra hver side. Men det kommer der selvfølgelig ikke noget ud af; Rikke Hvilshøj skriver om regeringens politik ret løsrevet fra bogen i øvrigt, og Henrik Sass Larsens konklusion er fornuftigvis, at nogle af konklusionerne synes at være spændende og overraskende, men »først og fremmest får (de) mig til at efterlyse endnu mere forskning på området« før politiske konklusioner drages. Det er et stort spørgsmål, om det længere er muligt at nå ud til både forskningsmiljøerne, den dannede offentlighed og beslutningstagerne med en og samme publikation. I vore dage er man nok nødt til at målrette kommunikationen langt mere – det er ikke længere de samme, der læser kronikker og fagøkonomiske tidsskrifter.

Hvis sigtet er den velorienterede beslutningstager, embedsmand og rådgiver, er det lidt overflødigt med de omfattende indledninger med opsummeringer af velkendt stof om dansk indvander- og boligpolitik. Redaktøren synes at have glemt den fra Socialkommissionen i begyndelsen af 1990'erne kendte »Aase Olesens lov«, der siger, at for samfundsorienterede rapporter er summen af antal sider og antal læsere altid lig 500. Med en noget hårdere redigering og en udeladelse af politikerkommentarerne kunne man så have fået godt 100 flere læsere, og det fortjener bogens problemstillinger og resultater.

Litteratur

- Christiansen, C. Carøe & G. Schmidt. 2002. *Man-ge veje til integration. Resultater og perspektiver fra Socialforskningsinstituttets forskning om etniske minoriteter*, Socialforskningsinstituttets rapport 02:25.
- Kærgård, N. 1989. Estimation criterion, residuals and prediction evaluation, *Computational Statistics & Data Analysis*, vol. 5 s. 443-50.

Mobilitet af håndværksydelser i Danmark

Nikolaj Malchow-Møller

CEBR og University of Southern Denmark, E-mail: nmm.cebr@cbs.dk

Jakob Roland Munch

University of Copenhagen, E-mail: jakob.roland.munch@econ.ku.dk

Jan Rose Skaksen

Copenhagen Business School og CEBR, E-mail: jrs.eco@cbs.dk

SUMMARY: In this paper, we analyse the mobility of construction services across different regional markets in Denmark. Although the markets appear to be segmented – both in terms of production and price changes – we show that a considerable amount of trade is taking place between regions. This is done through the use of unique Danish data that allow us to estimate the extent and character of this interregional trade. Furthermore, we analyse the characteristics of the individuals and firms exporting these services. It turns out that the level of export is particularly high in the Western part of Denmark. In addition, it appears to be the most productive individuals in the most productive – or the most specialised – firms who export the most. Finally, we find no evidence of distortionary effects of the existing tax subsidies to this export.

1. Introduktion

I dette papir analyserer vi mobiliteten mellem de regionale markeder for håndværksydelser i Danmark. Dette gøres i kraft af helt unikke danske data, der tillader os at opgøre omfanget og karakteren af handlen med håndværksydelser på tværs af regionale markeder. Spørgsmålet om, hvor mobile/handlede sådanne ydelser egentlig er, er i stigende grad blevet relevant – og er kommet højt på den politiske dagsorden – med den seneste EU-udvidelse, der åbner op for denne type af handel med de østeuropæiske lande.

Det er velkendt, at der eksisterer store prisforskelle på håndværksydelser – og serviceydelser mere generelt – på tværs af forskellige geografiske regioner, f.eks. i

Dette papir er en del af et fælles projekt mellem CEBR og Rockwool Fondens Forskningsenhed. Vi vil gerne rette en særlig tak til Rockwool Fonden for finansiel støtte til projektet. Vi vil også gerne takke Torben Tranæs og to anonyme referees for mange konstruktive kommentarer og forslag. Vibeke Borchsenius og Jonas Helth Lønborg takkes for fremragende forskningsassistance.

Danmark i forhold til i Polen, men også i mindre grad mellem f.eks. København og Sønderjylland. Ifølge den gængse handelsteori burde sådanne forskelle helt (eller delvist) elimineres som følge af handel med disse ydelser og muligvis også via permanent reallokering af produktionsfaktorer, idet disse må forventes at søge hen, hvor aflønningen er størst, se f.eks. Leamer og Levinsohn (1995). Den internationale handel med serviceydelser er dog langt mindre end den tilsvarende handel med varer, se Lipsey (2006).¹

En oplagt forklaring på den mindre handel med serviceydelser er, at der internationalt har været en række restriktioner på denne handel. Kun inden for EU er der således fri bevægelighed af produktionsfaktorer og fri handel med f.eks. håndværksydelser på tværs af lande. Med udvidelsen af EU sker der dog en åbning på dette område (og en ikke uvæsentlig én af slagsen), idet der med de nye landes tilgang opstår markant større prisforskelle end hidtil mellem lande inden for unionen.

Inden for et lands grænser har der ikke været den slags hindringer. Men der er forhold, der bevirker, at handel med håndværksydelser alligevel må forventes at adskille sig fra traditionel handel med varer, se van Welsum (2003). For det første, fordi selve arbejdskraften, der skal producere ydelser, i en periode skal flyttes med. Dette giver helt anderledes og sandsynligvis større handelsomkostninger i form af leveomkostninger end ved handel med traditionelle varer. For det andet, fordi det kan være relativt svært for kunderne at få information om kvaliteten af ydelser, inden den er leveret. Denne asymmetriske information må formodes at mindske potentialet for handel.

Umiddelbart vil man derfor forvente, at der selv inden for Danmarks grænser er mange relativt lokale eller segmenterede markeder for håndværksydelser. Dette gør det dog samtidig interessant at se på den handel, der trods alt foregår mellem de regionale markeder. Hvad er omfanget af den, og hvem er det, der »eksporterer« disse ydelser? Med andre ord, hvor mobile/handlede er håndværksydelser i Danmark? I dette papir udnytter vi eksistensen af helt særlige danske data til at få et indblik i dette fænomen, der pga. manglende data er meget sparsomt belyst i litteraturen, se Lipsey (2006).

I den første del af dette papir kigger vi nærmere på forudsætningerne for interregional handel med håndværksydelser i Danmark. Dvs., hvad kan motivere en sådan handel, og er der umiddelbare tegn i produktions- og prisdata på, at disse forudsætninger

1. WTO skelner mellem fire typer (»modes«) af international handel med serviceydelser, se f.eks. Bhagwati, Panagariya og Srinivasan (2004): »Mode 1«, hvor en forbruger i land A modtager en service fra land B via post eller telekommunikation. Eksempler er call centre og software programmering. »Mode 2«, hvor forbrugere fra land A forbruger serviceydelsen direkte i land B. Dette dækker f.eks. turisme. »Mode 3«, hvor et firma fra land B sælger servicen i land A via en fast kommerciel repræsentation i land A. Denne type involverer således et element af direkte udenlandske investeringer fra land B til land A. »Mode 4«, hvor en person fra land B leverer serviceydelsen i land A – enten som udsendt fra et firma i land B eller ved at tage midlertidigt arbejde i et lokalt firma i land A. Handel med håndværksydelser tilhører denne sidste kategori.

er til stede, og at handel derfor finder sted? I analysens anden del udnytter vi data om udbetaling af skattefri rejsegodtgørelse og fradrag på selvangivelsen i forbindelse med udstationering til mere direkte at estimere omfanget og karakteren af denne handel på tværs af regioner. Vi undersøger også, hvad der karakteriserer de firmaer og personer, der eksporterer disse ydelser.

I analysens første del finder vi, at der er visse tegn på, at interregional handel med håndværksydelser i Danmark finder sted, selvom den relativt jævne geografiske fordeling af produktionen ikke umiddelbart tyder på, at handel kan være motiveret af regionale stordriftsfordele i form af klynger. Til gengæld eksisterer der betydelige regionale prisforskelle, der kan danne basis for handel mellem regionerne på trods af de høje handelsomkostninger. Det viser sig f.eks., at lønnen til en gennemsnitlig håndværker er knap 20% højere i København end i Sønderjylland – en forskel der har været stort set konstant i perioden 1993 til 2004. Vi finder dog også, at sammenhænge i lønudviklingen for geografisk mere adskilte markeder er svagere, hvilket peger i retning af, at handlen ikke er så udbredt, at de regionale markeder kan opfattes som fuldt integrerede.

Anvender vi vores nye og mere direkte estimater af den interregionale handel, finder vi imidlertid en betydelig handel med håndværksydelser. Der foregår handel ud af både højt lønsområder og lavtlønsområder, hvilket tyder på, at produktdifferentiering forklarer en del af denne handel. Dette understøttes yderligere af, at det er de mest produktive personer i de mest produktive virksomheder – eller de virksomheder, der er bedst til at produktdifferentiere – der især deltager i den interregionale handel.² Det er dog især håndværkere fra Jylland og Fyn, der sælger til kunder uden for lokalområdet. Dvs. det er især håndværkere, der er lokaliseret i områder, hvor lønningerne er relativt lave, der eksporterer ud af lokalområdet. Under visse forudsætninger finder vi f.eks., at jyske og fynske murere og tømrere i 2004 leverede mellem 6,7% og 20,6% af de håndværksydelser, der blev solgt i København. Den anden vej er handlen betragtelig mindre, idet håndværkere fra Sjælland og Lolland-Falster i 2004 leverede mindre end 2% af håndværksydelserne i Jylland og på Fyn.

Der er ikke noget der tyder på, at den skattefri rejsegodtgørelse og fradraget på selvangivelsen overkompenserer håndværkere for deres omkostninger ved at være udstationeret. Der er til gengæld heller ikke noget, der tyder på, at virksomheder betaler deres medarbejdere ekstra for at være udstationeret. Håndværkere, der eksporterer, tjener typisk mere, mens de er udstationerede, men dette synes primært at være et resultat af mere overarbejde og højere kvalifikationer.

Den resterende del af papiret er opbygget som følger: Afsnit 2 præsenterer data, mens vi i afsnit 3 belyser forudsætningerne for og indikationerne på interregional handel med

2. Dette resultat er helt analogt til, at Munch og Skaksen (2006) finder, at det inden for industrien er de mest produktive virksomheder – eller de virksomheder, der er bedst til at produktdifferentiere – der eksporterer.

håndværksydelse i Danmark. I afsnit 4 estimerer vi omfanget af denne handel, mens vi i afsnit 5 undersøger karakteren af den. Endelig ser vi i afsnit 6 på sammenhængen mellem udstationering og aflønning. Afsnit 7 konkluderer.

2. Data

Vores primære datakilde er Integreret Database for Arbejdsmarkedsforskning (IDA) fra Danmarks Statistik. IDA indeholder data på individniveau bla. omkring arbejdsmarkedsforhold (løn, beskæftigelse, m.m.) og personlige baggrundskarakteristika (alder, uddannelse, familieforhold, m.m.). Individuer er desuden koblet til arbejdssteder, så data kan aggregeres på arbejdsstedsniveau, og information om arbejdsstedet kan anvendes på individniveau.

Vi har endvidere data på individniveau for udbetaling af skattefri rejsegodtgørelse og fradrag på selvangivelsen for ekstraomkostninger i forbindelse med arbejde på et midlertidigt arbejdssted (udstationering). Disse data kan bruges til at opgøre omfanget af »udedage« og forklares yderligere i afsnit 4. Vi anvender data for årene 1993-2004, idet oplysninger om skattefri rejsegodtgørelse og fradrag for udstationering kun er tilgængelige for disse år.

Vi definerer håndværksydelse som anvendelsen af håndværkere inden for bygge- og anlægsbranchen. I analyserne skelner vi mellem følgende syv typer af håndværkere: Tømrere, murere, snedkere, malere, elektrikere, smede, og øvrige bygge- og anlægsarbejdere. Håndværkertyperne er defineret ud fra deres uddannelsesniveau, og vi begrænser os til dem, der er ansat i bygge- og anlægsbranchen.

For at analysere de lokale markeder for håndværksydelse inddeler vi Danmark i lokalområder, hvor vi anvender de gamle amter som afgrænsninger. Dog slår vi Københavns Amt sammen med Københavns og Frederiksberg Kommuner og betegner dette område København. Dette giver os ialt 14 regioner/lokalområder.

3. Forudsætningerne for interregional handel med håndværksydelse i Danmark

Formålet med dette afsnit er at belyse forudsætningerne for interregional handel med håndværksydelse i Danmark. Er der tegn på, at interregional handel kan finde sted – og finder sted – når vi betragter forskelle i produktion og priser på tværs af danske regioner?

Handel med håndværksydelse kan opstå af flere grunde. For det første, hvis bestemte typer af håndværkere er koncentreret geografisk, vil det være tegn på eksport af håndværksydelse fra områder med relativt mange håndværkere af en bestemt type til områder med relativt få håndværkere af denne type. Dette kunne f.eks. være tilfældet, hvis der er stordriftsfordele i produktionen. Vi vil derfor undersøge, hvorvidt produk-

tionen af forskellige håndværksydelser er karakteriseret ved at foregå i klynger, eller den er spredt forholdsvis jævnt ud geografisk.

For det andet, kan der være så store lønforskelle mellem regionerne, at det kan betale sig for højt-lønsregioner at købe håndværksydelser fra lavtlønsregionerne på trods af de forholdsvis høje handelsomkostninger, der er forbundet med levering af netop håndværksydelser. Derfor undersøger vi, i hvor høj grad lønninger til håndværkere varierer over de forskellige regionale markeder. Vi analyserer også sammenhænge i lønudviklinger på tværs af de forskellige regioner for at afgøre, hvor integrerede de regionale markeder er. Dette giver en yderligere indikation af, hvor mobile (eller handlede) håndværksydelser potentielt er på tværs af regionale markeder.

For det tredje kan håndværksydelser selv inden for en given håndværkstype være differentierede, hvilket kan give anledning til, at samme region både importerer og eksporterer f.eks. tømrerydelser, fordi virksomhederne specialiserer sig i produktionen af visse typer af disse.³ Data giver os ikke direkte mulighed for at belyse en sådan specialisering i produktionen i de forskellige regioner. Vi kan dog indirekte få en indsigt i betydningen af dette forhold, når vi analyserer omfanget og retningen af handel mellem regionerne i afsnit 4, samt når vi analyserer, hvad der karakteriserer de virksomheder og håndværkere, der deltager i den interregionale handel i afsnit 5 og 6.

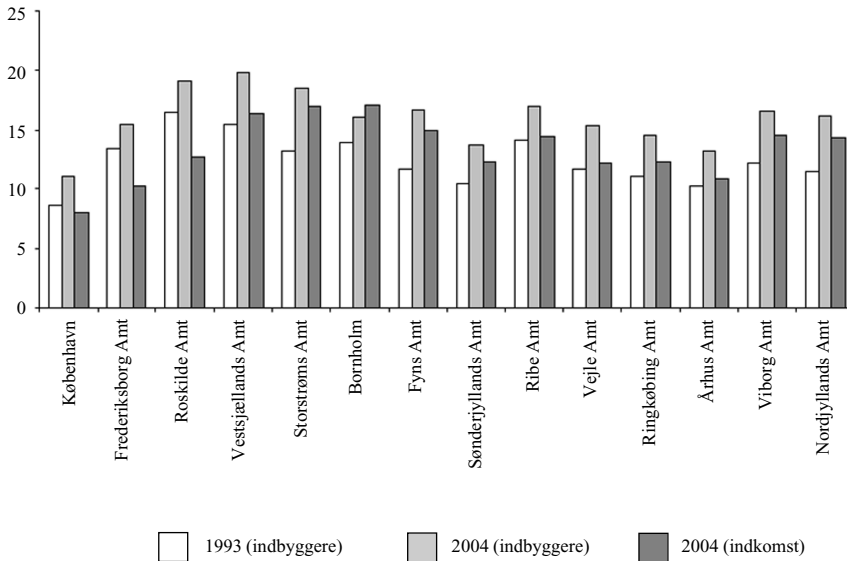
Endelig kan offentlige subsidier til handel med serviceydelser som f.eks. de danske muligheder for skattefri rejsegodtgørelse og fradrag på selvangivelsen for ekstraomkostninger i forbindelse med udstationering bidrage til at skabe grundlag for (tovejs) handel med håndværksydelser – selv i en situation hvor ingen af de ovennævnte forudsætninger for handel er til stede. Vi vender tilbage til dette aspekt senere.

3.1 Geografisk koncentration

Hvis håndværkerne er lokaliseret i geografiske klynger – eller måske kun et enkelt sted – er det et stærkt tegn på interregional handel med håndværksydelser. Er håndværkerne omvendt forholdsvis jævnt fordelt i Danmark, er det en indikation af, at produktionen og leveringen af ydelsen foregår lokalt. Det er kendt fra studier af USA, at produktionen af visse serviceydelser (f.eks. forsikring, underholdning og finansielle ydelser) har karakter af at være meget geografisk koncentreret, mens andre er mere jævnt fordelt, se Krugman (1991). Ved at betragte forskellige mål for den geografiske koncentration af håndværkere forsøger vi derfor at få indblik i hvilke håndværksydelser, der er mest tilbøjelige til at blive handlet på tværs af regioner.⁴

3. Dette forudsætter dog typisk også, at der er stordriftsfordele forbundet med produktionen af de forskellige tømrerydelser.

4. Med en lignende tilgang undersøger Jensen og Kletzer (2005) hvilke stillingskategorier i servicesektoren, der er mest udsat for at blive outsourcet.



Figur 1. Antal håndværkere i bygge- og anlægsbranchen pr. 1000 indbyggere og pr. 100 millioner kr. lønindkomst.

Note: Antal håndværkere er opgjort som antallet af fuldtids- og deltidsbeskæftigede håndværkere ansat i bygge- og anlægsbranchen i et givet lokalområde i november måned det pågældende år. Antal indbyggere er fundet på www.statistikbanken.dk som antallet af personer i et givent amt pr. 1. januar det givne år. København består af Københavns og Frederiksberg kommuner samt Københavns Amt. Indkomst består af lønindkomst, jubilæumsgratiale og diverse vederlag (opgjort som »løn mv.« på www.statistikbanken.dk).

Figur 1 viser antallet af håndværkere beskæftiget inden for bygge- og anlægsbranchen i de forskellige lokalområder i forhold til indbyggertallet i området. Vi ser, at der er en ret jævn fordeling af håndværkere, der arbejder inden for bygge- og anlægssektoren i de forskellige lokalområder. Dette er tegn på, at håndværksydelser produceres og leveres lokalt, fordi de har den særlige karakter, at håndværkeren typisk »følger med« leveringen af ydelsen, og der derfor er forholdsvis høje handelsomkostninger.

Et alternativt mål for omfanget af den lokale efterspørgsel er den samlede lønindkomst i et område. Figur 1 viser derfor også antallet af håndværkere i forhold til lønindkomsten i et område i 2004. Når lønindkomsten anvendes i stedet for indbyggertallet, ser der ud til at være en anelse større variation på tværs af lokalområder.

Figur 1 dækker dog over en del (mindre) forskelle mellem de forskellige håndværkertyper. Dette kan illustreres ved at se på variationen over lokalområder i antallet af de syv forskellige håndværkertyper pr. 1000 indbyggere. Tabel 1 præsenterer variationskoefficienterne for 1993, 1998 og 2004:

Table 1. Variationskoefficienter for antal håndværkere pr. 1000 indbyggere på tværs af lokalområder.

	Murere	Tømrere	Snedkere	Malere	Elektrik.	Smede	Øvrig B&A	I alt
1993	0,189	0,249	0,299	0,113	0,147	0,247	0,226	0,125
1998	0,198	0,164	0,224	0,085	0,130	0,181	0,199	0,083
2004	0,185	0,176	0,192	0,085	0,103	0,172	0,172	0,092

Note: Antal håndværkere er opgjort som antallet af fuldtids- og deltidsbeskæftigede håndværkere ansat i bygge- og anlægsbranchen i november måned det pågældende år. Variationskoefficienterne er for hver af de pågældende håndværkertyper defineret som standardafvigelsen divideret med gennemsnittet og er opgjort på tværs af lokalområder. Antal indbyggere er fundet på www.statistikbanken.dk som antallet af personer i et givent amt pr. 1. januar det givne år. København består af Københavns og Frederiksberg kommuner samt Københavns Amt.

Vi ser, at den mindste variation findes for malere og elektrikere, mens en større geografisk variation findes for snedkerne. Det er i øvrigt interessant at bemærke, at der ikke er nogen klar tendens til, at spredningen ændrer sig over tiden.

At spredningen er lidt større inden for nogle håndværksgrupper end inden for andre kunne være en første indikation af, at der foregår mere interregional handel med de typer af håndværksydelser, hvor spredningen er størst. Det behøver dog ikke at være tilfældet. Man kan f.eks. forestille sig, at der er mere efterspørgsel efter murere i områder, hvor der er meget nybyggeri, mens der er mere efterspørgsel efter snedkere i områder, hvor der foregår en del renovering af ældre huse.

3.2 Lønninger og lønudvikling

Selv i en situation med høje handelsomkostninger vil betydelige prisforskelle på håndværksydelser mellem regionale markeder kunne skabe grundlag for en vis handel med disse ydelser fra lavtlønsområder til højtølsområder.

Som mål for de regionale priser på håndværksydelser kan vi benytte gennemsnitstimelønningerne, idet arbejdskraft udgør langt det væsentligste input, og eftersom systematiske forskelle i produktiviteten på tværs af regioner må formodes at være små. I figur 2 er vist timelønningerne for en gennemsnitlig håndværker i de forskellige regioner i forhold til landsgennemsnittet.

Vi ser, at der er betydelige forskelle på timelønningerne i de forskellige områder af Danmark. De højeste timelønninger findes i København, og generelt er lønningerne højere på Sjælland end i de øvrige områder. De laveste lønninger findes på Bornholm og dernæst i Sønderjylland. I 2004 var lønnen til en gennemsnitlig håndværker således næsten 20% højere i København end i Sønderjylland. Det er desuden bemærkelsesværdigt, hvor konstante lønforskellene har været. Hvis vi f.eks. tager forskellen på



Figur 2. Lønninger for en gennemsnitshåndværker i lokalområder.

Note: Lønnen er beregnet som timelønnen for en gennemsnitlig fuldtidsansat håndværker ansat i et givet lokalområde. Vægtene af de 7 håndværkertyper er givet ved deres andele af den samlede beskæftigelse inden for bygge- og anlægsbranchen i 2004. Den gennemsnitlige timeløn i hele Danmark er i hvert år indekseret til 1.

lønningerne i København og i Sønderjylland, var det også i 1993 tilfældet, at lønningerne var ca. 20% højere i København end i Sønderjylland.

Hvis man ser på de forskellige håndværkertyper, er billedet stort set det samme. Set over hele perioden er der således en ikke ubetydelig forskel på lønningerne mellem de forskellige områder, og den ser ud til på lang sigt at holde sig ret konstant. Det kunne således tyde på, at der er et vist prismæssigt grundlag for handel med disse ydelser – i hvert fald fra Vest- til Østdanmark. Dette kan i særlig grad være tilfældet, når der som i Danmark ydes skattemæssig kompensation for udgifter i forbindelse med levering af håndværksydelser over større afstande (udstationering).

Store prisforskelle behøver dog ikke være tilstrækkelige til at skabe handel mellem regionerne, hvis handelsomkostningerne er tilsvarende høje. For at få en yderligere ide om, hvor handlede/mobile håndværksydelser egentlig er, dvs. hvor integrerede de lokale markeder er, kan vi se på graden af korrelation mellem lønændringer i de forskellige områder. Hvis en lønændring i én region spreder sig til andre regioner, er det tegn på, at disse markeder er tæt integrerede, hvilket indikerer, at håndværksydelser er relativt mobile. Omvendt vil lave korrelationer tyde på en større grad af segmentering.

Table 2. Sammenhænge i lønændringer på tværs af lokalområder.

Konstant	0,6100 (3,64) ***	0,5871 (3,34) ***
Murer	0,2076 (0,95)	0,2076 (0,93)
Snedker	2,3486 (10,76) ***	2,3486 (10,54) ***
Maler	3,0581 (14,02) ***	3,0581 (13,72) ***
Elektriker	1,3109 (6,01) ***	1,3109 (5,88) ***
Øvrig B&A	0,5558 (2,55) **	0,5558 (2,49) **
Smed	1,1032 (5,06) ***	1,1032 (4,95) ***
PERIODE2	0,0868 (0,74)	0,0868 (0,73)
NABO	1,4207 (9,71) ***	
STOREBÆLT		0,7476 (6,17) ***
Antal obs.	1274	1274
R ²	0,2472	0,2147

Note: Den afhængige variabel er korrelationen mellem de årlige lønændringer i en delperiode for en given håndværker-type i to forskellige lokalområder. Se tekst og appendiks for yderligere forklaring. *t*-statistikker i parenteser. *, ** og *** betyder, at estimatet er signifikant på hhv. 10%, 5% og 1% niveau.

Til brug for denne analyse beregner vi korrelationer mellem de årlige lønændringer i perioderne 1993-1998 og 1999-2004 for de forskellige håndværkertyper i de forskellige områder. I appendiks A er det mere præcist angivet, hvordan disse korrelationer er beregnet.⁵ Opdelingen på to delperioder vælges for at kunne teste for, om der er sket ændringer i »markedsintegrationen« over tid.⁶ Et første tegn på, at de regionale markeder ikke er fuldt integrerede, vil være, at lønkorrelationerne er højere mellem markeder, der ligger geografisk tættere på hinanden, end mellem markeder, der ligger langt fra hinanden. For at afgøre, hvorvidt dette er tilfældet, kan vi estimere følgende model:

$$r_{ijht} = \alpha_0 + \alpha_1 TYPE_i + \alpha_2 NABO_{jh} + \alpha_3 PERIODE2_t + \varepsilon_{ijht} \quad (1)$$

5. Beregningerne er inspireret af en metode udviklet af Frankel og Rose (1998), hvor det testes, hvorvidt international handel giver tættere korrelerede konjunkturcykler. I Andersen, Haldrup og Sørensen (2000) benyttes samme metode til at analysere, hvorvidt international handel giver tættere korrelerede lønninger.

6. Desuden er der et trade-off mellem antallet af delperioder og antallet af observationer i den afsluttende regression forstået på den måde, at hvis man vælger ikke at opdele på to delperioder og i stedet beregner korrelationerne for den fulde periode, så halveres antallet af observationer til den afsluttende regression.

hvor r_{ijht} er en logistisk transformation af korrelationskoefficienten mellem lønændringerne i periode t i region i og j for håndværkstype h (se appendix A for nærmere forklaring). $TYPE_i$ er en vektor af dummyvariable for de forskellige håndværker typer (*tømrer* er den udeladte kategori), $PERIODE2_t$ er en dummyvariabel for periode 2 (1999-2004), og $NABO_{jh}$ er en dummyvariabel, der angiver hvorvidt der er tale om regioner, der grænser op til hinanden. Regioner, der er adskilt af Storebælt, antages ikke at være naboregioner, hvorimod de to regioner, der grænser op til Lillebælt antages at være naboregioner. Hvis markederne er delvist segmenterede, vil vi forvente, at koefficienten til $NABO_{jh}$ er signifikant positiv, således at lønændringer er tættere korrelerede i naboregioner.

Resultaterne fra regressionen er angivet i første søjle i tabel 2. Vi ser for det første, at estimatet på α_2 er stærkt signifikant positivt, hvilket viser, at korrelationen mellem lønændringer i naboregioner er stærkere end mellem andre regioner. For det andet, eftersom estimatet på α_3 ikke er signifikant, betyder det, at der ikke er tegn på, at korrelationerne har ændret sig i de senere år i forhold til tidligere. Endelig viser regressionen også, at lønændringer for f.eks. malere er tættere korrelerede på tværs af regionale markeder end lønændringer for f.eks. murere og tømrere.

En alternativ tilgangsvinkel kunne være at betragte Storebælt som en særlig barriere, der gør, at markederne på samme side af Storebælt er tættere integrerede med hinanden end markeder på hver sin side af Storebælt. For at undersøge dette kan vi i ovenstående regression erstatte $NABO_{jh}$ med $STOREBÆLT_{jh}$, som er en dummy for om to regioner ligger på samme side af Storebælt.⁷ Resultaterne fra denne regression fremgår af den anden søjle i tabel 2.

Vi ser, at estimatet på α_2 igen er signifikant positivt, men betragteligt mindre end i den foregående analyse. Det støtter, at lønændringer i regioner på samme side af Storebælt er tættere korrelerede, men samtidig mindre end lønændringer i naboregioner. Resultaterne tyder altså på, at markederne ikke blot er segmenterede som følge af Storebælt, men til en vis grad også inden for f.eks. Jylland.

Samlet set peger analyserne i dette afsnit på, at der er et vist grundlag for interregional handel med håndværksydelser i Danmark. Den geografiske spredning af produktionen tyder dog ikke umiddelbart på, at handel er motiveret af regionale stordriftsfordele. Omvendt tyder de betydelige prisforskelle mellem regionale markeder på, at der kan være en vis handel med håndværksydelser fra lavtløns- til højt lønsområder; noget der kan forstærkes af de skattemæssige subsidier til denne type af handel. Den lavere korrelation mellem lønudviklinger på markeder, der ligger længere væk for hinanden rent geografisk, tyder dog samtidig på, at handlen ikke er så udbredt, at de regionale markeder er blevet fuldt integrerede.

7. Da Bornholm ligger ret isoleret, antages det, at Bornholm hverken tilhører området vest eller øst for Storebælt.

4. Handlen med håndværksydelser

I dette afsnit vil vi se på den handel med håndværksydelser, der foregår ud af de forskellige regionale markeder. Tidligere har det været umuligt at opgøre denne handel, da den ikke registreres direkte. Skattelovgivningen giver imidlertid mulighed for indirekte at måle en del af denne handel og dermed estimere denne handels andel i den samlede produktion.

4.1 Opgørelse af handlen

Nøglen til opgørelsen af den interregionale handel er, at handel med håndværksydelser forudsætter, at arbejdskraften (håndværkeren) midlertidigt opholder sig på leveringsadressen. Og hvis en person i forbindelse med dette har ekstraomkostninger til kost og logi (det vil i praksis sige, at han/hun skal overnatte uden for hjemmet), kan vedkommende modtage skattemæssig kompensation. Dette gælder både hvis »udstationeringen« foregår via virksomheden, som personen normalt er ansat i, eller den foregår ved at personen selv tager midlertidigt arbejde i en lokal virksomhed på destinationen. Begge typer af »udstationering« betragtes som handel med håndværksydelser i det følgende, men langt den overvejende del af udstationeringen foregår via virksomheden.⁸

Kompensationen kan opnås på tre måder:

1. Ved udbetaling af *skattefri rejsegodtgørelse* fra arbejdsgiveren efter faste satser vedtaget af Ligningsrådet.
2. Ved *refusion efter udlæg* ifølge regning fra arbejdsgiveren.
3. Ved *fradrag på selvangivelsen* efter faste satser vedtaget af Ligningsrådet.

De faste satser vedtaget af Ligningsrådet, der både gælder ved udbetaling af skattefri rejsegodtgørelse fra arbejdsgiveren og ved fradrag på selvangivelsen, er gengivet i tabel 3.

Når der benyttes faste satser som under punkt 1 og 3, registreres udbetalingen af skattemyndighederne, mens dette ikke er tilfældet ved refusion efter udlæg. Da satserne er forholdsvis attraktive at benytte for udstationerede håndværkere, vurderes den manglende registrering af refusion efter udlæg som værende af forholdsvis begrænset omfang.⁹ Der er naturligvis betingelser, der skal være opfyldt for at opnå enten skatte-

8. Dette er helt på linje med WTO's definition af mode-4 handel med serviceydelser, jf. fodnote 1.

9. Selvom det fra arbejdstagerens synspunkt umiddelbart kunne synes fordelagtigt, at arbejdsgiveren betalte for logi og kost efter udlæg, da udgifter til hotel eller lignende hurtigt vil overstige satsen for den skattefri rejsegodtgørelse, så ville dette betyde, at den leverede ydelse ville blive meget dyr for arbejdsgiveren og dermed kunden.

Tabel 3. Oversigt over nominelle satser og totale beløb.

	Sats pr. døgn (kr.)	Udbetalt skattefri godtgørelse (mio. kr.)	Fradrag ifølge selvangivelsen (mio. kr.)	I alt (mio. kr.)
1993	368,00	239,16	81,12	320,29
1994	377,30	234,13	149,83	383,96
1995	381,80	240,21	102,75	342,96
1996	391,20	212,26	73,36	285,61
1997	398,30	213,20	59,02	272,22
1998	409,45	221,15	37,54	258,69
1999	420,50	202,95	25,72	228,67
2000	501,00	235,66	34,13	269,78
2001	519,00	255,53	40,96	296,49
2002	534,00	258,53	35,95	294,48
2003	553,00	269,03	34,89	303,92
2004	570,00	272,58	35,93	308,51

Note: Satsen pr. døgn er opgjort som den samlede mulige skattefri godtgørelse, bestående af en sats for logi og en sats for kost. Alle observationer indeholdende udbetalt skattefri godtgørelse er medtaget (inklusive negative værdier) for individer med håndværksmæssig uddannelse ansat i bygge- og anlægsbranchen i november måned det pågældende år. Fradrag iflg. selvangivelsen er tillagt bundgrænsen, såfremt der er registreret et positivt fradrag i det pågældende år.

fri rejsegodtgørelse eller fradrag. Disse betingelser samt de væsentligste ændringer, der har været i reglerne i den periode, vi betragter, er kort opsummeret i appendiks B.

Som det fremgår af tabel 3, så har de nominelle satser været støt stigende. Den skattemæssige værdi af fradraget eller godtgørelsen er dog kun mellem 1/3 og 2/3 heraf.¹⁰

Oplysningerne om *skattefri rejsegodtgørelse* og *fradrag på selvangivelsen* kan benyttes til at estimere omfanget af udstationering. Da reglerne er således, at man enten kan få skattefri rejsegodtgørelse fra arbejdsgiveren eller fradrag på selvangivelsen, kan de to beløb for den enkelte person adderes.¹¹ Ved at dividere det resulterende beløb med de faste satser pr. rejsedag fås et estimat på omfanget af rejsedage (eller »ude-dage«) for hver enkelt person og dermed et estimat på omfanget af den interregionale handel med håndværksydelser, idet den største del af denne handel må antages at forudsætte udstationering med overnatning.

I forbindelse med opgørelsen af fradrag på selvangivelsen er der det problem, at indberetningen sker under rubrikken »Øvrige lønmodtagerudgifter«, der også omfat-

10. Ligningsmæssige fradrag som f.eks. rejseudgifter fradrages i 2004 kun i amts-, kommune og kirkeskatten, hvorved den skattemæssige værdi af fradraget bliver omkring 30-35% afhængig af de lokale skattesatser. Den skattemæssige værdi af rejsegodtgørelsen er knapt så entydig, men hvis vi forestiller os, at individet modtager skattefri godtgørelse i stedet for almindelig løn, således at arbejdsgiverens fortjeneste er upåvirket, så sparer lønmodtageren marginalskatten, som for de højt lønnede er ca. 63%.

11. Det er også tilfældet, at hvis arbejdsgiveren udbetaler mindre i skattefri rejsegodtgørelse end bestemt ved de faste satser, så kan den ansatte fradrage differencen på selvangivelsen.

ter andre udgifter end rejseudgifter, der har været nødvendige for arbejdets udførelse. Det drejer sig især om kursusudgifter og telefonudgifter. I specielle tilfælde kan der desuden være fradrag til beklædning, arbejdsværelse, computer, faglitteratur m.m. Det forekommer dog rimeligt at antage, at for håndværkere består »Øvrige lønmodtagerudgifter« i langt overvejende grad af rejseudgifter. Men i det omfang andre udgifter optræder under denne rubrik, vil det naturligvis give en opadrettet bias på vores estimat på antallet af udedage.

En række andre forhold bevirker imidlertid, at vores estimat på omfanget af *reelle* udedage samlet set vil være konservativt. For det første får man kun fradrag, hvis beløbet overstiger en given bundgrænse. I appendix B findes en tabel med størrelserne af denne bundgrænse i de forskellige år. I de følgende beregninger vil vi antage, at for de personer, der ikke har fået registreret noget fradrag, har der ikke været nogen rejseaktivitet udover, hvad der kan udledes af en eventuel udbetalt skattefri rejsegodtgørelse. For de, der har fået registreret et fradrag, anvendes det registrerede fradrag plus bundgrænsen.

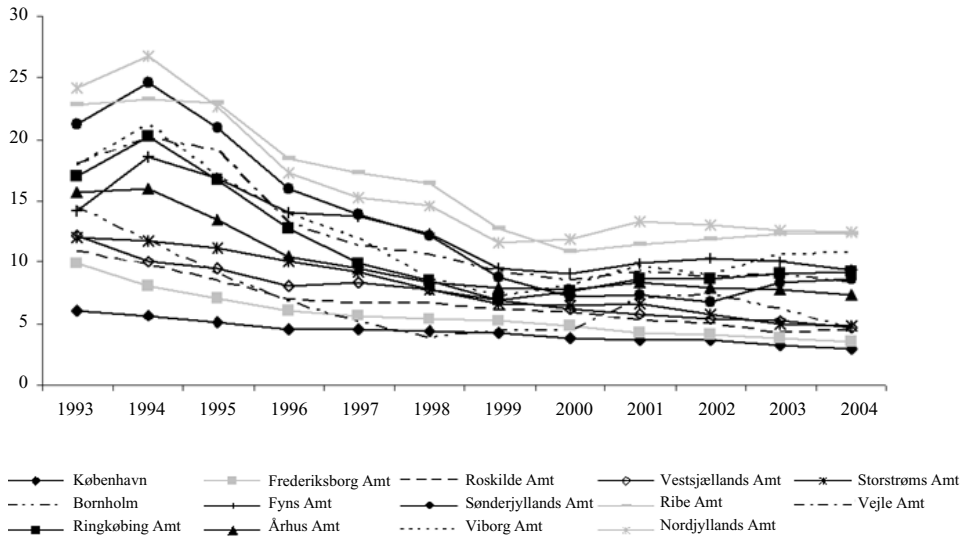
For det andet har vi som nævnt ikke data på *refusion efter udlæg*, hvorved rejseaktivitet under denne regel ikke medregnes. For det tredje gælder reglerne ovenfor kun for lønmodtagere. Dvs. selvstændiges rejseaktivitet medregnes ikke.

For det fjerde dividerer vi de udbetalte og fratrukne beløb med de fulde satser, selv om nogle personer i perioder kun har haft ret til reducerede satser (se appendix B). Det betyder også, at vi undervurderer omfanget af udedage for personer, der f.eks. er væk fra mandag morgen til torsdag aften og dermed kun får godtgørelse for 3,5 døgn, selv om de har været udstationeret i fire fulde arbejdsdage. Dette problem er naturligvis forholdsvis større, hvis man kun er borte i to arbejdsdage og får godtgørelse for 1,5 døgn.

For det femte fanger tallene ikke de folk, der pendler dagligt til et fjerntliggende arbejdssted i forbindelse med levering af en håndværksydelse.

4.2 Omfanget af handlen

Tabel 3 viser også det samlede omfang af udbetalt skattefri rejsegodtgørelse og fradrag på selvangivelsen for håndværkere i bygge- og anlægsbranchen i perioden 1993-2004. Som vi kan se af tallene, er udbetaling af skattefri rejsegodtgørelse den hyppigst anvendte metode, hvilket kan hænge sammen med, at skatteværdien af rejsegodtgørelsen er potentielt højere (se fodnote 10). De problemer, der således knytter sig til opgørelsen af udedage som følge af fradrag på selvangivelsen, vurderes derfor som relativt begrænsede, da eventuel støj herfra kun i begrænset omfang vil påvirke opgørelsen af det totale omfang af reelle udedage. Samlet set vil det resulterende estimat på omfanget af reelle udedage dog være relativt konservativt.



Figur 3. Gennemsnitligt antal udstationerede dage fordelt på lokalområder.

Note: Antal udedage er for hvert enkelt individ begrænset til intervallet [0;365]. Personer med »<0« eller »>365« udedage er således sat til henholdsvis 0 og 365 udedage, inden det gennemsnitlige antal udedage er beregnet for håndværkere bosat i et givet lokalområde.

I figur 3 er vist udviklingen i det gennemsnitlige antal udstationerede dage for håndværkere i de forskellige regioner.

Vi ser, at der er en klar tendens til, at håndværkere bosat i Jylland er mere udstationerede end håndværkere på Sjælland. Der er desuden et markant højere niveau af udedage i de første 3-4 år af den betragtede periode, hvilket kan hænge sammen med en generelt højere arbejdsløshed i de år og dermed et større incitament til at søge arbejdsopgaver i udlandet. Det relativt konstante niveau for København tyder dog på, at det ikke er f.eks. andre lønmodtagerfradrag, der initialt »forurener« vores estimater. Det yderligere fald omkring 1998-99 kan meget vel tænkes at skyldes ibrugtagningen af Storebæltsbroen, der kan have resulteret i et ændret mønster med hensyn til, hvornår man er udstationeret, og hvornår man dagligt pendler mellem hjem og arbejdssted. Perioden fra 1999 og frem afspejler i øvrigt en langt mere stabil udvikling end de tidligere år.

En yderligere opdeling af udedage på håndværkertyper og regioner er givet i tabel 4. Tabellen bekræfter, at udstationering benyttes mere blandt jyske håndværkere end blandt sjællandske. Desuden er malerne den mindst udstationerede gruppe. Her er det

Table 4. Gennemsnitligt antal udstationerede dage fordelt på håndværkertyper og lokalområder, 1993 og 2004.

	Murere		Tømrere		Snedkere		Malere		Elektrikere		Smede		Øvrig B&A	
	1993	2004	1993	2004	1993	2004	1993	2004	1993	2004	1993	2004	1993	2004
København	7,66	2,79	10,60	2,81	5,62	1,99	3,90	3,13	7,08	4,28	7,26	4,93	2,87	1,76
Frederiksborg Amt	11,41	4,24	12,76	3,43	8,03	2,41	7,87	3,44	9,61	5,08	11,18	4,22	5,97	2,45
Roskilde Amt	17,09	4,75	13,73	3,44	10,56	2,16	6,01	3,93	10,45	5,64	7,55	8,07	7,07	3,43
Vestsjællands Amt	12,19	2,87	17,19	4,16	15,70	2,41	5,69	2,34	8,49	4,50	11,52	12,15	7,07	5,25
Storstrøms Amt	14,03	3,95	13,86	4,89	9,70	3,47	7,50	3,86	8,34	4,87	17,90	7,60	10,39	4,98
Bornholm	14,37	3,54	24,16	5,40	19,16	1,70	2,57	4,71	4,15	1,74	9,61	11,44	2,18	1,60
Fyns Amt	19,15	9,85	17,10	10,05	14,45	8,43	4,19	4,54	13,18	9,69	13,04	10,18	11,33	11,49
Sønderjyllands Amt	24,65	7,72	32,12	11,55	16,70	6,00	4,64	5,09	12,57	6,64	18,39	17,01	33,93	8,06
Ribe Amt	27,93	11,58	35,71	17,72	23,33	13,11	5,70	6,40	7,55	5,81	17,58	16,19	16,28	7,50
Vejle Amt	21,87	7,05	20,59	8,73	11,59	8,48	7,15	4,75	17,12	7,94	22,07	15,74	18,61	8,66
Ringkøbing Amt	21,49	7,86	21,52	11,50	21,81	8,82	6,95	3,25	4,15	5,88	16,68	16,00	19,84	9,14
Århus Amt	22,38	6,79	19,05	6,68	13,53	7,61	9,29	3,68	10,62	8,39	13,99	11,27	12,67	9,45
Viborg Amt	24,63	8,83	19,46	10,41	22,47	15,80	4,93	4,76	9,33	6,25	21,00	19,59	13,80	17,07
Nordjyllands Amt	34,87	9,72	31,04	12,98	27,00	12,99	5,80	9,89	13,60	10,40	20,30	21,74	19,37	12,14

Note: Antal udedage er for hvert enkelt individ begrænset til intervallet [0;365]. Personer med " ≤ 0 " eller " > 365 " udedage er således sat til henholdsvis 0 og 365 udedage, inden det gennemsnitlige antal udedage er beregnet for de forskellige typer af håndværkere bosat i et givet lokalområde.

iøvrigt interessant at bemærke, at malerne også var den gruppe, der havde den mest jævne geografiske fordeling af produktionen (jf. tabel 1).

At antallet af udedage er størst i de jyske og fynske »lavtlønsområder« tyder på, at handlen med håndværksydelser i høj grad er motiveret af prisforskelle. Det faktum, at ikke kun håndværkere fra lavtlønsområder, men også håndværkere fra f.eks. København er udstationerede, tyder dog på, at produktdifferentiering eller de danske subsidier i form af skattefradrag/-frihed også spiller en rolle. I de første kolonner af tabel 5 har vi opgjort, hvor meget eksporten af håndværksydelser fylder i produktionen i henholdsvis Øst- og Vestdanmark. I Vestdanmark ligger estimatet for 2004 på 2,3-6,8%, mens eksporten kun udgør 1,0-3,3% i Østdanmark. Tidligere i den betragtede periode har tallene for især Vestdanmark været dobbelt så store. Som vi har argumenteret for, er dette iøvrigt relativt konservative estimater.

Et andet interessant spørgsmål er naturligvis, hvor meget udstationering bidrager til udbuddet af håndværksydelser i de forskellige områder? Kan vi f.eks. sige noget om, hvor meget de jyske håndværkere »fylder« på Sjælland og evt. i København?

Lad os i første omgang betragte Jylland og Fyn som ét område og Sjælland og Lolland-Falster som et andet område. Lad os yderligere antage, at al udstationering fra Jylland og Fyn er til Sjælland og Lolland-Falster og omvendt. Den totale mængde håndværksydelser, der leveres på Sjælland og Lolland-Falster, kan nu måles som den totale beskæftigelse på Sjælland og Lolland-Falster plus udstationeringen af håndværkere fra Jylland og Fyn minus udstationeringen af håndværkere fra Sjælland og Lolland-Falster. Hvis udstationeringen af jyske og fynske håndværkere derefter sættes i forhold til den totale mængde håndværksydelser, får vi et udtryk for, hvor meget jyske og fynske håndværkere bidrager til udbuddet af håndværksydelser på Sjælland og Lolland-Falster. På tilsvarende måde kan vi beregne, hvor meget håndværkere fra Sjælland og Lolland-Falster bidrager til udbuddet af håndværksydelser i Jylland og på Fyn. Søjle 3 og 4 i tabel 5 viser disse beregninger for 2004. Vi ser for det første, at håndværkere fra Jylland og Fyn fylder langt mere på Sjælland og Lolland-Falster end omvendt. Således udgør det jysk-fynske udbud af tømrere over 7% på Sjælland og Lolland-Falster, mens tømrere fra Sjælland og Lolland-Falster kun fylder ca. 1% i Jylland og på Fyn.

I søjle 5 og 6 er det vist, hvordan nettoudstationeringen – dvs forskellen mellem indstationeringen og udstationeringen – bidrager til udbuddet af håndværksydelser. Vi ser, at udstationering har en forholdsvis stor positiv nettoeffekt på Sjælland og Lolland-Falster, mens der er en negativ nettoeffekt i Jylland og på Fyn.

Da det især er i København, at prisniveauet på håndværksydelser er højt, forekommer det sandsynligt, at jysk-fynske håndværkere primært tager til hovedstaden. Derfor har vi i den næstsidste søjle antaget, at jyske og fynske håndværkere kun arbejder i

Table 5. Udstationeringens bidrag til udbudet af håndværksydelser i 2004.

	Eksport fra Vest- danmark	Eksport fra Øst- danmark	Jyder på Sjælland	Sjællændere i Jylland	Nettoudstationering		Jyder i København	
					på Sjælland	i Jylland	Brutto	Netto
Murere	0,0379	0,0156	0,0665	0,0087	0,0517	-0,0304	0,1827	0,1725
Tømrere	0,0473	0,0160	0,0712	0,0104	0,0561	-0,0387	0,2061	0,1962
Snedkere	0,0438	0,0103	0,0463	0,0097	0,0363	-0,0357	0,1051	0,0972
Malere	0,0234	0,0143	0,0242	0,0138	0,0101	-0,0098	0,0490	0,0357
Elektrikere	0,0362	0,0207	0,0403	0,0186	0,0200	-0,0183	0,0872	0,0697
Smede	0,0677	0,0326	0,0992	0,0218	0,0688	-0,0493	0,2631	0,2468
Øvrig B&A	0,0454	0,0129	0,0344	0,0172	0,0218	-0,0296	0,0721	0,0648

Note: Udbuddet af håndværksydelser i et lokalområde er opgjort som antallet af fuldtids- og deltidsbeskæftigede pr. 1. november. Et årsværk af udedage er opgjort som 227 udedage. Antal udedage er for hvert enkelt individ begrænset til intervallet [0;365]. Personer med »<0« eller »>365« udedage er således sat til henholdsvis 0 og 365 udedage, inden antal udedage er beregnet for håndværkere bosat i et givet lokalområde.

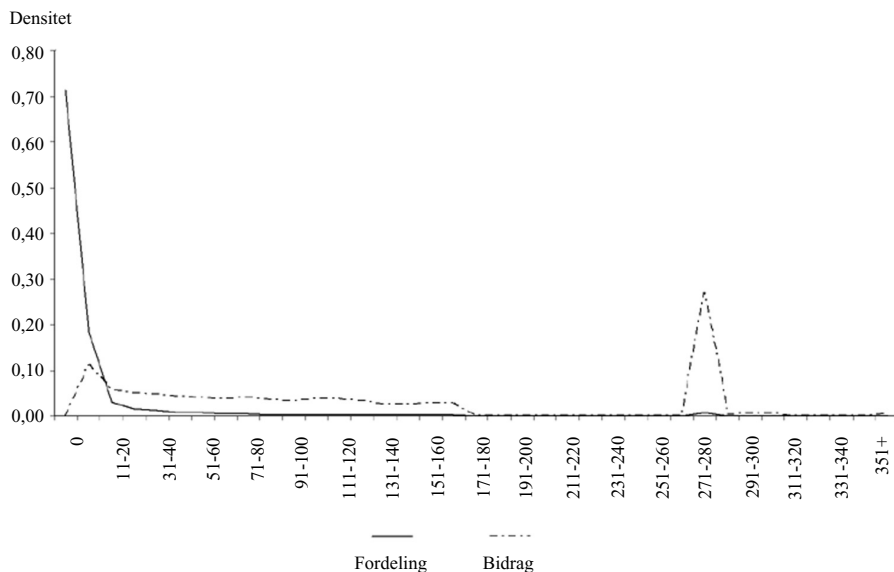
København, og det er da beregnet, hvor meget de der bidrager til udbuddet af håndværksydelser. Under denne antagelse ser vi, at f.eks. jyske og fynske tømrere udgør godt 20% af udbuddet i hovedstaden. I den sidste søjle ser vi, at dette giver en nettoeffekt af udstationering af tømrere i hovedstaden på godt 19%.¹²

Af tabellen kan vi også se, at det især er murere, tømrere og smede fra Jylland og Fyn, der »fylder« på Sjælland og Lolland-Falster. Til gengæld er der kun meget få jysk-fynske malere.

Som nævnt er der en række grunde til at tro, at det estimerede omfang af reelle udedage er relativt konservativt. Det betyder, at de beregnede bidrag fra udstationering til udbuddet af håndværksydelser vil være underestimeret. Omvendt betyder antagelsen om, at alle de jysk-fynske udedage anvendes på Sjælland eller i København, at estimatet på bidraget bliver for højt. F.eks. er det en mulighed, at de i et vist omfang anvendes i udlandet eller inden for Jylland og Fyn. Mange håndværkere er traditionelt taget til f.eks. Norge for at arbejde i perioder. Til gengæld har vi heller ikke indstationeringen af udlændinge med i vores beregninger af, hvor meget handlen fylder i det regionale udbud.¹³ Under alle omstændigheder tyder analysen på en betragtelig handel med disse ydelser på tværs af de danske regioner – en handel der synes at være drevet af de ek-

12. Hvis man yderligere antager, at tømrere fra Vestsjælland og Lolland-Falster udstationeres i København i stedet for i Jylland og på Fyn, så kommer nettoeffekten af udstationering op på over 21% i København. Det er dog tvivlsomt om håndværkere fra det øvrige Sjælland kan betragtes som udstationerede, når de arbejder i København.

13. Denne indstationering kan dog være anderledes fordelt på tværs af regioner end den tilsvarende udstationering til udlandet og kan dermed påvirke det regionale mønster.



Figur 4. Fordeling af antal udstationerede dage og bidrag til samlede antal dage i 2004 – individer.

Note: Antal udedage er for hvert enkelt individ begrænset til intervallet [0;365] inden figuren er opgjort. Ved opgørelsen af bidraget er intervallsandsynlighederne ganget med midtpunktet i intervallet.

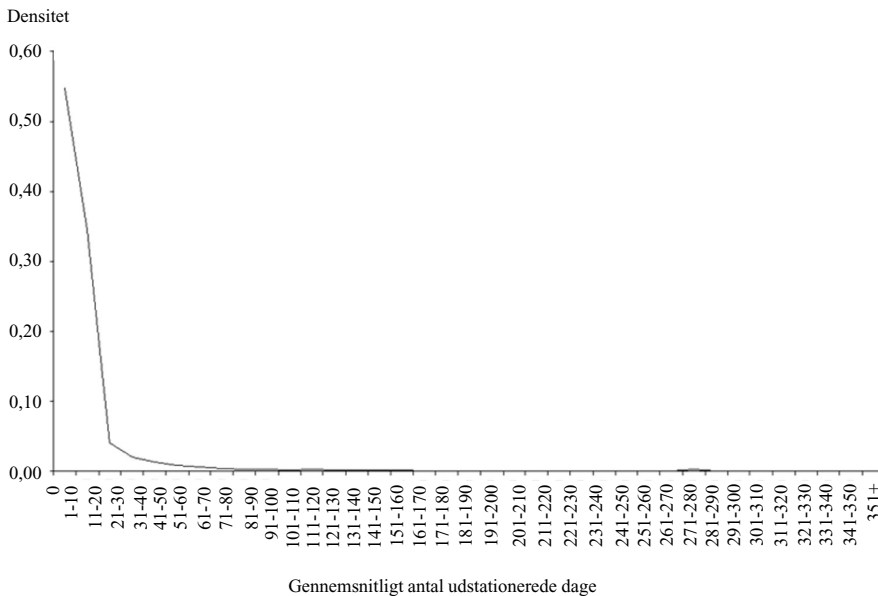
sterende prisforskelle kombineret med produktdifferentiering og/eller skattesubsidier. I de følgende afsnit prøver vi at belyse dette yderligere.

5. Hvem eksporterer håndværksydelser?

Foregående afsnit dokumenterede en betragtelig (interregional) handel med håndværksydelser i Danmark. I indeværende afsnit kigger vi nærmere på, hvem der leverer disse ydelser. Er det jævnt fordelt, eller er der virksomheder og individer, der specialiserer sig i denne »handel« med håndværksydelser?

Figur 4 og 5 viser, hvordan eksporten, dvs. antallet af udedage, fordeler sig på tværs af henholdsvis individer og arbejdssteder i 2004. Figur 4 indeholder endvidere fordelingen af bidraget til det samlede antal udedage. Som vi kan se, så har langt størstedelen af håndværkerne (omkring 70%) og arbejdsstederne (omkring 50%) ingen registrerede udedage. Blandt de personer og arbejdssteder, der har registrerede udedage, har de fleste relativt få (<20), men der er også en betydelig andel med over 100 udedage. Faktisk er der en mindre »pukkel« af individer med 271-280 estimerede udedage.

Det billede, der umiddelbart tegner sig, er altså en vis specialisering i handel både på individ- og arbejdsstedsniveau. Dette kunne indikere, at handlen i et vist omfang er



Figur 5. Fordeling af gennemsnitligt antal udstationerede dage i 2004 – arbejdssteder.

Note: Det gennemsnitlige antal udedage for et arbejdssted er beregnet som det gennemsnitligt antal udstationerede dage for håndværkere tilknyttet arbejdsstedet pr. 1. november.

drevet af produktdifferentiering i den forstand, at visse virksomheder producerer særlige ydelser, der efterspørges over hele landet. I det følgende prøver vi at afgøre, hvad der karakteriserer de individer/firmaer, der vælger at eksportere håndværksydelser ud af lokalområdet. Dette gøres ved at estimere Tobit modeller på både arbejdssteds- og individniveau for perioden 1999-2004, hvor den afhængige variabel er antallet af udedage (pr. ansat når estimationen er på arbejdsstedsniveau).¹⁴

De forklarende variable, der anvendes i regressionerne, er en række arbejdssteds-karakteristika, der siger noget om arbejdsstedets størrelse og den typiske ansatte håndværker. Vi anvender således antal ansatte¹⁵ på arbejdsstedet samt den gennemsnitlige alder, erfaring og tenure og andelen af mænd blandt de ansatte håndværkere. I regressionerne på individniveau inkluderes endvidere en række individkarakteristika, der siger noget om individets muligheder for og omkostninger ved udstationering, herunder

14. En Tobit model anvendes, fordi der er en stor gruppe med værdien nul for den afhængige variabel. Vi anvender kun årene, 1999-2004, i disse og efterfølgende regressioner, fordi mønsteret ser markant anderledes ud inden 1999, jf. figur 3.

15. I regressionerne i dette og det kommende afsnit følger vi en stor del af litteraturen ved at anvende logaritmen til antal ansatte i stedet for antal ansatte.

Tabel 6. Tobit modeller for antal udedage, 1999-2004.

Afhængig variabel: Antal udedage	Arbejdssteds- niveau	Individ- niveau
<i>Arbejdssteds karakteristika</i>		
Log(ansatte)	9,595 (83,460) ***	8,473 (107,010) ***
Gennemsnitlig alder	-0,377 (-18,45) ***	0,159 (5,080) ***
Andel mænd	0,109 (0,090)	15,793 (12,110) ***
Gennemsnitlig erfaring	0,579 (21,140) ***	0,248 (6,010) ***
Gennemsnitlig tenure	-0,914 (-27,80) ***	-1,736 (-41,16) ***
<i>Individkarakteristika</i>		
Gift		-1,935 (-7,91) ***
Børn 0-6 år		-5,627 (-19,70) ***
Børn 7-14 år		-1,113 (-3,94) ***
Alder 25-29		-0,266 (-0,66)
Alder 30-39		0,361 (0,750)
Alder 40-49		1,299 (2,310) **
Alder 50-59		-0,196 (-0,32)
Alder 60+		-0,193 (-0,22)
Fuldtidsansat		36,139 (67,650) ***
Erfaring		-0,130 (-1,99) **
Erfaring kvadreret		0,000 (-0,02)
Tenure		-1,831 (-29,30) ***

fortsættes næste side ...

fortsat ...

Tenure kvadreret		0,068 (18,150)***
Ansæt siden 1980		-1,749 (-1,81)*
<i>Dummyvariable</i>		
Region	JA	JA
Håndværkertype	JA	JA
År	JA	JA
Antal observationer	82412	429868
LR χ^2	11624,53	40630,57
Pseudo R^2	0,028	0,219

Note: I regressionerne på arbejdsstedsniveau er den afhængige variabel det gennemsnitlige antal udedage pr. ansæt pr. 1. november. *t*-statistikker i parenteser. *,** og *** betyder, at estimatet er signifikant på hhv. 10%, 5% og 1% niveau.

dummyvariable for alder og børn i forskellige aldersgrupper samt erfaring og tenure. Endelig inkluderes dummyvariable for håndværkertype, år og region. Yderligere forklaring af variable findes i appendiks C.

Vi ser, at især større arbejdssteder har mange udedage pr. person. Endvidere er det især arbejdssteder, der er karakteriseret ved håndværkere med lav tenure. Dvs. estimaterne tyder på, at udstationerede arbejdssteder har større udskiftning blandt deres ansatte håndværkere. Individregressionerne peger desuden på, at det især er ugifte, barnløse håndværkere i alderen 40-50 år og med lav tenure, der har mange udedage. Samlet set bekræfter regressionerne derfor billedet af en vis specialisering i eksport – både blandt store arbejdssteder med kortere ansættelsesforhold men også blandt relativt nyansatte individer med lave transportomkostninger, dvs. ugifte og barnløse.

Groft sagt kan vi tale om tre grupper af individer/arbejdssteder. »De ikke-eksportende«, dvs. dem uden udedage, »deltidseksportørerne«, dvs. dem med begrænsede udedage (<100), og »fuldtidseksportørerne«, dvs. dem med mange udedage (>100). Et alternativ til Tobit estimationerne kunne derfor være at estimere multinomial logit modeller, hvor den afhængige variabel kan antage tre forskellige værdier svarende til de tre grupper af individer og arbejdssteder. Resultaterne fra disse regressioner kan findes i appendiks C. De viser stort set det samme billede som Tobit regressionerne ovenfor.

6. Løn og udstationering

I afsnit 5 så vi på hvilke firmaer og personer, der leverer eksporten af håndværksydelser. I indeværende afsnit ser vi på, hvordan udstationering påvirker individernes

aflønning. Skal de udstationerede personer kompenseres i form af højere løn end deres ikke-udstationerede kolleger for de ekstra omkostninger, der er forbundet med udstationeringen? Det drejer sig om omkostninger til kost og logi, men også omkostninger af ikke-pekuniær art såsom ubehag ved at opholde sig uden for eget hjem og adskilt fra familien kan spille en rolle. Eller er den skattefrie godtgørelse og/eller fradraget på selvangivelsen en tilstrækkelig kompensation?

Hvis de udstationerede håndværkere er villige til at påtage sig arbejde til en løn, der er lavere end lønnen i forbindelse med arbejde, der ikke kræver udstationering, så er det tegn på, at den skattefrie godtgørelse og fradraget overkompenserer dem for deres omkostninger. Omvendt, hvis omkostningerne ved udstationering er højere, vil håndværkerne forlange en højere løn, hvis der er udstationering i forbindelse med et arbejde.¹⁶

Der kan dog også være andre årsager til, at lønnen for udstationerede kan afvige fra lønnen for ikke-udstationerede. Det kan være, at udstationerede personer adskiller sig ved at være mere eller mindre produktive end andre, eller det kan være, at udstationerede virksomheder betaler særligt høje eller lave lønninger. Munch og Skaksen (2006) finder f.eks., at eksporterende virksomheder betaler højere lønninger, mens Malchow-Møller, Markusen og Schjerning (2007) viser, at udenlandsk ejede virksomheder også betaler højere lønninger. Resultaterne i tabel 6 tyder på, at både de virksomheder, der udstationerer, og de personer, der bliver udstationeret, skiller sig ud.

I tabel 7 nedenfor præsenteres resultater fra en standard Mincer lønregression, se f.eks. Heckman, Lochner og Todd (2003), for perioden 1999-2004. På venstresiden i regressionen har vi timelønnen, der i de tre første søjler er beregnet på basis af et antal timer svarende til en normal arbejdsuge, dvs. der er ikke korrigeret for evt. overarbejde. Dette gøres derimod i søjle 4 og 5 (se nedenfor).¹⁷ Som forklarende variable benyttes de sædvanlige individkarakteristika (jf. afsnit 5), størrelsen af arbejdsstedet (logaritmen til antallet af ansatte) samt dummyvariable for regioner, håndværkertyper og år. Desuden inkluderes en dummy for, hvorvidt personen har haft imellem 1 og 25 udedage i året, og en dummy for, hvorvidt personen har haft over 25 udedage.

I den første søjle i tabel 7 vises resultatet fra en almindelig OLS regression. Vi ser her, at der er en lønpræmie ved udedage, og at præmien er klart størst ved mange udedage. En person med mere end 25 udedage får således en timeløn, der er 10,5% højere

16. Siden 1994 har direkte reduktion i lønnen i forbindelse med udbetaling af skattefrie rejsegodtgørelse ikke været tilladt. Det forhindrer dog ikke, at en håndværker kan acceptere en generelt lavere løn, hvis det forventes, at der er en del udstationering forbundet med jobbet.

17. Timelønsvariablen, der benyttes i dette afsnit, er baseret på indberetninger fra arbejdsgiverne («Lønstatistikens serviceregister»). Det har den fordel, at der kan tages højde for overarbejde, idet information herom også oplyses af arbejdsgiveren. Ulempen er, at arbejdssteder med færre end 10 ansatte ikke indgår i datamaterialet. Hvis IDA timelønnen anvendes i stedet for i søjle 1-3 således, at også arbejdssteder med mindre end 10 ansatte kan medtages, fås dog tilsvarende resultater.

Tabel 7. Lønninger og udedage, 1999-2004.

Afhængig variabel: log(timeløn).	1	2	3	4	5
	OLS	Arbejdssteds fixed effects	Job-spell fixed effects	Job-spell fixed effects	Individ fixed effects
<i>Arbejdssteds karakteristika</i>					
Log(ansatte)	0,041 (108,00) ***	0,015 (8,11) ***	0,011 (5,76) ***	0,008 (4,09) ***	0,012 (20,03) ***
<i>Individkarakteristika</i>					
Gift	0,017 (15,95) ***	0,013 (12,73) ***	0,002 (0,84)	0,002 (0,97)	0,001 (0,43)
Børn 0-6 år	-0,001 (-0,44)	-0,002 (-2,02) **	-0,002 (-1,40)	0,001 (0,69)	0,002 (1,2)
Børn 7-14 år	0,010 (8,40) ***	0,009 (7,61) ***	0,002 (1,28)	0,002 (1,08)	0,003 (2,8) *
Alder 30-39	0,002 (1,43)	0,001 (0,49)	0,002 (0,64)	0,001 (0,49)	0,005 (2,03) **
Alder 40-49	0,003 (1,64) *	0,001 (0,72)	0,002 (0,40)	-0,001 (-0,22)	0,002 (0,7)
Alder 50+	-0,015 (-6,40) ***	-0,011 (-4,98) ***	0,001 (0,17)	-0,001 (-0,11)	-0,000 (-0,02)
Mand	0,073 (21,93) ***	0,054 (16,30) ***			
Erfaring	0,007 (21,14) ***	0,005 (18,19) ***	0,014 (7,13) ***	0,009 (4,86) ***	0,012 (9,38) ***
Erfaring kvadreret (x1000)	-0,118 (-15,97) ***	-0,106 (-15,37) ***	-0,165 (-8,00) ***	-0,171 (-8,45) ***	0,000 (-9,14) ***
Tenure	0,001 (5,82) ***	0,005 (19,36) ***	0,002 (2,93) ***	0,002 (4,10) ***	0,002 (6,92) ***
Tenure kvadreret (x1000)	-0,049 (-3,87) ***	-0,103 (-8,16) ***	-0,117 (-4,71) ***	-0,176 (-7,20) ***	0,000 (-7,6) ***
0-25 udedage	0,028 (26,69) ***	0,015 (14,34) ***	0,006 (4,27) ***	0,002 (1,48)	0,002 (1,39)
25+ udedage	0,105 (64,08) ***	0,070 (41,75) ***	0,019 (7,09) ***	0,008 (2,95) ***	0,011 (4,96) ***
<i>Dummyvariable</i>					
Region	JA	JA	JA	JA	JA
Håndværkertype	JA	JA	JA	JA	JA
År	JA	JA	JA	JA	JA
Antal observationer	178294	178294	178294	178294	178294
R ²	0,2117	0,4326	0,8148	0,8125	0,7418

Note: I regressionerne i søjle 1-3 er den afhængige variabel timelønnen uden korrektion for overarbejde. I søjle 4-5 er det timelønnen med korrektion for overarbejde. Kun fuldtidsansatte på 25+ år med informationer fra »Lønstatistikens serviceregister« indgår. *t*-statistikker i parenteser. *, ** og *** betyder, at estimatet er signifikant på hhv. 10%, 5% og 1% niveau.

end personer, der ikke udstationeres, mens en person med mindre end 25 dage får ca. 2,8% mere.¹⁸

I den anden søjle er vist resultatet af en regression med »fixed effects« på arbejdsstedsniveau, dvs. der er kontrolleret for uobserverbare arbejdssteds karakteristika. Vi ser igen, at personer, der har udedage, får en signifikant højere løn end andre. Det er dog vigtigt at bemærke, at de estimerede effekter af udedage er markant lavere i anden søjle (henholdsvis 7,0% og 1,5%) end i første. Det viser, at en væsentlig årsag til, at udstationerede personer modtager en højere løn end andre, er, at de virksomheder, de er ansat i, betaler generelt højere lønninger til deres ansatte. Når disse virksomheder kan betale højere lønninger, tyder det på, at de enten er mere produktive end andre virksomheder¹⁹, eller som følge af produktdifferentiering er i stand til at tage en højere pris for det, de producerer. De estimerede effekter af udedage er dog fortsat signifikante i anden søjle, hvilket viser, at inden for virksomhederne, er det også de udstationerede personer, der tjener mest.

For at afklare, hvorvidt det er noget specielt ved de udstationerede personer, der gør, at de får højere lønninger, eller de højere lønninger skyldes, at de er udstationerede, viser vi i tredje søjle i tabel 7 resultatet af en regression, hvor der er anvendt såkaldte »job-spell fixed effects«, se f.eks. Munch og Skaksen (2006), dvs. fixed effects for enhver kombination af arbejdssted og individ. Koefficienterne til udedage er fortsat signifikant positive (1,9% og 0,6%), men de er langt mindre end koefficienterne i anden søjle. Personer, der er udstationerede, får altså højere løn, både fordi de er særligt produktive generelt set, men tilsyneladende også fordi de er udstationerede.

Timelønnen i de tre regressioner ovenfor minder om det timelønsbegreb, der anvendes i IDA. Det har den ulempe, at hvis en person har meget overarbejde, vil det give sig udslag i en højere timeløn. For at forsøge at tage højde for dette, rapporteres der i fjerde søjle resultatet af en regression, hvor der som afhængig variabel benyttes en timeløn, hvor der er korrigeret for dette. Ligesom ved estimation af koefficienterne i tredje søjle er der anvendt »job-spell fixed effects«. Vi ser nu, at koefficienterne til udedage bliver meget små (0,8% og 0,2%) og kun koefficienten til over 25 udedage er signifikant. Med andre ord, lønpræmien til udstationerede forsvinder stort set, når der kontrolleres for overarbejde som i søjle 4.

18. Bemærk, at koefficienten på »tenure« er meget lav. Et års ekstra tenure giver kun omkring 0,1% højere løn, hvilket er lavt i forhold til det, der typisk findes i litteraturen. Dette kan evt. forklares med, at man ikke opbygger meget »firmaspecifik human kapital« som håndværker, hvis f.eks. arbejdsgange er ens på tværs af virksomheder. Et argument for dette kunne være, at meget af arbejdet alligevel foregår ude hos kunden.

19. Dette er helt i overensstemmelse med Melitz (2003), der i en teoretisk model viser, at eksporterende virksomheder er en selektion af de mest produktive virksomheder. Malchow-Møller, Markusen og Schjerning (2007) viser endvidere, at de mere produktive virksomheder også kan udbetale højere lønninger selv i et kompetitivt arbejdsmarked.

Vi kan konkludere, at udstationerede håndværkere umiddelbart får højere løn end andre håndværkere. Det skyldes dog i høj grad, at udstationerende virksomheder generelt set betaler deres medarbejdere højere løn, samt at de personer, der bliver udstationeret, er personer, der typisk tjener mere end andre – uafhængigt af om de er udstationerede eller ej. Med andre ord, det er især højproduktive eller produktdifferentierende virksomheder, der udstationerer, og det er især højproduktive personer, der er udstationeret. Der er til gengæld ikke noget, der tyder på, at skattegodtgørelsen og/eller fradraget overkompenserer udstationerede håndværkere for deres omkostninger ved at være udstationeret – dvs. der er ikke noget, der tyder på, at virksomheder benytter skatte-subsidiet som et (netto)substitut for løn. Der er til gengæld heller ikke noget, der tyder på, at virksomheder betaler deres medarbejdere ekstra for at være udstationerede. Håndværkere, der er udstationerede, tjener typisk mere, mens de er udstationerede, men dette synes primært at være et resultat af mere overarbejde og højere kvalifikationer.

Kan vi være sikre på, at resultatet om, at antallet af udedage ikke påvirker lønnen inden for et ansættelsesforhold (jf. søjle 4), ikke blot skyldes, at det for nogle ansættelsesforhold er aftalt, at de giver både højere løn og flere udedage – selvom lønnen ikke løbende tilpasses antallet af udedage, bla. fordi det sidstnævnte ikke er tilladt? Dette kunne være forklaringen på, hvorfor vi ikke finder nogen effekt af udedage, når vi kontrollerer for uobserverbare job-spell karakteristika. Det faktum, at effekten af udedage stort set også forsvinder, hvis vi blot har fixed effects på individniveau (søjle 5), tyder imidlertid på, at selv når individet skifter fra ét ansættelsesforhold til et andet med flere/færre udedage, så påvirkes lønnen kun meget lidt.

7. Konklusion

I dette papir har vi forsøgt at belyse et hidtil relativt uberørt område: Handel med håndværksydelser (og serviceydelser mere generelt) på tværs af lokalområder. En række forhold bevirker, at handel med håndværksydelser mellem forskellige lande og regioner må forventes at være betragtelig mindre end handel med varer. For det første, fordi der er langt større handelsomkostninger forbundet med denne type af handel, og for det andet fordi der internationalt har været restriktioner på denne handel.

Østudvidelsen af EU åbner imidlertid op for helt nye perspektiver i den internationale handel med bl.a. håndværksydelser inden for EU, idet der nu rent prismæssigt er opstået langt større potentiale for en sådan handel. Kravet om, at østeuropæere, der får arbejde i Danmark, skal ansættes på danske overenskomstmæssige vilkår, vil lægge en naturlig dæmper på mulighederne for »traditionel« migration. Omvendt gør det det relativt mere attraktivt for østeuropæiske virksomheder at benytte sig af udstationering, som der også er åbnet op for i forbindelse med østudvidelsen. Disse udstationerede medarbejdere er nemlig ikke omfattet af krav om ansættelse på danske overenskomst-

mæssige vilkår. For at kunne skønne over det potentielle omfang af denne handel er det nødvendigt at vide, hvor mobile (dvs. hvor »handlede«) sådanne serviceydelse er. Dette papir giver nogle første svar på dette spørgsmål.

Vi finder, at der er store prisforskelle på håndværksydelser i forskellige områder af Danmark. Timelønnen for en gennemsnitlig håndværker er således knap 20% højere i København end i Sønderjylland. Markederne for håndværksydelser er i høj grad lokale. Der er dog en vis sammenhæng mellem lønudviklingen på disse lokale markeder – især mellem markeder, der er lokaliseret tæt på hinanden.

Vi finder, at der også foregår en del handel med håndværksydelser, der leveres over så store afstande, at det kræver udstationering. Denne handel foregår især ud af lokalområder i Jylland og på Fyn, hvor lønningerne er relativt lave. Der er dog også en vis handel ud af »højtlønsområder« som f.eks. København. Jyske og fynske håndværkere leverer en betragtelig større andel af håndværksydelserne på Sjælland end sjællandske håndværkere leverer i Jylland og på Fyn. Under visse forudsætninger finder vi, at f.eks. jyske og fynske murere og tømrere har en markedsandel på mellem 6,7% og 20,6% i København. Sjællandske håndværkere har til gengæld under 2% af markedet i Jylland.

De personer, der udstationeres er især personer, der formodentlig har relativt lave omkostninger ved at være væk fra eget hjem. Det er således især ugifte, barnløse håndværkere i alderen 40-50 år, der udstationeres.

Virksomheder med udstationerede håndværkere betaler generelt højere lønninger end andre virksomheder. Dette tyder på, at disse virksomheder enten er mere produktive end andre virksomheder, eller at de er bedre til at produktdifferentiere. De personer, der udstationeres synes også at være mere produktive end andre.

Endelig er der ikke noget, der tyder på, at udstationerende virksomheder benytter skattesubsidiet som et (netto)substitut for aflønning af udstationerede håndværkere.

Litteratur

- Andersen, T. M., N. Haldrup og J. R. Sørensen. 2000. Labour Market Implications of EU Product Market Integration, *Economic Policy*, 30, 105-33.
- Bhagwati, J., A. Panagariya og T. N. Srinivasan. 2004. The Muddles over Outsourcing, *Journal of Economic Perspectives*, 18, 93-114.
- Frankel, J. A. og A. K. Rose. 1998. The Endogeneity of the Optimum Currency Area Criteria, *Economic Journal*, 108, 1009-25.
- Heckman, J. J., L. J. Lochner og P. E. Todd. 2003. Fifty Years of Mincer Earnings Regressions, *IZA DP No. 775*.
- Jensen, B. og L. Kletzer. 2005. Tradable Services: Understanding the Scope and Impact of Services Offshoring, i L. Brainard og S.M. Collins, eds., *Offshoring White-Collar Work – Issues and Implications*, Brookings Trade Forum.
- Krugman, P. 1991. *Geography and Trade*, MIT Press, Cambridge, MA.
- Leamer, E. E. og J. Levinsohn. 1995. International Trade Theory: The Evidence, i G.M. Grossman og K. Rogoff, eds., *Handbook of International Economics*, Vol. 3, Elsevier, 1339-94.
- Lipsey, R.E. 2006. Measuring International

- Trade in Services, *NBER working paper* no. 12271.
- Malchow-Møller, N., J. R. Markusen og B. Schjerning. 2007. Foreign Firms, Domestic Wages, *NBER Working Paper* No. 13001.
- Melitz, M. J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, *Econometrica*, 71, 1695-1725.
- Munch, J. R. og J. R. Skaksen. 2006. Human Capital and Wages in Exporting Firms, *CEBR Discussion Paper* 2006-17.
- van Welsum, D. 2003. International Trade in Services: Issues and Concepts, Birkbeck College, School of Economics, Mathematics and Statistics, *Economics Working Paper* 04/03.

Appendiks

A. Beregning af lønkorrelationer

Først beregnes årlige lønændringer for hver af de syv håndværkertyper i hver af de 14 regioner. En årlig lønændring beregnes som logaritmen til gennemsnitslønnen i år $s + 1$ minus logaritmen til gennemsnitslønnen i år s .

Dernæst beregnes korrelationer mellem lønændringer for den samme type af håndværkere for hvert par af regioner for hver af de to perioder: 1993-1998 og 1999-2004. Korrelationskoefficienten, ρ_{ijht} , er således korrelationen mellem lønændringerne i regionerne i og j for håndværkstype h i periode t , hvor $t = 1, 2$ (svarende til de to perioder). Med 14 regioner og 7 håndværkertyper giver det i alt $\frac{1}{2} \cdot 14 \cdot 13 \cdot 7 \cdot 2 = 1272$ observationer.

I forbindelse med regressionsanalysen i sektion 3 er det nødvendigt at benytte en transformation af korrelationskoefficienten, ρ_{ijht} , da denne kun antager værdier mellem -1 og 1 . Vi anvender derfor følgende skalerede logistiske transformation:

$$r_{ijht} = \log \left(\frac{\frac{\rho_{ijht} + 1}{2}}{1 - \frac{\rho_{ijht} + 1}{2}} \right) \quad (2)$$

som kan antage værdier mellem $-\infty$ og ∞ ; se Andersen, Haldrup og Sørensen (2000) for yderligere detaljer.

B. Skattefri rejsegodtgørelse og fradrag

En væsentlig betingelse for at opnå enten skattefri rejsegodtgørelse eller fradrag på selvangivelsen er, at udstationeringsstedet ikke må have karakter af fast arbejdssted. Det betyder i praksis, at arbejdet på et bygge- og anlægsprojekt ikke må vare mere end 12 måneder. Desuden gælder reglerne kun for lønmodtagere.

Fradrag på selvangivelsen kan enten foregå efter dokumenterede udgifter eller faste satser, hvor de sidstnævnte vurderes at være så attraktive og administrativt enklere, at de udgør den langt overvejende del. Dagssatsen består af én sats for kost og én sats for logi. I nogle af de tidligere år (før 1999) har det dog forholdt sig således, at satserne nedsættes efter en given periode på det samme arbejdssted, en nedsættelse der imidlertid kan omgås ved midlertidigt arbejde andetsteds.

Fradraget på selvangivelsen kan kun opnås, hvis beløbet overstiger en given bundgrænse. I følgende tabel er udviklingen i denne bundgrænse angivet:

Tabel B1. Årlige bundgrænser for »Øvrige lønmodtagerudgifter.«

	Bundgrænse (kr.)
1993	3600
1994	3600
1995	3800
1996	3900
1997	3900
1998	4000
1999	4100
2000	4300
2001	4400
2002	4600
2003	4700
2004	4900

Kilde: Told&Skat.

Der har været større ændringer i reglerne for opnåelse af skattefri rejsegodtgørelse i 1994, 1997 og 1999. I 1994 ændrede man blandt andet reglerne, så det ikke længere var tilladt at reducere den almindelige løn i forbindelse med udbetaling af skattefri rejsegodtgørelse. En større ændring i reglerne blev vedtaget i 1997, hvor det blandt andet blev vedtaget, at udgifter til logi kun kunne fradrages efter dokumentation af faktiske udgifter. Disse ændringer skulle gælde fra 1. april 1999. I december 1999 bliver en ny lov dog vedtaget, som stort set ophæver ændringerne, der blev vedtaget i 1997 med tilbagevirkende kraft for hele 1999. Dvs. der har i hele den periode, vi ser på, været mulighed for at benytte faste satser. Man skal dog være opmærksom på, at fra 1. april 1999 og året ud forventede folk ikke, at der var faste satser for logi.

C. Variable og ML regressioner til afsnit 5

I regressionerne i afsnit 5 anvendes en række variable på arbejdssteds- og individniveau. På arbejdsstedsniveau drejer det sig om: *Log(ansatte)*, der er logaritmen til antallet af samtlige ansatte på arbejdsstedet i november det pågældende år; *gennemsnitlig alder*, *gennemsnitlig erfaring* og *gennemsnitlig tenure* er henholdsvis den gennemsnitlige alder, erfaring og tenure blandt de ansatte håndværkere på arbejdsstedet i november. Tilsvarende er *andel mænd* opgjort som andelen af mænd blandt de ansatte håndværkere i november.

På individniveau drejer det sig om: *Gift*, som er en dummy for, om personen er gift; *børn 0-6 år* og *børn 7-14 år*, som er dummyvariable for om individet har børn i de respektive aldersgrupper. *Alder 25-29*, *alder 30-39*, *alder 40-49*, *alder 50-59* og *alder 60+* er tilsvarende dummyvariable for personens alder (hvor »under 25 år« er den ude-

Tabel C1. Multinomial logit estimationer for antal udedage, 1999-2004.

Udeladt kategori: 0 udedage	Arbejdssteds- niveau		Individ- niveau	
	0-100 udedage	100+ udedage	0-100 udedage	100+ udedage
<i>Arbejdsstedskarakteristika</i>				
Log(ansatte)	1,047 (112,78) ***	0,755 (20,96) ***	1,439 (261,38) ***	1,449 (89,64) ***
Gennemsnitlig alder	-0,029 (-20,78) ***	-0,015 (-2,48) **	-0,012 (-10,15) ***	0,006 (1,07)
Andel mænd	-0,477 (-6,02) ***	6,057 (3,14) ***	-0,516 (-10,40) ***	11,250 (8,54) ***
Gennemsnitlig erfaring	0,032 (17,49) ***	0,055 (7,14) ***	0,025 (15,64) ***	-0,017 (-2,31) **
Gennemsnitlig tenure	-0,042 (-19,55) ***	-0,185 (-13,96) ***	-0,077 (-43,19) ***	-0,327 (-31,54) **
<i>Individkarakteristika</i>				
Gift			-0,010 (-0,90)	-0,160 (-3,54) ***
Børn 0-6 år			-0,041 (-3,22) ***	-0,304 (-5,72) ***
Børn 7-14 år			-0,010 (-0,78)	-0,018 (-0,36)
Alder 25-29			-0,035 (-1,90) *	0,341 (4,30) ***
Alder 30-39			-0,029 (-1,34)	0,591 (6,57) ***
Alder 40-49			0,006 (0,25)	0,719 (7,10) ***
Alder 50-59			0,077 (2,81) ***	0,512 (4,51) ***
Alder 60+			0,145 (3,77) ***	0,883 (5,27) ***
Fultidsansat			0,410 (23,89) ***	1,336 (12,15) ***
Erfaring			-0,012 (-4,38) ***	-0,012 (-1,00)
Erfaring kvadreret			0,000 (2,74) ***	0,000 (-0,41)

forsættes næste side ...

fortsat ...

Tenure		-0,055 (-19,91) ***	-0,016 (-1,19)
Tenure kvadreret		0,003 (18,87) ***	0,001 (0,67)
Ansatt siden 1980		-0,173 (-4,60) ***	-0,234 (-0,73)
<i>Dummyvariable</i>			
Region	JA		JA
Håndværkertype	JA		JA
År	JA		JA
Antal observationer	82412		429868
LR χ^2	22853,55		170638,47
Pseudo R^2	0,1887		0,3274

Note: I regressionerne på arbejdsstedsniveau er den afhængige variabel det gennemsnitlige antal udedage pr. ansat pr. 1. november. *t*-statistikker i parenteser. *, ** og *** betyder, at estimatet er signifikant på hhv. 10%, 5% og 1% niveau.

ladte kategori). *Erfaring og tenure* er individets erfaring og tenure målt i år, mens *ansatt siden 1980* er en dummy for om ansættelsen på arbejdsstedet går længere tilbage end 1980 (det første år i vores data). Endelig er *fuldtidsansat* en dummy for om individet er fuldtidsansat.

I tabellen ovenfor præsenteres resultaterne fra de omtalte multinomial logit regressioner i afsnit 5.

Danske lægers turnusordning

Jens Leth Hougaard

Økonomisk Institut, Københavns Universitetet, E-mail: jens. leth. hougaard@econ.ku.dk

Jacob Michelsen Kolind

Økonomisk Institut, Københavns Universitetet, E-mail: jacob. michelsen. kolind@econ.ku.dk

Lars Peter Østerdal

Økonomisk Institut, Københavns Universitetet, E-mail: lars. peter. osterdal@econ.ku.dk

SUMMARY: A medical internship is mandatory for all medical students graduating from a Danish university. The internships are distributed fairly even amongst most Danish hospitals, m.a. in order to secure a steady inflow of new doctors to all regions of Denmark. When allocating the graduates to an internship by lottery, the National Board of Health takes into account only the candidates preferences over the various internships, for reasons discussed in the paper. However, a problem, relating to the lottery which the NBH uses, is that the lottery itself is not efficient in an economic sense. By way of simulations this article shows how an alternative lottery built on the so-called EPS-algorithm developed by Katta and Sethuraman [Journal of Economic Theory 2006, 131, 231-250] performs better than the lottery used today. This, along with other problematic aspects such as black-market trading of internships, indicates that a replacement of the currently used lottery may be beneficial for the NBH as well as the graduating medical students.

Indledning

Turnusordningen er en betegnelse for det kliniske uddannelsesforløb som medicinske kandidater skal gennemføre efter den medicinske embedseksamen for at erhverve autorisation som selvstændig læge. Den består af et uddannelsesforløb hvor den enkelte læge får ansættelse, dels på et sygehus, dels i almen praksis. Hvert halve år sendes omkring 400 færdiguddannede læger i turnus på landets sygehuse.¹ Af regionale hen- syn og praktiske årsager er turnuspladserne geografisk fordelt over hele landet.

Forfatterne takker Søren Thorning fra Sundhedsstyrelsen og medarbejdere fra sygehusregionerne for fremskaffelse af oplysning om turnusordningen Vinteren 2006/2007, samt en anonym referee for nyttige kommentarer. Desuden en særlig tak til kommende turnuskandidater og kursusansvarlige på Rigshospitalet og Hvidovre Hospital for stor samarbejdsvilje i forbindelse med generering af et datasæt egnet til simuleringer. Korrespondance: Jacob Michelsen Kolind, Økonomisk Institut, Studiestræde 6, 1455 København K.

1. Jf. dokumenterne »Opgørelse over turnusordningen« på: www.sst.dk/Uddannelse/Laeger/Turnus.aspx

Der er mange årsager til, at den enkelte læge har præferencer for særlige sygehuse og derfor har en holdning til turnuspladsernes fordeling. For eksempel kan der være store omkostninger (inklusive sociale omkostninger) forbundet med at skulle flytte fra en ende af landet til en anden for at gennemføre turnusuddannelsen, hvis lægen allerede er etableret med familie eller i et socialt netværk. Alle pladserne er derfor ikke lige attraktive set fra de dimitterede medicinstuderendes synspunkt. Tilsvarende kan man let forestille sig at sygehusene heller ikke er ligeglade med hvilke kandidater de modtager. Problemet er således at forsøge at finde den bedst mulige matching mellem læger og praktikpladser.

En mulighed er at overlade matchingen til de involverede parter selv – altså til markedet. Erfaringer fra f.eks. USA peger imidlertid på, at sådan en løsning sjældent vil være optimal for nogen af parterne. Typisk vil sygehusene nemlig forsøge at hverve de bedste studerende allerede inden de er færdige med studiet. De studerende kommer således let til at binde sig tidligt (fordi de jo ikke ved, om der kommer et bedre tilbud og hellere vil være på den sikre side), og undgår måske derved muligheden for en bedre matching. Således kan man let forstille sig situationer, hvor en læge, der har indgået kontrakt med et hospital, hellere ville have kunnet takke ja til et senere tilbud fra et andet hospital, der tilmed må tage til takke med en anden kandidat, som de finder ringere.²

Der er således meget, der taler for, at man forsøger at finde en central løsning på matchingproblemet. I USA og England har man valgt en fremgangsmåde, hvor både hospitaler og turnussøgende får lov til at rangordne hinanden. En matching foregår derefter centralt i et »clearing house«. Populært sagt består matchingen af en række inklinationsrunder som til et afdansningsbal, der til slut afstedkommer en matching, der er stabil i den forstand, at der ikke findes et par (en turnussøgende og et sygehus), der vil ønske at bryde den matching, der er opnået. Specielt er der således ingen gensidigt fordelagtige byttehandler blandt de turnussøgende, og sygehusene får de bedste kandidater blandt de stabile udfald, se f.eks. Roth (1984, 2003).

Set med danske øje kan man fristes til at sige, at den matching, der opnås, fungerer lidt for godt, fordi dygtige læger typisk matches med store hospitaler i storbyerne. Resultatet vil således ofte være en skæv geografisk fordeling af talentmassen, hvilket man kan frygte vil have konsekvenser (på sigt) for kvaliteten af ydelserne på hospitaler i de geografiske randområder.³

2. Se Roth (1984, 2003).

3. I Sundhedsloven (LOV nr. 546 af 24/06/2005) hedder det i kapitel 1 § 2:

Loven fastsætter kravene til sundhedsvæsenet med henblik på at sikre respekt for det enkelte menneske, dets integritet og selvbestemmelse og at opfylde behovet for

- 1) let og lige adgang til sundhedsvæsenet,
- 2) behandling af høj kvalitet, ...

I rapporten »Sundhedsvæsenets organisering« (2003), betænkningsgrundlaget for Sundhedsloven, hedder det bl.a. om punkt 1 på s. 54. »Målsætningen om fri og lige adgang har også et geografisk aspekt, hvor hensigten er at sikre et nogenlunde ensartet kvalitets- og serviceniveau i landet som helhed.«

Da man i Danmark har et skattefinansieret nationalt sundhedssystem, er der fra politisk side lagt visse begrænsninger på, hvor frit man som planlægger kan designe en fordelingsprocedure, fordi borgerne principielt har krav på en ensartet behandling i hele landet. Der er derfor en politisk modvilje mod at lade sygehusene prioritere de studerende efter f.eks. karakterer, da dette givetvis, jf. de amerikanske tilstande, vil øge skævvridningen i talentmassen mellem provins og universitetsbyer. Også turnuspladsernes geografiske fordeling er en politisk beslutning og må som sådan tages for givet fra et planlægningsmæssigt synspunkt, spredt jævnt ud over landet.

Ud fra et fairness princip er det i Danmark også udelukket at kompensere lægerne økonomisk for at vælge uattraktive turnuspladser eller at udbyde turnuspladserne på en auktion. Situationen kan sidestilles med at studerende generelt ikke subsidieres i deres studievalg eller betaler for deres studier. En auktion ville ellers løse mange problemer set med planlæggeres øje, da udfaldet af en auktion i teorien vil stille alle agenter tilfredse med den plads, de får tildelt, og således ikke ville ønske at bytte med de andre agenter, givet deres betalingsvillighed.

Planlæggeren skal derfor designe en fordelingsprocedure

- der ikke involverer penge,
- hvor kun lægernes egne ønsker påvirker det endelige udfald (og ikke f.eks. deres karakterer eller sygehusenes prioriteter) og
- turnuspladsernes geografiske placering må tages for givet.

Den nuværende fordelingsprocedure opfylder (næsten) de tre ovenstående krav, men mangler til gengæld økonomisk motivation for at være valgt blandt flere mulige fordelingsprocedurer. »Næsten« i parentes har sneget sig ind, fordi den nuværende fordelingsprocedure indirekte forbryder sig mod det første krav: fordelingsproceduren må ikke involvere penge. Måden, proceduren er designet på i dag, åbner nemlig op for at bytte pladser. Et tilbud der ofte benyttes og undertiden involverer penge under bordet blandt de turnussøgende læger. Beløb på op imod 200.000 kr. bliver fra tid til anden nævnt i dagspressen⁴ og på internetfora for medicinstuderende. Alene dette problem ved den nuværende fordelingsprocedure har skabt så stor utilfredshed blandt de medicinstuderende, at der igennem flere år har været diskuteret krav om en ændring af fordelingsproceduren.

I princippet er efterfølgende byttehandler mellem de studerende (med eller uden penge under bordet) jo kun af det gode i den forstand, at de retter op på nogle af de skavanker, der er ved den politisk fastlagte fordelingsprocedure. Problemet er imidlertid, at fordi en turnusplads via sådanne byttehandler går hen og opnår en »markedsværdi«,

4. »Hvide kitler på den sorte børs« af Peter Burhøi, Berlingske Tidende 14. maj 2006.

så kan man frygte, at de studerende fristes til at manipulere med deres præferencer over turnuspladserne. F.eks. kunne man forestille sig, at en studerende, der er født og opvokset i Lægeløse, foretrækker at vende tilbage til Lægeløse Sygehus (der vist ikke eksisterer), men vælger at prioritere Rigshospitalet højest. Der er nu en vis sandsynlighed for at vedkommende rent faktisk får en plads på Riget, og denne plads kan således »sælges«, hvorefter en plads på Lægeløse Sygehus relativt let kan opnås da disse, af uransagelige årsager, er langt mindre eftertragtede. De problemer, der er involveret i at designe fordelingsproceduren, er således mere komplicerede end som så.

I det følgende ser vi først på, hvordan den nuværende danske fordelingsprocedure konkret fungerer og diskuterer de problematiske aspekter ved denne. Bagefter præsenteres en forbedret fordelingsprocedure, der i teoretisk forstand og ved simulationseksperimenter er den nuværende procedure overlegen. Den nye fordelingsprocedurer undgår de designfejl, der ligger i den gamle procedure. Vi konkluderer, at den nye fordelingsprocedure bør erstatte den gamle.

Den nuværende fordelingsprocedure

For de danske turnuskandidater foregår fordelingsproceduren som en lodtrækning. Hvert halve år er der start for nye turnushold på sygehusene, og godt et halvt år før turnusstart melder de næsten færdige medicinstuderende sig til turnusordningen. Når alle tilmeldinger er i hus hos Sundhedsstyrelsen (SST), oprettes et passende antal turnuspladser i 22 turnusområder.⁵ SST betegner dette som fase A af fordelingsproceduren. Før årsskiftet 06/07 var der 16 turnusområder, bestående af de enkelte amter, H:S og Bornholms Regionskommune. I det følgende refererer vi til grupperingen af hospitaler og sygehuse benyttet før 2007 (da vores simuleringer er baseret på data som refererer til disse), men det har ingen betydning for analysen. Efter oprettelsen af turnuspladser foregår en lodtrækning. Hver turnussøgende læge får tildelt et nummer, og skal derefter liste amter, H:S og Bornholms Regionskommune i en prioriteret rækkefølge, efter hvor den turnussøgende ønsker at komme i turnus. Sundhedsstyrelsen betegner dette som fase B af fordelingsproceduren. Efter at alle har indsendt deres prioriteringsskema fordeler SST de turnussøgende efter princippet om, at »det laveste nummer får lov at vælge først«. Et eksempel herpå kunne være den netop færdige læge A, der har fået tildelt nr. 31 og har udfyldt sin prioriteringsliste som følger:

1. H:S
2. Københavns Amt
3. Roskilde Amt
- ...

5. Se <http://www.sst.dk/Uddannelse/Laeger/Turnus.aspx>

Sundhedsstyrelsen har allerede fordelt lægerne med lodtrækningsnumre fra 1-30, da de kigger på As prioriteringsliste. Blandt de allerede fordelte er 26 blevet fordelt til H:S, der har plads til 26, og 3 blevet fordelt til Københavns Amt der har plads til 30. SST kan derfor ikke tildele A sin førsteprioritet, men kan imødekomme hendes andenprioritet. A får derfor tildelt en turnusplads i Københavns Amt.

Efter endt fase B er alle turnussøgende blevet fordelt blandt amter, H:S og Bornholms Regionskommune. Det er derefter op til det enkelte amtslige videreuddannelsesråd at fordele de turnussøgende ud på de enkelte hospitalspladser. Det er kutyme, at dette sker ved at lade de turnussøgende prioritere de enkelte pladser og så fordele dem efter deres oprindelige lodtrækningsnummer.

En krølle på historien er, at man tager særlige hensyn til ægtefæller, samlevende og læsemakkere.⁶ Disse kan efter ønske blive fordelt til samme amt. Dette foregår ved, at kun den ene får tildelt et nummer, og at den anden derefter bare følger med under SSTs fordeling af de turnussøgende. Parret bliver således tildelt det amt, der, når parret står for tur til at blive fordelt, har plads til to, og som parret helst ønsker.

Hvis den enkelte turnussøgende efter fordelingen af en hospitalsplads er utilfreds med, hvor han/hun er endt, kan han/hun bytte med en anden turnussøgende med samme problem. Det er op til den enkelte turnussøgende at finde en person, der vil bytte. Hvis det ikke er muligt at finde en tilfredsstillende byttehandel, må den turnussøgende acceptere udfaldet af fordelingsproceduren. Det er denne byttemulighed, der af nogen bliver brugt til handel med pladser.

Endelig indgår enkelte turnussøgende slet ikke i den generelle lodtrækning, da særlige vilkår gør sig gældende for dem. Det kan f.eks. være værnepligt eller sygdom i den nærmeste familie. Denne gruppe bliver fordelt inden lodtrækningen og kan derfor (set fra en planlæggeres synspunkt) ignoreres under analysen af forskellige fordelingsprocedurer.

Problematiske aspekter ved den nuværende danske turnusordning

Om den danske fordelingsprocedure kan man – i det mindste – sige, at den »virker«. Der kommer færdiguddannede læger ud fra turnusstillingerne. Som antydnet i indledningen er de økonomiske argumenter for at vælge netop den fordelingsprocedure (der er beskrevet i forrige afsnit) imidlertid ikke oplagte, og det er givetvis, fordi der ikke rigtigt er nogen. Utilfredsheden over handlen med turnuspladser kan ses som et symptom på, at fordelingsproceduren, som den bruges i dag, ikke fungerer optimalt.

Det er naturligt at stille sig selv det spørgsmål, hvorfor man ikke bare forbyder byttehandler for at komme problemet med penge under bordet til livs? Svaret ligger i et andet mærkeligt designaspekt ved den nuværende fordelingsprocedure. Valget af turnus-

6. Fra årsskiftet 2006/2007 udgår læsemakkerordningen.

plads foregår over to runder. I den første runde indsnævres den geografiske placering af turnuspladsen til amtsplan, og først i anden runde finder den turnussøgende, hvor han/hun endeligt placeres. Dette kan give anledning til en fordeling af de turnussøgende, der ikke er stabil⁷ (og derved ikke Pareto-optimal). Og i og med at dette kan forekomme, er det i princippet ufornuftigt at forbyde byttehandler. Et eksempel kan kaste lys på problemstillingen: A og B er begge turnussøgende fra Aarhus Universitet, og deres præferencer over sygehuse er som følger (øverst er bedst):

A:	B:
Aarhus Amtssygehus	Aarhus Amtssygehus
...	...
Haderslev	Randers
...	...
Randers	Haderslev
...	...

Ud fra disse præferencer udfylder de prioriteringsskemaerne til SST

A:	B:
Aarhus Amt	Aarhus Amt
...	...
Sønderjyllands Amt	Sønderjyllands Amt

Lodtrækningsnumrene vil, at A på yderste mandat bliver sendt til Aarhus Amt, mens B er uheldig med sit nummer og bliver sendt til Sønderjylland. B får en turnusplads på Haderslev Sygehus og A får, fordi hun som en af de sidste fik lov til at vælge Aarhus Amt, ikke Aarhus Amtssygehus som ønsket med derimod Randers Sygehus. Fordelingsproceduren opnår her et uoptimalt resultat. Ved at A og B bytter plads, bliver de begge stillet bedre.

Man kan også angribe selve lodtrækningsproceduren for ikke at være optimal. For at komme videre med denne pointe er det en fordel at skrælle alt overflødig af vores problemstilling og betragte fordelingsprocedurer inden for rammerne af en abstrakt økonomisk model.

At finde en fordelingsprocedure til turnusordningen hører under den kategori af problemer, som i litteraturen går under navnet »the random assignment problem«, se f.eks. Bogomolnaia og Moulin (2001) og Katta og Sethuraman (2006).

7. Der findes en ombytning af turnuspladserne blandt de medicinske kandidater, der stiller mindst en kandidat bedre og ingen værre.

Vi modellerer problemet med n agenter (de turnussøgende), blandt hvilke n udelelige objekter (turnuspladserne) skal fordeles. Agenterne har mulighed for at angive deres præferencer over objekterne (rangordne objekterne), men enhver form for økonomisk kompensation er ikke tilladt i modellen. Selve proceduren foregår ved at agenterne rangordnes ved en lodtrækning for efter tur at vælge sit foretrukne objekt. I litteraturen går denne måde at matche agenter med objekter under navnet Tilfældigt Sekventielt Valg (Random Serial Dictatorship). Når den lodtrækningsprocedure, der anvendes, benytter en ligefordeling over alle rangordninger af agenter, som det er tilfældet med den nuværende danske fordelingsprocedure, kaldes fordelingsmetoden Tilfældig Prioritet (Random Priority).

Et basalt krav til en fordelingsprocedure er såkaldt ex-post efficiens. Dette begreb dækker over, at alle udfald fra fordelingsproceduren skal være stabile. Der må ikke findes gensidigt fordelagtige byttehandler blandt agenterne, efter at de er blevet parret med et objekt af fordelingsproceduren. Med et krav om ex-post efficiens indsnævres antallet af mulige fordelingsprocedurer drastisk. Et resultat af Bogomolnaia og Moulin (2001) foreskriver, at ethvert stabilt udfald fra en fordelingsprocedure svarer til en rangordning af agenterne og en efterfølgende fordeling af objekter efter princippet højst rangeret vælger først. Med andre ord er det, givet et stabilt udfald og agenternes præferencer, muligt at konstruere en rangordning, der genskaber det stabile udfald ud fra princippet, højst rangeret vælger først. En vilkårlig fordelingsprocedure, der producerer stabile udfald, svarer derfor til en sandsynlighedsfordeling over alle mulige rangordninger af agenterne. At trække en rangordning fra en fordeling og efterfølgende fordele objekter blandt agenter efter princippet højst rangeret vælger først er derfor præcist definitionen på et »Random Serial Dictatorship«. Enhver ex-post efficient fordelingsprocedure er derfor et »Random Serial Dictatorship«.

En ny fordelingsprocedure for den danske turnusordning skal derfor søges blandt de procedurer, der fremkommer, når sandsynlighedsfordelingen over rangordninger ændres. Denne noget tekniske diskussion belyses med fordel ved et eksempel.⁸

Betragt 4 agenter: A1, A2, A3, A4 med strenge præferencer over 4 objekter: O1, O2, O3, O4. Agenternes præferencer følger nedenstående skema:

A1:	O1	>	O2	>	O3	>	O4
A2:	O1	>	O2	>	O3	>	O4
A3:	O2	>	O1	>	O4	>	O3
A4:	O2	>	O1	>	O4	>	O3

8. Eksemplet er taget fra Moulin og Bogomolnaias artikel.

Rangordningen (A1, A2, A3, A4) af agenterne leder gennem princippet højst rangeret vælger først til allokationen:

A1	A2	A3	A4
O1	O2	O4	O3

A1 vælger sit foretrukne objekt blandt alle 4 objekter – altså O1. A2 vælger blandt de 3 resterende objekter O2, O3, O4 sit foretrukne objekt O2 etc.

De $24 = 4!$ mulige rangordninger giver anledning til et antal forskellige allokationer:

Rangordning:	Allokation:			
	A1	A2	A3	A4
(A1, A2, A3, A4)	O1	O2	O4	O3
(A1, A2, A4, A3)	O1	O2	O3	O4
(A1, A3, A2, A4)	O1	O3	O2	O4
(A1, A3, A4, A2)	O1	O3	O2	O4
(A1, A4, A2, A3)	O1	O3	O4	O2
...			...	

En optælling viser at A1 10 ud af 24 gange får O1. Hvis sandsynligheden, for at hver rangordning forekommer, er lige stor, vil A1 få O1 med sandsynlighed $5/12$. En tilsvarende udregning kan foretages for alle kombinationer af agenter og objekter og opsummeres i en såkaldt bistokastisk matrix:⁹

	O1:	O2:	O3:	O4:
A1:	$5/12$	$1/12$	$5/12$	$1/12$
A2:	$5/12$	$1/12$	$5/12$	$1/12$
A3:	$1/12$	$5/12$	$1/12$	$5/12$
A4:	$1/12$	$5/12$	$1/12$	$5/12$

Matricen opsummerer de marginale¹⁰ sandsynligheder for, at en given agent får et givent objekt under »Random Priority« fordelingsproceduren. Betragt en anden sandsynlighedsfordeling over rangordningerne:

9. En bistokastisk matrix en matrix med ikke-negative reelle indgange, hvor alle søjler og rækker summer til 1.

10. Marginal skal forstås i sin sædvanlige sandsynlighedsteoretiske betydning: En indgang i den bistokastiske matrix angiver sandsynligheden for, at en given agent bliver parret med et givent objekt, når man ikke skeler til, hvordan andre objekter fordeles. Modsat simultane sandsynligheder.

Rangordning: Sandsynlighed:

(A1, A3, A2, A4) 1/2

(A2, A4, A1, A3) 1/2

Denne fordeling giver anledning til matricen:

	O1:	O2:	O3:	O4:
A1:	1/2	0	1/2	0
A2:	1/2	0	1/2	0
A3:	0	1/2	0	1/2
A4:	0	1/2	0	1/2

Bemærk at agent A1 har fået større sandsynlighed for at få sin 1. og 3. prioritet på bekostning af en mindre sandsynlighed for at få sin 2. og 4. prioritet. Det er endda muligt at fortolke skiftet i sandsynligheder, som at agent A1 har fået større sandsynlighed for at få sin 1. prioritet på bekostning af en mindre sandsynlighed for at få sin 2. prioritet og større sandsynlighed for at få sin 3. prioritet på bekostning af en mindre sandsynlighed for at få sin 4. prioritet. Den anden sandsynlighedsfordeling stiller dermed agent A1 bedre. Overraskende er det, at ikke bare A1 men alle agenter stilles bedre med den anden sandsynlighedsfordeling. Alle agenterne må foretrække, at lodtrækningen foregår efter den anden sandsynlighedsfordeling frem for »Random Priority«s lige-fordeling.

Af eksemplet fremgår, at selv når man kigger isoleret på fordelingen af turnussøgende ud på amter, H:S og Bornholms Regionskommune og ikke tager højde for problemer, der opstår, når de turnussøgende viderefordes af de amtslige videreuddannelsesråd, er der stadig problemer. Valget af den nuværende fordelingsprocedure kan ikke motiveres økonomisk.

Ovenstående udgør de problematiske aspekter ved den nuværende fordelingsprocedure. Lad os for overblikkets skyld opsummere:

- Opdeling af fordelingsproceduren i to dele – et landsplan og et amtsplan – åbner op for ikke-stabile slutmatchinger. Agenterne må derfor tillades at bytte pladser, hvilket åbner op for penge under bordet og giver et incitament til at angive urigtige præferencer.
- Random Priority algoritmen, der ligger til grund for fordelingsproceduren, er ikke økonomisk efficient. Givet visse familier af præferencer er det muligt at finde en alternativ fordelingsprocedure, som stiller alle agenter bedre ex ante.

En ny fordelingsprocedure

Eksemplet i forrige afsnit peger fremad mod, hvordan en alternativ, bedre fordelingsprocedure skal konstrueres. Da al essentiel information omkring en ex-post efficient fordelingsprocedure kan samles i en bistokastisk matrix, er det naturligt at tage udgangspunkt i disse matricer. Eksemplet viste, hvordan en matrix kunne være bedre end en anden set ud fra agenternes synspunkt. Dette formaliseres i begrebet *stokastisk dominans*.

*En bistokastisk matrix A stokastisk dominerer en bistokastisk matrix, B hvis enhver agents marginale sandsynligheder under A stokastisk dominerer de marginale sandsynligheder under B.*¹¹

Stokastisk dominans er en partiel ordensrelation på rummet af bistokastiske matricer. Maksimale elementer under denne ordning kaldes *ordinalt efficiente* eller o-efficiente. Alternativt er en bistokastisk matrix A o-efficient, hvis der ikke findes en bistokastisk matrix B, der stokastisk dominerer A givet agenternes præferencer. O-efficiens er en økonomisk tilfredsstillende egenskab ved en bistokastisk matrix, der repræsenterer en fordelingsprocedure: Det er umuligt at finde en alternativ fordelingsprocedure, som man ved, at alle agenter vil foretrække frem for den bestående, ud fra de givne rangordninger. Fordelingsprocedurer svarende til o-efficiente bistokastiske matricer er ex-post efficiente.¹²

En fordelingsprocedure, der altid har en o-efficient bistokastisk matrix, er ønskelig. Dermed er fordelingsproceduren i en vis forstand økonomisk efficient. Samtidig er en fordelingsprocedure, der foregår i én runde, også ønskelig, da det fjerner problemet med ikke-stabile slutmatchinger og derved et behov for efterfølgende bytteri. Begge krav er mulige at efterkomme (det andet helt oplagt).

En algoritme udviklet af Katta og Sethuraman (2006) kaldet EPS-algoritmen kan ud fra de turnussøgendes præferencer generere en o-efficient bistokastisk matrix. I vores implementering dekomponeres matricen efterfølgende til en konveks kombination af permutationsmatricer ved Birkhoff von-Neumann algoritmen.¹³ Vægtene fra de-

11. En agents marginale sandsynligheder under A stokastisk dominerer de tilsvarende marginale sandsynligheder under B, hvis

- Sandsynligheden for at agenten får sin 1. prioritet er mindst lige så stor under A som under B.
- Sandsynligheden for at agenten får sin 1. eller 2. prioritet er mindst lige så stor under A som under B.
- Sandsynligheden for at agenten får sin 1., 2. eller 3. prioritet er mindst lige så stor under A som under B.
- ...

12. Lemma 2 (ii) i Bogomolnaia og Moulins artikel.

13. Den udgave af Birkhoff von-Neumann algoritmen, vi benytter, er at finde i Bapat (1997). Både EPS-algoritmen og Birkhoff von-Neumann algoritmen har polynomielle køretider. Køretiden for begge algoritmer er domineret af $O(n^4)$.

kompositionen bruges efterfølgende som sandsynlighedsfordeling over alloktioner (permutationsmatricerne) af agenter til objekter. Altså et lotteri over mulige parringer af turnussøgende med turnussteder. Denne algoritme er derfor et muligt alternativ til Random Priority algoritmen – RP-algoritmen, der ligger til grund for den nuværende fordelingsprocedure.

EPS-algoritmen virker ved at fordele sandsynlighedsmasse i stedet for turnuspladser. Man skal forestille sig, at agenterne inden selve lodtrækningen forhandler om, hvad sandsynligheden for at agent A_i modtager objekt O_j skal være. Det gør de ved at dele objekterne op i infinitesimalt små anparter og så fordele anparterne mellem sig. Ejeren af en anpart modtager ikke, efter lodtrækningsproceduren er gennemført, en fysisk anpart af et objekt, men anparten giver ham en chance for at modtage hele objektet. Anparterne kan altså opfattes som sandsynlighedsmasse.

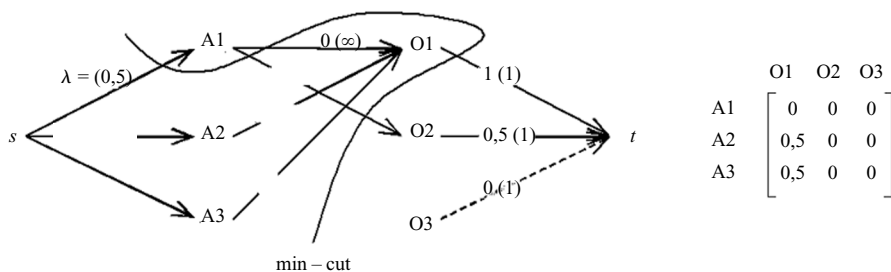
Et eksempel vil vise hvordan EPS-algoritmen fordele anparterne (eksemplet er mere formelt gennemregnet i Katta og Sethuraman (2006)). Vi betragter tre agenter A_1 , A_2 og A_3 med præferencer

A_1 :	O_1, O_2	$>$	O_3		
A_2 :	O_1	$>$	O_2	$>$	O_3
A_3 :	O_1	$>$	O_3	$>$	O_2

Agenterne starter med at blive enige om, hvor stor sandsynlighed de skal have for at få deres førsteprioritet. Der er konkurrence om O_1 , men da agent A_1 er indifferent mellem O_1 og O_2 , lader han de andre deles om O_1 . Indledningsvis kan hver agent få for 50% anparter i O_1 hhv. O_2 .

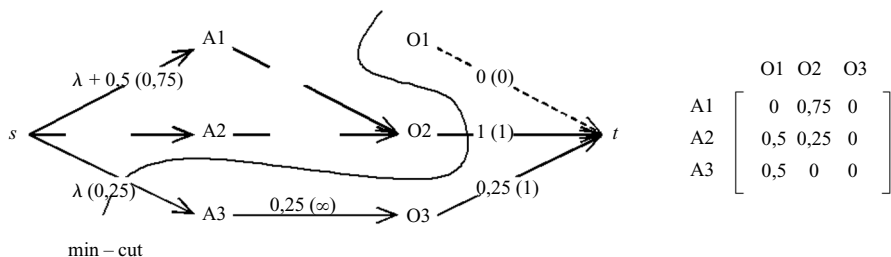
Fordelingen kan man forestille sig foregår i et netværk, hvor tallene i parentes angiver kantens kapacitet, mens tallene før parentes angiver flowet af andele gennem kanten (se figur 1). I første omgang er grafen konstrueret, så agenterne er forbundet til deres førsteprioriteter. Det samlede flow reguleres ved λ og øges, indtil der ikke kan løbe flere anparter igennem grafen.¹⁴ Vi fortolker dette, som at agenterne opnår størst mulig anpart i deres førsteprioritet, givet at alle agenter kan lægge beslag på den samme mængde af anparter. I dette tilfælde opstår den første bottleneck ved objekt O_1 , og man må stoppe med at øge kapaciteten hos agenterne. Anparterne i O_1 fordeles.

14. Den relevante værdi af λ findes ved at løse et parametric max-flow problem. Parametric max-flow algoritmen står beskrevet i Goldberg og Tarjan (1988).



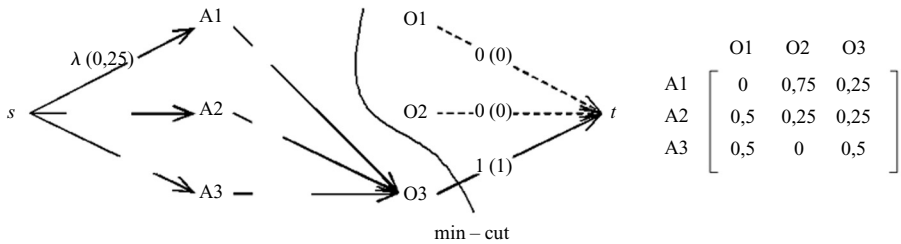
Figur 1.

Agent A1, der ikke fik del i O1, har 0,5 anparter til gode. Han ville gerne have anparterne i objekt O2, men objekt O2 har endnu ikke været en kritisk ressource. Agent A1 får 0,5 anparter forspring i forhold til de andre agenter i næste fordelingsrunde, hvor agenterne byder ind på deres højest prioriterede objekt blandt de tilbageværende objekter.



Figur 2.

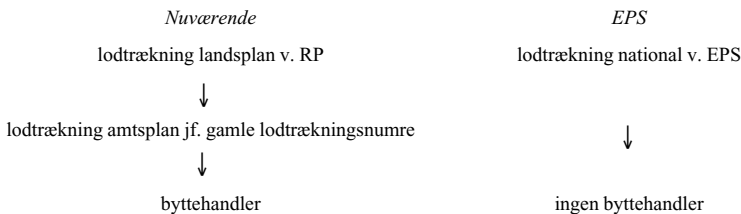
Her bliver objekt O2 først en kritisk ressource, da agent A1 kræver en anpart på 0,75 og agent A2 en anpart på 0,25. O2 fordeles, og nu er det agent A3, som har en anpart på 0,25 til gode. I en sidste fordelingsrunde fordeles objekt O3.



Figur 3.

Den resulterende bistokastiske matrix er o-efficient. En kort forklaring på dette er, at o-efficiens er ensbetydende med, at sandsynlighedsmassen er fordelt optimalt i forhold til agenternes præferencer, og EPS algoritmen tilstræber netop dette ved hele tiden at fordele mere og mere af sandsynlighedsmassen, så den enkelte agent hele tiden får det optimale ud af sin sandsynlighedsmasse.

EPS-algoritmen er grundstenen i en ny fordelingsprocedure. Skemaet nedenfor op-risder den nye fordelingsprocedure og sammenligner den med opbygningen af den nuværende.



Figur 4.

EPS-fordelingsproceduren har flere fortrin frem for den nuværende. Dels er den o-efficient, dels kan den håndtere indifferens mellem turnussteder blandt de enkelte agenter. Desuden er det muligt at vise, at ingen agent misunder en anden agent for den ex-ante sandsynlighedsfordeling over turnuspladser.

At åbne op for at agenterne kan være indifferente mellem turnussteder er nødvendigt, når fordelingen foregår i *et* skridt. Fuldstændig prioritering på turnusstedniveau er svært at forlange, når der er 52 forskellige turnussteder. Ulempen herved er, at det så

i få (teoretiske) tilfælde kan betale sig for agenterne at oplyse urigtige præferencer for at opnå en bedre ex-ante sandsynlighedsfordeling over turnuspladser. Det kræver dog, at den enkelte agent har adgang til de andre agents fulde præferencer, for at han med fordel kan manipulere sine egne præferencer. I praksis er dette aldrig tilfældet.

En ulempe ved EPS-fordelingsproceduren er, at ægtefæller/samlevere ikke kan håndteres af algoritmen. Dette kan løses ad to veje. Enten kan en ekstra plads oprettes ad hoc, hvis et par frem for en enkelt agent tildeles en turnusplads. Ellers kan, som det også hænder under den nuværende ordning, en plads reserveres ved samme turnussted til ægtefællen/samleveren et halvt år senere. De to starter derved forskudt på turnus, men da mange ægtefæller/samlevere alligevel ikke er færdige samtidigt, vil dette næppe være et stort problem i praksis.

Det er værd at understrege, at EPS-fordelingsproceduren ikke stokastisk dominerer den nuværende procedure, men er o-efficient. EPS-fordelingsproceduren er derfor ikke ubetinget bedre end den nuværende procedure, men må antages at performe bedre på sigt. Denne hypotese underbygger vi ved simulationseksperimenter.

Simulation

Inden vi diskuterer de konkrete simulationer, er det nødvendigt at afgøre, hvornår en procedure fordeler agenter »godt«. Vi har med andre ord brug for at kunne sammenligne fordelingsprocedurer ud fra en »bedre-end-relation«. Efter endt lodtrækning offentliggør SST selv en liste med følgende indhold:

- Antallet af agenter der opnåede deres 1. prioritet
- Antallet af agenter der opnåede deres 2. prioritet
- Antallet af agenter der opnåede deres 3. prioritet
- ...

Vi kalder en sådan liste knyttet til udfaldet af en enkelt lodtrækning for en *scoringsliste*. Scoringslister lader sig let konstruere for alle typer af fordelingsprocedurer, og da det ydermere giver mening at tale om det *gennemsnitlige* antal agenter x_i der opnåede deres i 'te prioritet, kan en liste med et sådant indhold nemt genereres ud fra 100 eller 1000 simulationer fra en fordelingsprocedure – simpelthen ved at tage gennemsnit koordinatvis af scoringslisterne genereret ud fra de enkelte udfald. Et gennemsnit af scoringslister vælger vi også at kalde en scoringsliste. Vektorudgaven af store tals lov foreskriver, at sådanne scoringslister vil konvergere mod en fast vektor. En fast vektor som det virker rimeligt at sammenligne de forskellige fordelingsprocedurer ud fra.

Selve sammenligningen af fordelingsprocedurer skal foregå ved Lorenzdominans.

En scoringsliste, (x_i) , over gennemsnitligt opnåede prioriteter er bedre end en anden scoringsliste, (y_i) , hvis:

- $x_1 \geq y_1$
- $x_1 + x_2 \geq y_1 + y_2$
- $x_1 + x_2 + x_3 \geq y_1 + y_2 + y_3$
- ...

Det er klart, at ikke alle scoringslister kan sammenlignes under Lorenzdominans, men en partiel ordning er tilstrækkelig til vores formål. Lad os forestille os, at to fordelingsprocedurer, A og B, tre gange benyttes til at fordele tre agenter:

antal	A			B		
	1. prioriteter	2. prioriteter	3. prioriteter	1. prioriteter	2. prioriteter	3. prioriteter
1. ford.	2	1	0	1	2	0
2. ford.	1	1	1	2	0	1
3. ford.	2	1	0	2	0	1
gns.	5/3	1	1/3	5/3	2/3	2/3

A dominerer B, da:

A	B
5/3	$\geq 5/3$
5/3+1	$\geq 5/3+2/3$
5/3+1+1/3	$\geq 5/3+2/3+2/3$

I gennemsnit fordeles 1/3 flere agenter på deres 2. prioritet frem for deres 3. prioritet under fordelingsproceduren A ceteris paribus.

Ovsnævnte måde at sammenligne fordelingsprocedurer har både fordele og ulemper. Største fordel er nok, at sammenligningen er intuitiv forståelig. Største ulempe er, at det er et aggregeret mål. Alle agenter vægtes lige, og der tages ikke højde for, at springet fra en 2. til en 3. prioritet for en agent kan være meget værre, end springet fra en 4. til en 17. prioritet er for en anden agent. Alligevel virker sammenligning ved Lorenzdominans som et fornuftigt mål for effektiviteten af fordelingsprocedurer, da sammenligningen er tilpas detaljeret til at fange mindre forskelle og ikke så detaljeret, at en sammenligning bliver umulig, som tilfældet er med stokastisk dominans.

Grundlaget for simulationerne er en spørgeskemaundersøgelse foretaget blandt me-

dicinstuderende på Københavns Universitetet i januar 2007, der stod for at skulle i turnus. De studerende blev spurgt om dels at prioritere blandt amter, H:S og Bornholms Regionskommune, dels blandt de enkelte turnussteder. Nedenstående er en liste over turnussteder fordelt på amt, H:S og Bornholms Regionskommune:

<i>Bornholm:</i>	<i>Nordjyllands Amt:</i>	<i>Storstrøms Amt:</i>
Rønne	Aalborg	Næstved
	Dronninglund/Aalborg	Nykøbing Falster
<i>Frederiksborg:</i>	Farsø/Aalborg	Fakse
Hillerød	Frederikshavn	Nakskov
Helsingør	Hjørring	
Frederikssund	Hobro	
		<i>Vejle Amt:</i>
		Horsens
<i>Fyn:</i>	<i>Ribe Amt:</i>	Vejle
Nyborg	Esbjerg/Varde	Fredericia
Middelfart	Grindsted	Kolding
Svendborg		
Faaborg	<i>Ringkøbing Amt:</i>	<i>Vestsjællands Amt:</i>
Odense	Holstebro	Kalundborg
Odense-Svendborg	Herning	Holbæk
	Ringkøbing	Slagelse
		Ringsted
<i>H:S:</i>	<i>Roskilde Amt:</i>	
Bispebjerg	Køge	<i>Viborg Amt:</i>
Hvidovre	Roskilde	Viborg-Kjellerup
Frederiksberg		Sygehus Nord,
Amager		Nykøbing-Thisted
<i>Københavns Amt:</i>	<i>Sønderjyllands amt:</i>	
Glostrup	Haderslev	<i>Århus Amt:</i>
Gentofte	Sønderborg	Silkeborg
Herlev	Tønder	Randers
Amager		Aarhus Amtssygehus
		Århus Kommunehospital

Det er vigtigt at pointere at prioriteringen af amt, H:S og Bornholms Regionskommune skete adskilt fra prioriteringen af turnussteder. Vi valgte at spørge de studerende

på denne måde for at kunne undersøge om de prioriterer konsistent, eller om en situation lignende den, vi har skitseret for de to turnussøgende fra Aarhus Universitet, kan opstå. Det viste sig at ud af 111 udfyldte spørgeskemaer var ikke ét spørgeskema konsistent forstået således, at hvis et amt var prioriteret højere end et andet, ville også alle turnussteder i det højt prioriterede amt være højere prioriterede end turnuspladserne i det lavt prioriterede amt. I spørgeskemaundersøgelsen blev ægtefælle/samleverordningen ignoreret.

111 turnussøgende, alle fra Københavns Universitet, er selvfølgelig ikke en repræsentativ population i forhold til en rigtig turnusrunde, hverken i antal eller prioriteringer. Antallet af normalt oprettede turnuspladser pr. turnussted valgte vi derfor at nedskalere med en faktor ca. 3,5, så antal turnuspladser og antal turnussøgende svarede til hinanden. Dette gjorde vi for at »presse« algoritmerne, så turnussøgende i simulationerne også bliver fordelt på upopulære pladser, mere i tråd med en fordeling ved en rigtig turnusrunde. Men da prioriteringerne stadig ikke er repræsentative for en normal turnuspopulation, er fordelingsresultaterne fra simulationerne derfor ikke direkte sammenlignelige med en rigtig turnusrunde.

Simulationerne sammenholder fire forskellige fordelingsprocedurer:

- RP two-step: Referencefordelingsproceduren. Designet med henblik på at ligne den nuværende fordelingsprocedure så meget som muligt. Ud fra en ligefordeling trækkes en rangordning af agenter. Agenterne fordeles efter deres prioritering af amter, H:S og Bornholms Regionskommune ud fra princippet højest rangeret vælger først. Agenterne beholder deres »lodtrækningsnummer« og bliver derefter i hvert amt, H:S og Bornholms Regionskommune fordelt efter nummeret ud fra deres turnussted-prioritering, igen ud fra princippet højest rangeret vælger først.
- RP one-step: Ud fra en ligefordeling trækkes en rangordning af agenter, som fordeles efter deres prioriteringer af turnussteder ud fra princippet højest rangeret vælger først. Fordelingen af agenter sker i ét skridt, men selve algoritmen, hvorved fordelingen foregår, er identisk med den brugt i første runde af den nuværende fordelingsprocedure.
- EPS one-step: Fordelingen sker som skitseret i sidste afsnit ved hjælp af EPS algoritmen. Algoritmen får som input agenternes prioriteringslister over turnussteder og genererer en o-efficient bistokastisk matrix. Ud fra den bistokastiske matrix trækkes en allokation af agenter på turnussteder.
- EPS two-step: En hybrid mellem EPS one-step og RP two-step. Ud fra en ligefordeling trækkes en rangordning af agenter. Agenterne fordeles efter deres prioritering af amter, H:S og Bornholms Regionskommune ud fra princippet højest rangeret vælger først. Derefter fordeles agenterne på turnuspladser i de enkelte amter, H:S og

Bornholms Regionskommune ved hjælp af EPS algoritmen uden indflydelse fra lodtrækningsnumrene fra første fordelingsrunde. EPS two-step er medtaget, fordi den vil kunne løse det problem, der ligger i, at ægtefælle/samleverordningen ignoreres. Da agenterne fordeles efter RP-metoden på amtsplan, er det muligt at sende ægtefæller/samlevere til samme amt uden komplikationer. Ægtefælle/samleverordningen garanterer ægtefæller/samlevere at ende i samme amt, H:S eller Bornholms Regionskommune.

I simulationssammenhæng ignoreres i alle tilfælde eventuelle efterfølgende fordelagtige byttehandler blandt agenter. Sådanne kan opstå, som vi før har argumenteret, når agenterne fordeles over to runder. Fordelagtige byttehandler kan også opstå på grund af den »tie break rule«, man bliver nødt til at introducere i RP-algoritmen. »Tie break« reglen bruges til at at bestemme, hvilket objekt der allokeres til agent A_i , når agent A_i står for tur til at vælge sit foretrukne objekt, men faktisk har flere »mest foretrukne« objekter.

Alle simulationer blev kørt mindst 100 gange for at opnå tilstrækkelig konvergens af scoringslisterne. I figur 5 er kumulerede summer af scoringslister fra de fire fordelingsprocedurer plottet. y -aksen angiver det gennemsnitlige antal agenter, som mindst opnår en given prioritet. Hvis f.eks. 17 på x -aksen svarer til 57,3 på y -aksen, kan det tolkes som det gennemsnitlige antal agenter, der allokeres til en turnusplads, de har angivet som 1.-17.-prioritet er 57,3. Fortolkningen af plottet er, at hvis grafen for en fordelingsprocedure helt ligger over grafen for en anden, så Lorenzdominerer scoringslisten for den første fordelingsprocedure scoringslisten for den anden.

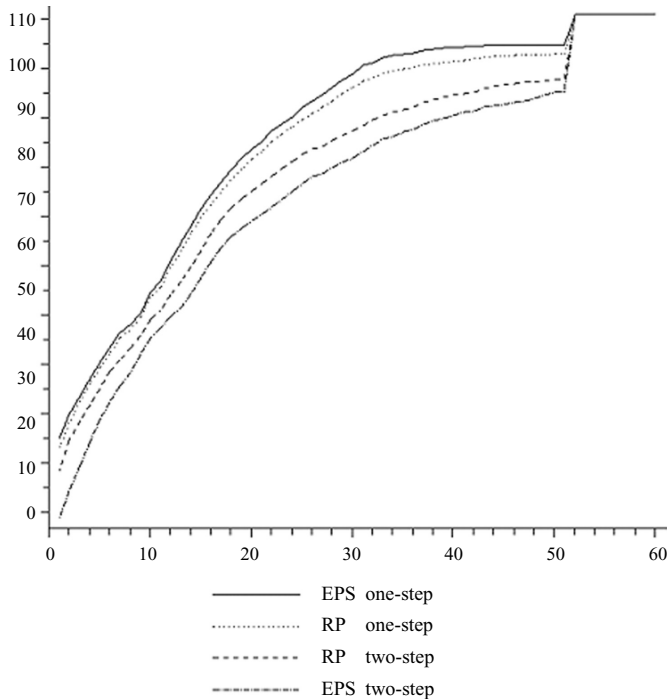
Af plottet fremgår, at de fire fordelingsprocedurer alle står i relation til hinanden og (Lorenzdominans er transitiv):

EPS one-step > RP one-step > RP two-step > EPS two-step

Forklaringen på hoppet i grafen ud for værdien 52 på x -aksen skal søges i, at der er 52 turnussteder. Alle turnussteder, der af agenterne enten ikke er blevet prioriteret eller er prioriteret lavest, er registreret som en 52. prioritet. Dem er der mange af, og en betydelig del af agenterne er derfor blevet tildelt en 52. prioritet.

Konklusion og afsluttende bemærkninger

Ud fra simulationen ser EPS one-step ud til mere effektivt at fordele de turnus-søgende end RP two-step, den nuværende fordelingsprocedure. I det konkrete eksempel opnår 25,0 agenter mod 18,6 deres førsteprioritet. Det er en forbedring på over 25%. Samtidig løser EPS one-step fordelingsproceduren problemerne i forbindelse



Figur 5.

med byttehandler ved at forbyde dem. Og forbuddet kan begrundes økonomisk med, at udfaldet af EPS-fordelingsproceduren er ex-post efficient. Det virker altså fornuftigt at skifte fordelingsprocedure, om end der ikke er noget strengt økonomisk argument for at gøre det i alle tilfælde: EPS one-step stokastisk dominerer ikke konsekvent RP two-step.

RP one-step performer næsten lige så godt som EPS one-step og er derfor et interessant alternativ. Fordelen ved RP one-step er, at algoritmen er nem at implementere. Det kan undre, at SST ikke allerede har valgt denne løsning, da udfaldene af RP one-step er ex-post efficiente, og man derfor ville kunne have forbudt byttehandler. Desuden vil en RP one-step løsning nemt kunne håndtere ægtefæller/samlevere.

Da konklusionerne i nærværende artikel er baseret på et datasæt, der kun indeholder prioriteringslister fra studerende fra Københavns Universitetet, må de tages med visse forbehold. Mere præcise konklusioner kunne opnås, hvis et datasæt for en hel turnusrunde blev anvendt i simulationerne. Ud fra det nuværende datamateriale og den økonomiske analyse af turnusordningens nuværende fordelingsprocedure vil vi dog anbefale, at man tester brugen af den nye procedure for evt. på sigt at skifte helt over til denne.

Endelig kan det nævnes, at det os bekendt er første gang, at EPS-algoritmen er blevet implementeret og forsøgt sammenlignet med RP-algoritmen i en konkret sammenhæng. Sammenligningen af de to »rene« algoritmer er værd at bemærke, da vores resultater underbygger en noget uhåndgribelig teoretisk pointe, om at EPS-algoritmen, da denne er o-efficient, er bedre end den ikke o-efficiente RP-algoritme. Dog skal det også bemærkes, at vi med vores simulationer kun har eftervist denne påstand for et enkelt om end på ingen måde trivielt tilfælde. Også måden hvorved sammenligningen foregår (Lorenzdominans mellem scoringslister) er værd at bemærke. Sammenligningsmetoden er et løsningsbud på det problem, at algoritmerne ikke er sammenlignelige under stokastisk dominans.

Litteratur

- Bapat, R. B, Raghavan, T. E. S. 1997. *Non-negative Matrices and Applications*, Cambridge University Press, New York, 1997.
- Bogomolnaia, A., Moulin, H. 2001. A new solution to the random assignment problem. *Journal of Economic Theory* 100, 295-328.
- Burhøi, P. 2006. Hvide kitler på den sorte børs. Artikel i Berlingske Tidende, 14. maj 2006.
- Goldberg, A. V., Tarjan, R. E. 1988. A new approach to the maximum flow problem. *Journal of the Association for Computing Machinery* 35, 921-40.
- Katta, A.-K., Sethuraman J. 2006. A solution to the random assignment problem on the full preference domain. *Journal of Economic Theory* 131, 231-50.
- Pedersen, K. M. m.fl. 2003. *Sundhedsvæsenets organisering – sygehuse, incitamenter, amter og alternativer*. Rapport fra Indenrigs- og Sundhedsministerens rådgivende udvalg.
- Roth A. E. 1984. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy* 92, 991-1016.
- Roth A. E. 2003. The origins, history, and design of the resident match. *Journal of the American Medical Association* 289, 909-12.

Samfundsøkonomisk afkast af nordiske fiskerier

Max Nielsen, Bruno Cozzari og Jørgen Løkkegaard

Fødevareøkonomisk Institut: E-mail: max@foi.dk, bc@foi.dk og jl@foi.dk

Guri Eriksen og Ola Flaaten

Universitetet i Tromsø, Norge, E-mail: guri.eriksen@nfh.uit.no, olaf@nfh.uit.no

Eyjolfur Gudmundsson

Universitetet i Akureyri, Island, E-mail: eyjolfur@unak.is

Kari Petersen

Færøernes Landsbank, Færøerne, E-mail: kari@landsbank.fo

Staffan Waldo

Livsmedelekonomiska Institutet, Sverige, E-mail: staffan.waldo@sli.lu.se

SUMMARY: The fishing laws of the Nordic countries state that the fish stocks are owned by the Society. This implies that the Society has an opportunity to increase economic returns from this ownership, as is the case in the tenure of other non-renewable resources such as oil and natural gas. In this paper, the economic return from fish stocks is analysed in selected fisheries of five Nordic countries. The size of the Society's return and the allocation between labour, capital and the public authorities are identified. The Societies acquire returns from the fisheries to a different extent. The Society's return is positive in only three out of five case studies, and the public revenue from the existence of fisheries exceeds expenses in only two cases. On this basis, it is examined how the Society may increase the return through improved fisheries management, as well as reallocate the revenue through user payment and taxation.

1. Indledning

»Staten har et saare naturligt Middel til at bringe det hele i Lave, nemlig at opkræve Grundrente af de gode [Fiske] Grunde. Hvis Staten tager Differencen mellem Udbyttet pr. Mand

Korresponderende forfatter: Max Nielsen, Fødevareøkonomisk Institut, Det Biovidenskabelige Fakultet, Københavns Universitet, Rolighedsvej 25, 1958 Frederiksberg C. Tlf.: +45 35336894. Fax: +45 35286801. Forfatterne takker Hans Frost, Fødevareøkonomisk Institut, for at stille modelapparat til rådighed, såvel som Hans Frost og to anonyme referees takkes for værdifulde bemærkninger til tidligere udgaver af artiklen. Nordisk Ministerråds Arbejdsgruppe for Fiskeriforskning takkes for økonomisk støtte.

ved Maksimumsdrift og den gængse løn (+ Omk.), vil der netop komme det Antal Fiskere paa hver [Fiske] Grund, som betinger Maksimumsproduktionen« [citat Jens Warming, Om Grundrente af Fiskegrunde, Nationaløkonomisk Tidsskrift, 1911, s. 501].

Har fiskeriforvalterne i de nordiske lande lært af Jens Warming over det sidste århundrede?

Artiklen »Om Grundrente af Fiskegrunde« af den danske økonom Jens Warming fra 1911 i Nationaløkonomisk Tidsskrift lagde grunden til den moderne fiskeriøkonomi. Dette skete mere end 40 år før Scott Gordon i 1954 publicerede artiklen »*The economic theory of a common property resource: the fishery*«, som i den engelsksprogede litteratur anses som grundlaget for fiskeriøkonomien. Jens Warmings artikel er senere oversat til engelsk, Andersen (1983). Nu, næsten 100 år efter at Jens Warming introducerede begrebet »grundrente« i fiskeriet, er det interessant at undersøge i hvilket omfang, der rent faktisk genereres samfundsøkonomisk afkast i fiskeriet. Har ideerne vundet indpas i fiskeriforvaltningen, så der nu skabes positive samfundsøkonomiske afkast? Eller er der stadig lang vej til at opnå det maksimale samfundsøkonomiske afkast? I nærværende artikel identificeres størrelsen og fordelingen af det samfundsøkonomiske afkast i fem nordiske fiskerier med meget forskellig forvaltning.

Traditionelt har biologiske hensyn udgjort grundlaget for forvaltningen af fiskerierne. Fiskeriforvaltningen i de fleste lande er baseret på princippet om, at bestandene skal fiskes på et biologisk bæredygtigt niveau. Dette niveau er typisk bestemt af et forsigtighedsprincip, hvor risikoen for, at rekrutteringen af fisk svigter, minimeres. Forsigtighedsprincippet anvendes bl.a. i det Internationale Havforskningsråds rådgivning. Forvaltningen har i varierende omfang været rettet mod det økonomiske udbytte for fiskerierhvervet og samfundet. Det er til en vis grad lykkedes at holde fiskeriet på et biologisk bæredygtigt niveau, men det er ikke lykkedes at fokusere på økonomien i fiskerierne.

Dette indebærer, at der i mange fiskerier er et væsentligt potentiale for at forøge det samfundsøkonomiske udbytte af fiskebestandene. Warming (1911) viste, at i et ureguleret fiskeri med fri tilgang og afgang vil en overnormal profit for den marginale fiskerindsats betyde en stigning i fiskeriindsatsen, en reduktion i fiskebestanden og dermed en stigning i fangstomkostningerne pr. fanget kg. fisk. Dette vil til sidst fjerne den overnormale aflønning. Der vil opstå en ligevægt uden nogen ressourcerente, og de marginale enheder af arbejdskraft og kapital vil få en aflønning svarende til aflønningen i andre erhverv, tilpasset arbejdsvilkår og risiko i fiskeriet. Begrænses fiskeriet gennem en forvaltning kan der genereres et positivt samfundsøkonomisk afkast af fiskebestandene, Warming (1911), Gordon (1954) og Clark (1990). Afkastet kan opnås på flere forskellige måder, hvor det afgørende er, at der sker en værdiansættelse af

fiskebestandene, således at eksternaliteten i fiskeriet fjernes. Etablering af ejendomsrettigheder eller opkrævning af betaling for at fiske vil kunne skabe et afkast.

Formålet med nærværende artikel er at identificere det økonomiske afkast af udvalgte og forskelligt forvaltede fiskerier. Samtidig identificeres størrelsen, potentialet og fordelingen af afkastet, og det vurderes, hvordan det økonomiske afkast kan forøges og omfordeles. På baggrund heraf vurderes om fiskeriforvaltere har lært af Jens Warmings »læresætning« i Nationaløkonomisk Tidsskrift i 1911. Grundlaget for artiklen er beregninger baseret på en bio-økonomisk model og resultaterne er vurderet ud fra teorien om fornybare ressourcer i ressourceøkonomien.

Fokus på økonomien i fiskeriet er interessant i en situation, hvor der i mange fiskerier kan være et potentiale for på længere sigt at forøge det samfundsøkonomiske afkast af fiskebestandene. De fleste fiskerier er underlagt forskellige begrænsninger, som omfatter adgang til fiskeri, fartøjskapacitet, kvoter, indsats, redskabsanvendelse og fiskeriområder. Implementeringen af de forskellige forvaltningsinstrumenter er imidlertid forskellig, og det økonomiske afkast varierer. I nærværende artikel vurderes størrelsen og potentialet i fem konkrete fiskerier, som er forvaltet forskelligt.

Som første hypotese anvendes, at der i dag eksisterer et positiv samfundsøkonomisk afkast i de fleste fiskerier, men at det potentielt er muligt at forøge afkastet med en forbedret forvaltning. Det positive afkast skal ses på baggrund af, at mange lande gradvist over de sidste tre årtier har forbedret forvaltningen efter oprettelsen af eksklusive økonomiske zoner, EØZ, UN (1982).

Fordelingen af det samfundsøkonomiske afkast af fiskeressourcer skal vurderes i forhold til, at flere lande i fiskerilovgivningen angiver, at fiskebestandene ejes af samfundet. Fiskerne får i første omgang afkastet, men det generelle skattesystem sikrer beskatning af både arbejdskraft og kapital ved fiskeriaktiviteter, hvorved den offentlige sektor får del i afkastet. Omvendt har den offentlige sektor udgifter til fiskeriet i form af tilskudsordninger, gennemførelse af kontrol, biologiske undersøgelser og rådgivning. Fordelingen af det samfundsøkonomiske afkast er derfor komplekst. Den anden hypotese er, at den offentlige sektor i dag får en beskedent andel af det samfundsøkonomiske afkast i de fleste fiskerier. Afkastet kan principielt omfordeles til den offentlige sektor ved en ressourcebeskatning. En sådan beskatning er imidlertid næsten ukendt i nordiske fiskerier. Kun i Island er der i 2004 indført en afgift på ejerskab af fiskerirettigheder som, når den er fuldt indført i 2008, svarer til 2% af landingsværdien. Til sammenligning opnår den offentlige sektor for andre naturressourcer gennem særlige afgifter en større andel af ressourceafkastet. På norsk olieudvinding er der f.eks. en selskabsskat på 50% af overskuddet, hvilket er udover den generelle selskabsskat, og i Danmark eksisterer en aftale om særskat på olie og naturgas.

2. Metode

Warmings (1911) »Grundrente af fiskegrunde« svarer til ressourcerenten i den moderne fiskeriøkonomiske litteratur. Den angiver »den overnormale profit som opstår ved udnyttelse af fiskebestande og som kan opretholdes på langt sigt«. Ressourcerenten er således det vedvarende økonomiske afkast, samfundet får ud af at eje en fiskebestand. Teorien tilsiger, at ressourcerenten ved frit fiskeri forsvinder, og fiskeriet fortsætter på et niveau, hvor profitten svarer til, hvad der opnås ved andre aktiviteter. Yderligere kan fiskeriet afkaste en producentrente, også benævnt infra-marginal rente, der angiver »det afkast som nogle fiskere opnår, fordi de er dygtigere end andre«. Den infra-marginale rente opnås ved frit fiskeri og medfører et samfundsøkonomisk afkast, der ligeledes opnås i et reguleret fiskeri.¹

Det samfundsøkonomiske afkast af regulerede fiskerier kan måles enten som ressourcerenten alene eller som summen af ressourcerenten og den infra-marginale rente. Anvendes alene ressourcerenten kan optimal skat fastsættes på grundlag af afkastet, idet alle individuelle agenter kan beskattes ens. Afkastet kan også anvendes til at prifsætte individuelle fiskekvoter på et perfekt marked uden transaktionsomkostninger. Anvendes summen af ressourcerenten og den infra-marginale rente opnås et mere fuldstændigt billede af det økonomiske potentiale ved effektiv fiskeriforvaltning, idet størrelsen af den infra-marginale rente også påvirkes af omlægninger i forvaltningen.

Det samfundsøkonomiske afkast af fiskeri beregnes i nærværende artikel som summen af ressourcerenten og den infra-marginale rente, da formålet med artiklen er at identificere gevinsten ved en mere effektiv fiskeriforvaltning. Samtidig kan der ikke i udgangssituationen skelnes mellem ressourcerenten og den infra-marginale rente, da ressourcerenten inkl. den infra-marginale rente beregnes ud fra regnskabsresultater.

Der anvendes følgende to definitioner på samfundsøkonomisk afkast:

- (A) »det nettooverskud, der på et givet tidspunkt er tilbage til aflønning af kapital og arbejdskraft ud over, hvad der opnås i andre erhverv«.
- (B) »det nettooverskud, der på et givet tidspunkt er tilbage til aflønning af kapital og arbejdskraft ud over, hvad der opnås i andre erhverv og efter fradrag af offentlige nettoudgifter til fiskeriet«.

Det samfundsøkonomiske afkast er således defineret før og efter fradrag af offentlige nettoudgifter til fiskeriet (hhv. type A og B). Maksimeringen af det samfundsøko-

1. Endvidere eksisterer et tredje rentebegreb, konsumentrenten, som angiver »forskellen mellem forbrugernes betalingsvillighed og den faktiske pris der betales for en given vare«. Fisk skaber også konsumentrente, men med henblik på at simplificere analysen betragtes konsumentrenten ikke yderligere i nærværende artikel. Samtidig eksporteres hovedparten af fangster fra de fem fiskerier, der betragtes, hvorfor hovedparten af konsumentrenten tilfalder forbrugerne i de lande, der eksporteres til.

nomiske afkast relateres til type A, hvorved usikkerhed ved beregning af ændrede fiskeriforvaltningsomkostninger undgås. Fordelingen af det samfundsøkonomiske afkast relateres til type B, hvorved der opnås et mere fuldstændigt billede af den reelle fordeling, når der tages højde for offentlige nettoudgifter til fiskeriet.

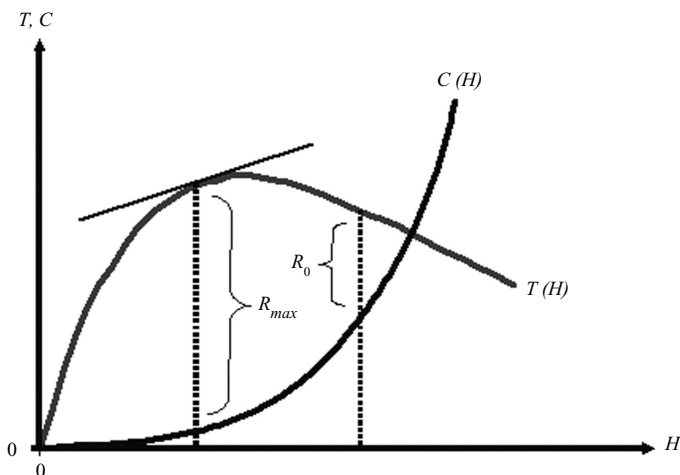
I nærværende artikel bestemmes det samfundsøkonomiske afkast ved en regnskabsmæssig beregning for en given periode. En eksisterende rente kan beregnes uden fradrag af faste omkostninger. Denne rente betegnes quasi-renten. Det nuværende samfundsøkonomiske afkast er imidlertid fradraget faste omkostninger for at kunne sammenligne med en ny fremtidig situation, hvor alle omkostninger er variable. Forudsætningen for, at en sådan sammenligning er realistisk, er, at fiskeriindsatsen er bæredygtig i udgangssituationen, dvs., at det nuværende fangstniveau kan opretholdes med den nuværende fiskeriindsats ud i al fremtid. Er dette tilfældet, og er fangstmulighederne fuldt udnyttet, opnås et mål for de langsigtede økonomiske gevinster ved bedre fiskeriforvaltning. Er fiskeriindsatsen bæredygtig, men fangstmulighederne ikke fuldt udnyttet, kan det samfundsøkonomiske afkast være undervurderet. Er fiskeriindsatsen ikke bæredygtig, kan det samfundsøkonomiske afkast ikke opretholdes på langt sigt.

Det samfundsøkonomiske afkast beregnes som nettooverskud, hvor omkostninger til kapital og arbejdskraft indgår i alternativ anvendelse, som angivet i appendiks. Aflønning af arbejdskraft og kapital i alternativ anvendelse fastsættes på det niveau, som realistisk vurderet, kan opnås ved henholdsvis ansættelse i landbaserede erhverv og ved investering i andre erhverv. Aflønning af kapital i andre erhverv består af rente af investeringer og overskud af virksomheder. Den infra-marginale rente i andre erhverv end fiskeri indgår således i aflønningen af kapital i alternativ anvendelse, hvorved der er sammenlignelighed mellem fiskeriet og andre erhverv. Den investerede kapital opgøres ekskl. værdien af fiskerirettigheder. Fiskerirettigheder behandles i nærværende rapport alene i forbindelse med fordeling af det samfundsøkonomiske afkast, og er ikke direkte inddraget i beregningen af størrelsen af afkastet.

Det maksimale samfundsøkonomiske afkast defineres i nærværende artikel ud fra type A, nettooverskuddet, som:

»det årlige maksimale overskud, der på langt sigt er tilbage efter aflønning af kapital og arbejdskraft ud over, hvad der opnås i andre erhverv«.

Det maksimale samfundsøkonomiske afkast defineres uden fradrag af offentlige nettoudgifter til fiskeriet, idet specielt reguleringsomkostningerne kan ændre sig betydeligt ved overgang til en mere effektiv forvaltning. Ved frit fiskeri kan reguleringsomkostningerne således være små eller nul, hvor tilstedeværelsen af forvaltning nød-



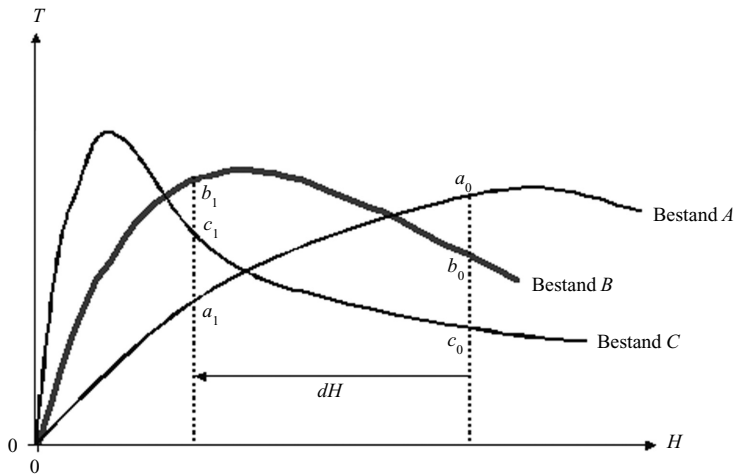
Figur 1. Nuværende og maksimalt samfundsøkonomisk afkast.

vendiggor omkostninger til eksempelvis administration, kontrol og biologiske undersøgelser.

Det maksimale samfundsøkonomiske afkast beregnes i en bio-økonomisk model ved fastlæggelse af den fiskeriindsats, der giver det maksimale overskud på langt sigt for en given fartøjsgruppe, når aflønning af kapital og arbejdskraft måles ved alternativ anvendelse. Modellen er formelt beskrevet i appendiks. Fiskeriindsatsen måles som antal havdage for flåden. Det maksimale samfundsøkonomiske afkast fastlægges ved det antal havdage (H), hvor forskellen mellem indtægter (T) og omkostninger (C) er størst. Dette er illustreret grafisk for et enartsfiskeri i figur 1.

I flerartsfiskerier antages yderligere, at de enkelte fartøjsgruppers fiskeri på forskellige arter er uafhængigt af hinanden, samt at de enkelte fartøjsgruppers relative tidsforbrug på de forskellige bestande er konstant. Det maksimale samfundsøkonomiske afkast bestemmes i flerartsfiskerier ved at måle effekten på det samfundsøkonomiske afkast i fiskeriet på alle bestande med samtidige og lige store ændringer i antallet af havdage. Dette er illustreret grafisk i figur 2, som sammenligner den nuværende situation med en tilfældig ny situation ved reduktion i fiskeriindsatsen.

Det beregnede maksimale samfundsøkonomiske afkast holdes op mod antagelsen om ligevægt i udgangssituationen (R_0), svarende til at være på totalomsætningskurven (T) i figur 1. Dette er en rimelig antagelse, såfremt fiskeriindsatsen i den nuværende situation er bæredygtig og fangstmulighederne fuldt udnyttede. Er fangstmulighederne ikke fuldt udnyttede, er det samfundsøkonomiske afkast undervurderet. Er fiskeriindsatsen ikke bæredygtig, kan fangster og omsætning ikke opretholdes i fremtiden. Konsekvensen er, at det maksimale samfundsøkonomiske afkast er overvurderet.



Figur 2. Nuværende og ændret samfundsøkonomisk afkast i flerartsfiskerier ved ændret fiskeriindsats.

Fordelingen af det samfundsøkonomiske afkast identificeres efter fradrag af offentlige nettoudgifter til fiskeriet (type B) ved at beregne andelen for kapitalen og arbejdskraften i fartøjsgruppen og den offentlige sektor, som beskrevet i appendiks. Kapitalens og arbejdskraftens andel beregnes uden hensyntagen til, om den tilfalder nuværende eller tidligere fiskere. Salg af fiskerirettigheder fra tidligere fiskere til nuværende er ikke inddraget i beregningen.

3. Data

Til belysning af det samfundsøkonomiske afkast er udvalgt fem fiskerier fra nordiske lande, der repræsenterer forskellige forvaltningsregimer, forskellig rentabilitet og lande med forskellig afhængighed af fiskeriet. De fem case studier er trawlerfiskeri i Island, kystfiskeri i Norge, partrawlerfiskeri i Færøerne, stimefiskefiskeri i Sverige og blåmuslingefiskeri i Danmark.

De udvalgte fiskerier forvaltes ved anvendelse af en lang række forskellige instrumenter, omfattende tekniske bevaringsforanstaltninger som maskestørrelse og mindstemål, adgangs begrænsning for fartøjer, kapacitetsbegrænsninger, indsatsforvaltning med havdage samt kvoteregulering. Fiskerirettighederne er i forskelligt omfang omsættelige, enten særskilt eller i forbindelse med salg af fartøjer.

Der eksisterer adgangs begrænsninger og tekniske bevaringsforanstaltninger i alle de fem fiskerier, og fiskeriforvaltningen baserer sig i alle fiskerier på rådgivning om bæredygtigheden målt ved fiskeridødeligheder og gydebiomasser. Det centrale ele-

ment i forvaltningen er i det islandske trawlerfiskeri individuelt omsættelige kvoter, i det norske kystfiskeri individuelt ikke-omsættelige kvoter, i det færøske partrawlerfiskeri havdageregulering, i det svenske stimefiskefiskeri rationsregulering og i det danske muslingefiskeri licensregulering kombineret med kapacitetsbegrænsning. I det islandske fiskeri er yderligere introduceret en afgift på ejerskab af fiskerirettigheder og i det norske kystfiskeri en mindre strukturafgift på landingsværdien.

Data fra fire af de fem case studier er fra 2001-2003, data for det norske kystfiskeri er fra 2003-2004. Data inkluderer fysiske, økonomiske og biologiske data. De fysiske data kendes fra nationale fartøjsregistre og fangststatistikker. Fuldtidsbeskæftigelsen er bemanningen registreret i fartøjsregistre.

Regnskabsstatistik stammer fra nationale statistikker om økonomien i det primære fiskeri. Alternativ aflønning af arbejdskraft er kendt fra de nationale statistikbureauers arbejdsmarkedsstatistikker. Udgangspunktet er den alternative løn, der kan opnås ved relevant arbejde på land. Denne løn beregnes med udgangspunkt i gennemsnitslønnen i landet i alle andre erhverv end fiskeriet, korrigeret for forventede afvigelser i forhold til alternative beskæftigelsesmuligheder i lokalområdet. Alternativ aflønning af kapital er skønnet på grundlag af viden om renter på langt løbende obligationer fra Danmarks Nationalbank og oplysninger om overskud i industrien fra de nationale statistikbureauers regnskabsstatistik. Offentlige udgifter til fiskeriet kendes fra OECD (2004, 2005), fra Færøerne dog fra Færøernes Landsbank (2006). De offentlige udgifter til fiskeriet inkluderer tilskud og fiskeriforvaltningsomkostninger. Disse kendes alene totalt for de nationale fiskerisektorer og andelen beregnes derfor i nærværende analyse ud fra landingsværdien i de forskellige fiskerier. Selskabs- og indkomstskat kendes fra nationale statistikbureauer. Indkomstskatteprocenterne er forskellige i fiskeriet og i alternative erhverv, idet der tages hensyn til bundfradrag, fiskerfradrag, arbejdsgiverudgifter og skattesystemets progression.

Fysiske og økonomiske data er angivet i tabel 1 som årlige gennemsnit.

Af de udvalgte fiskerier er fiskeriet i Island størst og mindst i Danmark, både når der måles i beskæftigelse og omsætning. Blandt de øvrige er det norske kystfiskeri størst og det færøske og svenske på samme niveau. Det svenske stimefiskefiskeri og det islandske trawlerfiskeri udgør næsten halvdelen af de nationale fiskerier, hvor de øvrige er af mindre betydning. Antallet af fartøjer er med 1.145 størst i Norge. Fiskeriet foregår med store fartøjer i Island (bemanning ombord ca. 20) og små i Norge og Danmark (1-2 mand). De svenske og færøske fartøjers bemanning er 4 til 7 personer. Størrelsen af fartøjerne afspejles i kapacitetsudnyttelsen, hvor de islandske og færøske fartøjer er på havet i flest dage om året. De norske fartøjer er aktive færrest dage, bl.a. som følge af at fiskeriet er sæsonbetonet. De svenske fartøjers lave aktivitet hænger

Tabel 1. Fysiske og økonomiske data.

	Island	Norge	Færøerne	Sverige	Danmark
<i>Fysiske data</i>					
Fuldtidsbeskæftigelse	2.100	1.870	322	344	75
Antal fartøjer	76	1.145	29	57	63
Forsikringsværdi pr. fartøj (mio. DKK)	35	2	12	19	2
Antal havdage pr. fartøj	267	171	244	203	203
Andel af total landingsværdi (%)	42	9	21	46	4
Arter	Torsk Rødfisk	Torsk	Sej Torsk Kuller	Sild Brisling Tobis	Blåmusling
<i>Regnskabsdata (mio. DKK)</i>					
Omsætning	2.776	792	302	387	158
Variable omkostninger	599	122	68	132	16
Fartøjsomkostninger	370	151	50	52	16
Afskrivninger	380	57	19	33	12
Lønomkostninger	1.092	422	123	89	55
Kapitalomkostninger	158	36	5	38	0
<i>Alternativ aflønning</i>					
Årsløn i andre erhverv (1.000 DKK)	240	270	183	267	360
Kapitalforrentning i andre erhverv (%)	6	6	6	6	6
<i>Offentlige udgifter til fiskeriet (mio. DKK)</i>					
Tilskud	0	-9	-5	-26	-8
Fiskeriforvaltning	-90	-84	-20	-66	-14
Brugerbetaling for kontrol	41	0	0	0	0
<i>Skat</i>					
Selskabsskat (pct. af overskud)	18	28	20	28	38
Indkomstskat i fiskeri (pct. af br. indk.)	35	23	36	48	50
Indkomstskat i andre erhverv (pct. af br. indk.)	34	33	38	48	41

Kilde: Nielsen m.fl. (2006).

sammen med den anvendte rationsregulering. De udvalgte islandske, færøske og norske fiskerier er målrettet bundfisk, det danske blåmusling og det svenske stimefisk.

Forrentningen af kapital i alternativ anvendelse er antaget ens i alle de udvalgte fiskerier. Aflønningen af arbejdskraft er forskellig. Disse antagelser bygger implicit på, at de internationale kapitalmarkeder er sammenhængende og på begrænset mobilitet af arbejdskraft.

De offentlige udgifter til de udvalgte fiskerier består dels af offentlige tilskud til op-hør og ophugning af fartøjer, fornyelse af flåden, modernisering af fartøjer, omkostningsreducerende transfereringer til fiskeriet og i Færøerne indkomststøtte til fiskere. Dels af omkostninger til fiskeriforvaltning, herunder til administration, kontrol, biolo-

Tabel 2. Faktisk gydebiomasse, gydebiomasser der giver de største vedvarende fangster og fiskeridødelighedsniveauer, tons.

	Kilde	Gydebiomasse		Fiskeridødelighedsniveau
		Faktisk 2003	Maksimal ¹	
<i>Danmark:</i>				
Blåmusling	DFU (2003)	363.000	300.000	Bæredygtigt
<i>Sverige:</i>				
Sild (Østersøen)	STECF (2003)	618.000	1.004.000	Risiko for ikke at være bæredygtigt i øst Referencepunkter ikke defineret i vest
Sild (Skagerrak-Kattegat)	ICES (2003)	2.334.000	2.061.000	Bæredygtigt
Brisling	ICES (2003)	871.000	1.018.000	Bæredygtigt
Tobis	ICES (2003)	211.000	600.000	Referencepunkter ikke defineret
Makrel	ICES (2003)	1.821.000	2.300.000	Ikke bæredygtigt
Blåhvilling	ICES (2003)	5.730.000	2.250.000	Ikke bæredygtigt
<i>Norge:</i>				
Torsk	ICES (2003)	642.000	460.000	Højere end intention i forvaltningsplan
Kuller	ICES (2003)	126.000	80.000	Bæredygtigt
Sej	ICES (2003)	448.000	150.000	Bæredygtigt
<i>Island:</i>				
Torsk	ICES (2003)	189.000	179.000	Fiskeridødelighed for høj
Kuller	ICES (2003)	147.000	78.200	Referencepunkter ikke defineret
Sej	ICES (2003)	129.000	107.000	Forøget risiko for ikke at være bæredygtigt
<i>Færøerne:</i>				
Torsk	ICES (2003)	51.000	40.000	Ikke bæredygtigt (Færø plateau)
Kuller	ICES (2003)	115.000	55.000	Forøget risiko for ikke at være bæredygtigt
Sej	ICES (2003)	83.500	85.000	Ikke bæredygtigt

Note: (1) Den gydebiomasse, der giver de største vedvarende fangster.

giske undersøgelser og rådgivning. I Island er der samtidig indført brugerbetaling for visse kontrolomkostninger.

Indkomstskatteprocenten viser, at der er bundfradrag og progression i skattesystemet i alle de nordiske lande, samt at dette alene for Danmark og Island slår igennem i højere indkomstskat i fiskeriet end i andre erhverv.

Biologiske data kendes fra det Internationale Havforskningsråd (ICES), som årligt rådgiver om fiskeriforvaltningen for de vigtigste fiskebestande i det nordøstlige Atlanterhav. Herudover anvendes data, som indgår i rådgivningen fra Kommissionens rådgivende tekniske, videnskabelige og økonomiske udvalg for fiskeri (STECF). For det danske muslingefiskeri anvendes bestandsvurderinger foretaget af Danmarks Fiskeriundersøgelser. Der anvendes tal for den nuværende gydebiomasse og den gydebiomasse, der giver de maksimale vedvarende fangster samt tal for faktiske fangster. Den gydebiomasse, der giver de størst mulige fangster på langt sigt, forudsættes at svare til det maksimale niveau, når forsigtighedsprincippet anvendes som grundlag for fiskeriforvaltningen. Dette vurderes af ICES. Gydebiomasse og fiskeridødelighedsniveau for bestandene i de fem udvalgte fiskerier er angivet i tabel 2. I Island anvendes forsigtighedsprincippet ikke, hvorfor den gydebiomasse, der giver de størst mulige vedvarende fangster, er fastsat som det gennemsnitlige niveau de sidste ti år.

Gydebiomasserne for bestandene i det islandske, færøske (ekskl. sej), danske og norske fiskeri har i 2003 i alle tilfælde fuld reproduktionskapacitet, da de nuværende gydebiomasser er større end det niveau, der giver de største vedvarende fangster på langt sigt (Maximum Sustainable Yield). For Sverige er dette kun tilfældet for blåhvilning og sild i Skagerrak og Kattegat.

Det fremgår endvidere af tabel 2, at antagelsen om, at fiskeridødeligheden i udgangssituationen er bæredygtig, er rimelig i Danmark og til dels i Norge, men ikke i Island, Færøerne og Sverige. Dette indebærer, at det beregnede maksimale samfundsøkonomiske afkast i Danmark er undervurderet, såfremt fangstmulighederne er underudnyttet. Er fangstmuligheden fuldt udnyttet, er afkastet bestemt præcist. For Norge er indikationerne ikke entydige, hvor det beregnede maksimale samfundsøkonomiske afkast i de tre øvrige lande kan være overvurderet, da fiskeriindsatsen ikke er bæredygtig i udgangssituationen.

4. Resultater

Størrelsen af det samfundsøkonomiske afkast som årligt gennemsnit over 2001-2003 er angivet for de fem udvalgte fiskerier i tabel 3.

Det samfundsøkonomiske afkast før fradrag af offentlige nettoudgifter til fiskeriet (type A) er med 765 mio. DKK størst i det islandske trawlerfiskeri, på samme niveau i det færøske partrawlerfiskeri og det danske muslingefiskeri (henholdsvis 84 og 80 mio. DKK), 12 mio. DKK i det svenske stimefiskefiskeri og minus 176 mio. DKK i det norske kystfiskeri. Målt i forhold til landingsværdien er det samfundsøkonomiske afkast med 60% størst i det danske muslingefiskeri, på niveau i Island og Færøerne (henholdsvis 30% og 28%), 3% i Sverige og minus 22% i Norge.

Det samfundsøkonomiske afkast efter fradrag af offentlige nettoudgifter til fiskeriet

Tabel 3. Størrelsen af det samfundsøkonomiske afkast 2001-2003, mio. DKK og pct. af fangstværdi.

	Island	Norge	Færøerne	Sverige	Danmark
Samfundsøkonomisk afkast (type A)	765	-176	84	12	80
– pct. af fangstværdi	30%	-22%	28%	3%	60%
<i>Offentlige nettoudgifter¹:</i>					
Tilskud	0	-9	-5	-26	-8
Fiskeriforvaltningsomkostninger ²	-90	-84	-20	-662	-14
Brugerbetaling for kontrol	41	0	0	0	0
Total	-49	-93	-25	-92	-22
– pct. af fangstværdi	2%	12%	8%	23%	17%
Samfundsøkonomisk afkast (type B)	716	-269	59	-80	58
– pct. af fangstværdi	28%	-34%	20%	-20%	44%

Noter: (1) De angivne udgifter er landets samlede offentlige udgifter til fiskeriet vægtet med fartøjsgruppernes andel af samlede landinger. (2) Fiskeriforvaltningsomkostninger i Danmark er estimerede på grundlag af fartøjsgruppernes andel af de totale omkostninger i EU.

Kilder: OECD (2004, 2005) og Færøernes Landsbank (2006).

tegner samme relative fordeling mellem landene. Den største relative ændring ses i det svenske stimefiskefiskeri, hvor det samfundsøkonomiske afkast som følge af tilstedeværelsen af store offentlige nettoudgifter til fiskeriet ændrer sig fra at være 12 til minus 80 mio. DKK.

Det samfundsøkonomiske afkast er beregnet under antagelse af, at fiskeridødeligheden er bæredygtig i udgangssituationen. Dette er, jf. tabel 2, rimeligt i Danmark og til dels i Norge, men ikke i Island, Færøerne og Sverige. Konsekvensen er, at det samfundsøkonomiske afkast i Island, Færøerne og Sverige kan være overvurderet.

Det store samfundsøkonomiske afkast i det danske muslingefiskeri skyldes en restriktiv licensregulering, hvor tilgangen af fartøjer såvel som kapacitetsforøgelser af eksisterende fartøjer i flere årtier har været reguleret. Også anvendelsen af periodelukkninger og kvoter er af betydning. I sammenligning med de øvrige fire fiskerier hænger den ekstraordinært gode økonomi i det danske muslingefiskeri sammen med, at omkostningerne ved fiskeriet er lave, da bestandene er stationære, samtidig med, at der i fiskeriet ikke er væsentlig overkapacitet.

I det islandske trawlerfiskeri og det færøske partrawlerfiskeri formår forvaltningen med henholdsvis individuelt omsættelige kvoter og havdage at generere forholdsvis store samfundsøkonomiske afkast. Det færøske resultat skal ses i sammenhæng med, at perioden 2001-2003 biologisk var gode år med ekstraordinær stor rekruttering af fisk. Det islandske resultat kan primært forklares ved den strukturelle økonomiske ud-

vikling i fiskeriet over en årrække, selvom den vigtige torskebestand i perioden var historisk lav.

I Sverige skal det beskedne samfundsøkonomiske afkast ses i forhold til en økonomisk mindre effektiv forvaltning med rationer, som i kortere perioder giver adgang til at fiske en given mængde. Dette indebærer i et fiskeri med overkapacitet, at de faste omkostninger er væsentlige, uden at der er mulighed for at udnytte kapitalapparatet optimalt.

Det negative samfundsøkonomiske afkast i det norske kystfiskeri er overraskende, idet det samfundsøkonomiske afkast selv i et ureguleret fiskeri i langsigtet ligevægt ifølge teorien ikke forventes at være negativ. Der er flere årsager. For det første har offentlige tilskud historisk medvirket til at opbygge overkapacitet i Norge. Siden 1990 er offentlige tilskud til fiskeriet blevet reduceret, så de i dag er tæt på nul, Hermansen og Flaaten (2004). Den nuværende regulering af den norske flåde med individuelle ikke-omsættelige kvoter har ikke medvirket i tilstrækkeligt omfang til at reducere denne overkapacitet. For det andet kan den anvendte metode stille det norske kystfiskeri i et uforholdsmæssigt dårligt lys. Det skyldes antagelsen om, at fiskeridødeligheden i udgangssituationen er bæredygtig, hvilket er rimeligt for Norge. For Island, Færøerne og Sverige er fiskeridødeligheden i udgangssituationen bæredygtig i mindre grad end i Norge, hvorved forskellen mellem de samfundsøkonomiske afkast i Norge og disse lande overvurderes.² For det tredje, at reguleringen af de bestande, som kystfiskeriet baseres på, ikke er tilstrækkelig restriktiv. Bestandene er fælles med Rusland og russiske fartøjer driver et betydeligt overfiskeri, Fiskeridirektoratet i Norge (2006). Disse forhold kan potentielt forklare et afkast, der er tæt på nul. Kombineres dette endvidere med, at mulighederne for at finde arbejde i andre erhverv er lille i de små og afsidesliggende fiskerisamfund langs den norske kyst, forekommer et negativt samfundsøkonomisk afkast realistisk. Samtidig finder nogle fiskere det mere attraktivt at arbejde på havet end på land, og mobiliteten på arbejdsmarkedet er således beskednen.

Det maksimale samfundsøkonomiske afkast er beregnet i tabel 4, hvor det nuværende afkast er angivet til sammenligning. Samtidig er den nødvendige flådetilpasning for at opnå det maksimale samfundsøkonomiske afkast vist.

Det maksimale samfundsøkonomiske afkast er størst i det danske muslingefiskeri (69%), og på samme niveau i Island (60%), Sverige (53%) og Færøerne (47%). Det høje

2. Metoden for beregning af ressourcerente kan også stille det norske kystfiskeri med små fartøjer i et uforholdsmæssigt dårligt lys af andre årsager. Således er eksempelvis kapitalafflønningen i alternativ anvendelse hele tre gange så stor som i faktisk anvendelse, såvel som aflønning af arbejdskraft er 25% højere i alternativ end faktisk anvendelse. For Island, som har den højeste ressourcerente af flerartsfiskerierne, er forholdet helt modsat. Aflønning af kapital og arbejdskraft i alternativ anvendelse er kun halvdelen af aflønningen i faktisk anvendelse.

Tabel 4. Det maksimale samfundsøkonomiske afkast og flådetilpasning.

	Island	Norge	Færøerne	Sverige	Danmark
<i>Samfundsøkonomisk afkast</i>					
<i>(mio. DKK):</i>					
Nuværende	765	-176	84	12	80
Maksimal	1.533	142	142	210	92
Potentiel stigning	768	318	58	198	12
<i>Samfundsøkonomisk afkast</i>					
<i>(i % af landingsværdi):</i>					
Nuværende	30%	-22%	28%	3%	60%
Maksimal	60%	18%	47%	53%	69%
Potentiel stigning	30%	40%	19%	50%	9%
<i>Antal fartøjer:</i>					
Nuværende	76	1.145	29	57	63
Maksimal	27	310	14	15	41
Reduktion	-49	-835	-15	-42	-22
– heraf reduktion af fartøjer i fartøjsgruppen	-46	-737	-12	-35	-18
– heraf kapacitetsudnyttelse (300/225 fiskedage om året)	-3	-98	-3	-7	-4

maksimale samfundsøkonomiske afkast i det danske case skal ses i sammenhæng med, at omkostningerne ved at fiske muslinger er små, og at antallet af fartøjer er lille.

Det samfundsøkonomiske afkast kan forøges i alle de fem udvalgte fiskerier uden at ændre den nuværende flådestruktur. Den potentielle stigning er størst i Sverige, hvor det samfundsøkonomiske afkast kan forøges med, hvad der svarer til 50% af landingsværdien. Også i Norge (40%) og i Island (30%) kan der opnås væsentlige forøgelser i afkastet. I Færøerne (19%) og i Danmark (9%) er de potentielle forøgelser mindre. Gevinsten vurderes endvidere at kunne forøges yderligere, såfremt der føres en politik, der ikke sætter begrænsninger på størrelsen og derved effektiviteten af fiskerifartøjer.

Det maksimale samfundsøkonomiske afkast opnås ved at reducere antallet af fartøjer. I Danmark reduceres med ca. en tredjedel, i Færøerne halveres antallet, mens det reduceres med 2/3 eller mere i de øvrige fiskerier. I flerartsfiskerierne kommer hovedeffekten fra reducerede omkostninger og stigende fangster som følge af flådereduktion. Bedre kapacitetsudnyttelse af fartøjerne giver kun en mindre effekt.

Det maksimale samfundsøkonomiske afkast er beregnet under antagelse af, at fiskeridødeligheden er bæredygtig i udgangssituationen. Dette er, jf. tabel 2, rimeligt i Danmark og til dels i Norge, men ikke i Island, Færøerne og Sverige. Konsekvensen

Tabel 5. Fordeling af det samfundsøkonomiske afkast, 2001-2003, mio. DKK og pct¹.

	Island	Norge	Færøerne	Sverige	Danmark
Kapital	24	-22	28	27	56
Arbejdskraft	52	-5	76	-12	0
Den offentlige sektor	24	-73	-4	-126	24
Total	100	-100	100	-100	100

Note:(1) Minus angiver, at sektoren bidrager negativt til det samfundsøkonomiske afkast.

er, at det maksimale samfundsøkonomiske afkast i Island, Færøerne og Sverige kan være overvurderet.

Fordelingen af det samfundsøkonomiske afkast er angivet i tabel 5 efter fradrag af offentlige nettoudgifter til fiskeriet. Kapitalens og arbejdskraftens andel er angivet efter fradrag af henholdsvis selskabsskat og indkomstskat. Den offentlige sektors andel består af skatteprovenu fra selskabs- og indkomstbeskatning, samt af indtægter fra brugerbetaling, fradraget tilskud og omkostninger til fiskeriforvaltning.

Den offentlige sektors andel af det samfundsøkonomiske afkast er positiv i Island og Danmark og udgør 24% af landingsværdien. Den offentlige sektors andel er negativ i de øvrige tre lande, mest i Sverige (-126%) og Norge (-73%) og mindre i Færøerne (-4%). I disse lande er fiskerisektorens bidrag til den offentlige sektor således mindre, end den er i andre sektorer, dvs., at det offentlige har relativt flere udgifter end indtægter fra fiskeriet sammenlignet med andre sektorer.

Kapitalens andel af det samfundsøkonomiske afkast er med 56% størst i Danmark og ca. på det halve niveau i Færøerne, Sverige og Island. Forrentningen efter skat er otte gange større end i andre erhverv i Danmark, 133% større i Island, dobbelt så stor på Færøerne og 46% større i Sverige. I Norge er kapitalens andel negativ. Kapitalafløningen er således væsentligt større i fiskeriet end i andre erhverv i alle de udvalgte fiskerier bortset fra kystfiskeriet i Norge.

Beregnes arbejdskraftens andel af det samfundsøkonomiske afkast er den med 76% størst i Færøerne, men også højere end i andre erhverv i Island og Danmark. Lønnen efter skat er i Færøerne og i Island mere end dobbelt så stor som i andre erhverv (henholdsvis 117% og 111% større). I Danmark er aflønningen 72% større. I Norge og Sverige er lønniveauet i de udvalgte fiskerier nogenlunde på niveau med andre erhverv.

5. Diskussion

Konsekvensen af, at det maksimale samfundsøkonomiske afkast i alle de udvalgte fiskerier er større end det nuværende, er, at der i de analyserede fiskerier er økonomisk potentiale. Skal dette potentiale realiseres er »prisen«, at muligheden for at prioritere

regionalpolitiske mål om beskæftigelse, bestemte bosætningsmønstre og tilstedeværelse af eksempelvis kystfiskeri med mindre fartøjer, begrænses.

Instrumentet til forøget samfundsøkonomisk afkast er bedre fiskeriforvaltning. Det grundlæggende princip er, at der skal ske en værdiansættelse af fiskebestandene, således at eksternaliteter fjernes. Fiskebestande skal forvaltes bæredygtigt, flådekapacitet, målt som investeret kapital, skal bringes i overensstemmelse med økonomi og biologi, og der skal sikres automatisk strukturtilpasning. Større samfundsøkonomisk afkast kan sikres på flere måder, eksempelvis gennem etablering af ejendomsrettigheder eller ved introduktion af betalingssystemer.

I de danske og islandske case er der i dag større økonomiske afkast end i andre erhverv. Det nuværende samfundsøkonomiske afkast i det danske muslingefiskeri er tæt på det maksimale. Årsagen er en restriktiv adgangs- og kapacitetsbegrænsning kombineret med område- og periodelukninger. Det nuværende samfundsøkonomiske afkast i det islandske trawlerfiskeri er ca. halvdelen af den maksimale. Det allerede opnåede økonomiske afkast kan således fordobles og forventes at kunne realiseres uden væsentlige ændringer i den nuværende forvaltning. Grunden til, at det samfundsøkonomiske afkast i dag ikke er tættere på det maksimale, på trods af at individuelt omsættelige kvoter blev introduceret for ca. 25 år siden, er ikke entydig. En forklaring kan være, at der ved introduktionen var væsentlig overkapacitet, samtidig med at strukturtilpasning tager tid. At de biologiske forhold var ugunstige i perioden med faldende rekruttering, kan ligeledes være af betydning.

I det norske kystfiskeri er det samfundsøkonomiske afkast negativt, hvilket kan forklares af tilstedeværelsen af overkapacitet i den norske fiskeflåde uden en velfungerende strukturtilpasning, kombineret med utilstrækkelig norsk-russisk regulering af fiskeriet i Barentshavet. Såfremt det maksimale økonomiske afkast på 1/5 af landingsværdien i dette fiskeri skal opnås, er et regimeskrift fra det nuværende ikke-omsættelige kvotesystem nødvendigt.

Resultaternes anvendelse på andre fiskerier afhænger af tilstedeværelsen af den konkrete type fiskeriforvaltning. Generelt vurderes der kun at være økonomisk effektiv forvaltning i forholdsvis få fiskerier globalt. Eksempelvis i visse fiskerier i New Zealand, Holland og Island med individuelt omsættelige kvoter, som under ideelle forudsætninger kan fungere økonomisk effektivt. Globalt set er forvaltningen af de fleste fiskerier et sted imellem en form for reguleret åben adgang og reguleret restriktiv adgang. Disse systemer kan gennem effektiv forvaltning skabe det maksimale samfundsøkonomiske afkast, men kan også skabe et negativ afkast. Dette afhænger af den konkrete forvaltning. De fem case studier repræsenterer variationen i forvaltningen af de globale fiskerier. En måling af det samfundsøkonomiske afkast i udvalgte fiskerier i dag, næsten 100 år efter at Jens Warming (1911) introducerede »Grundrente af Fiske-

grunde« indikerer således, at ideerne om fiskeriforvaltning med henblik på at generere samfundsøkonomisk afkast har vundet frem, men at der stadig er lang vej igen, før det fulde økonomiske potentiale er opnået.

Konsekvensen af, at den offentlige sektors andel af det samfundsøkonomiske afkast er negativ i tre af de fem udvalgte fiskerier, er, at den offentlige sektor gennem omfordeling potentielt kan opnå større andel af det samfundsøkonomiske afkast. Instrumenter til omfordeling omfatter salg af fiskerirettigheder f.eks. på auktion, opkrævning af fiskeriforvaltningsomkostninger, reduktion af tilskud og fiskerifradrag, samt afgifter på landinger og ejerskab af fiskerirettigheder.

Den identificerede fordeling af det samfundsøkonomiske afkast viser, at der er et potentiale for, at den offentlige sektor kan forøge sin andel. Kapitalens og arbejdskraftens andel af det samfundsøkonomiske afkast er imidlertid beregnet uden at se på fordelingen mellem nuværende og tidligere fiskere. Såfremt tidligere rettighedshavere enten har fået tildelt fiskerirettigheder gratis (førstegenerations rettighedshavere) eller har købt rettigheder på et tidspunkt, hvor det samfundsøkonomiske afkast var mindre og prisen på rettigheder derfor lav, har de fået »foræret« ressourcerenten. Når nye generationer fiskere skal ind i fiskeriet, skal de købe deres rettigheder enten særskilt eller tilknyttet fartøjerne og betale de tidligere rettighedshavere. Derved er det kun de »gamle« generationer af fiskere, der vinder ved tildeling af eksklusive rettigheder, nyere generationer af fiskere vinder kun, hvis det samfundsøkonomiske afkast stiger.

I eksempelvis de norske stimefiskefiskerier blev licenser for notfartøjer indført allerede i 1973 og i kombination med ophørstilskud samt delvis omsættelighed af licenser, blev antallet af fartøjer reduceret til det halve i løbet af ca. 10 år. I 1983 var 75% af de tilbageværende fartøjer fortsat ejet af fiskere, som havde fået tildelt licenserne gratis, mens 25% var ejet af fiskere, som havde købt sig ind i fiskeriet som andengenerations rettighedshavere. De førstnævnte fartøjer havde overskud, hvor de sidstnævnte havde underskud, Flaaten m.fl. (1995). Overgangen fra første- til andengenerations rettighedshavere med betalte rettigheder kan således tage lang tid.

Såfremt der introduceres et princip om at »ressourcerenten kan alene opkræves fra fiskere, der har fået tildelt den«, bliver potentialet for at forøge den offentlige sektors andel af det samfundsøkonomiske afkast reduceret. Tidligere generationer af fiskere har fået rettighederne foræret, og det bliver gradvist for sent at opkræve deres andel. Under anvendelse af dette princip kan ressourcerenten alene opkræves fra de nuværende fiskere, hvis den har været stigende, efter at de gik ind i fiskeriet. Muligheden er dog at opkræve ressourcerenten fra fremtidige fiskere ved at varsle omfordeling nogle år i forvejen, således at markedet for fiskerirettigheder kan indrette sig herefter. Såfremt omfordeling introduceres sideløbende med, at forbedringer af fiskeriforvaltningen forøger ressourcerenten, kan den offentlige sektor på længere sigt opnå en større andel af det samfundsøkonomiske afkast.

Litteratur

- Andersen, P. 1983. On rents of fishing grounds: a translation of Jens Warming's 1911 article with an introduction, *History of Political Economy*, 15 (3), 391-6.
- Clark, C. W. 1990. *Mathematical Bioeconomics – the optimal management of renewable resources*, second edition, John Wiley & Sons, Inc.
- Copes, P. 1972. Factor Rents, Sole Ownership and the Optimum Level of Fisheries Exploitation, *The Manchester School of Social and Economic Studies*, 40, 145-63.
- Fiskeridirektoratet i Norge. 2005. *Lønnsomhedsundersøgelse for fiskeflåten 2004 – helårsdrevne fiskefartøyer i størrelsen 8 meter største længde og over*, tilgængelig på <http://www.fdir.no/>.
- Fiskeridirektoratet i Norge. 2006. *Status report for 2005 – Russian cod and haddock fishing and transshipment at sea*, Bergen, March, tilgængelig på <http://www.fdir.no/>.
- Flaaten, O., K. Heen og KG. Salvanes. 1995. The invisible resource rent in the limited entry and quota managed fisheries – the case of Norwegian purse seine fisheries. *Marine Resource Economics*, 10, 341-56.
- Færøernes Landsbank. 2006. *Personlig kommunikation med Kari Petersen*.
- Gompertz, B. 1825. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies, *Philosophical Transactions of the Royal Society of London*, 115, 513-85.
- Gordon, H. Scott. 1954. The economic theory of a common property resource: the fishery. *Journal of Political Economy* 62: 124-42.
- Hermansen, Ø. og O. Flaaten. 2004. *Government financial transfers to the fish harvesting, processing and aquaculture industries in Norway 1990-2002*, NORUT Samfunnsforskning AS. Rapport nr. SF 09/204.
- Lindebo, E., A. Hoff og N. Vestergaard. 2002. *Economic and physical measures of capacity: A comparative analysis of Danish trawlers*, paper presented at the XIth biannual Conference in the International Institute of Fisheries Economics and Trade, Wellington, New Zealand, August 2002.
- Nielsen, M., B. Cozzari, G. Eriksen, O. Flaaten, E. Gudmundsson, J. Løkkegaard, K. Petersen og S. Waldo. 2006. *Økonomien i de nordiske fiskerier – Fokus på ressourcen*, Nordisk Ministerråd, Tema Nord 2006:540.
- Organisation of Economic Cooperation and Development. 2004. *The costs of managing fisheries*, Paris.
- Organisation of Economic Cooperation and Development. 2005. *Review of fisheries in 2001-2003*, Paris.
- The International Council for Exploration of the Sea. 2003. *Report from the Advisory Committee of Fisheries Management*, available at <http://www.ices.dk/indexfla.asp>.
- United Nations. 1982. *Convention on Law of the Sea*, Rome.
- Warming, J. 1911. Om grundrente af fiskegrunde, *Nationaløkonomisk Tidsskrift*, 49, 499-505.
- Warming J. 1931. Aalegaardsretten, *Nationaløkonomisk Tidsskrift*, 69, 151-62.

Appendiks

Det samfundsøkonomiske afkast

Det samfundsøkonomiske afkast i den nuværende situation er for de enkelte fartøjsgrupper givet ved (1) og (2), før og efter fradrag af offentlige nettoudgifter til fiskeriet (henholdsvis type A og B).

$$R_0^A = \sum_{i=1}^n L_i^0 p_i + OT - VC_0 - FC_0 \quad (1)$$

$$R_0^B = \sum_{i=1}^n L_i^0 p_i + OT - VC_0 - FC_0 + U_0 - S_0 - M_0 \quad (2)$$

hvor R_0^A og R_0^B er det nuværende samfundsøkonomiske afkast henholdsvis før og efter fradrag af offentlige nettoudgifter til fiskeriet ud over andre erhverv, L_i^0 = landinger af den kvoterede fiskeart i i udgangssituationen, i = fiskeart, hvor $i=1 \dots n$, p = pris, OT = anden indtægt, VC_0 = variable omkostninger i udgangssituationen, herunder omkostninger til aflønning af arbejdskraft i alternativ anvendelse, FC_0 = faste omkostninger i udgangssituationen, herunder omkostninger til aflønning af kapital i alternativ anvendelse, U_0 = offentlige indtægter fra brugerbetaling i fiskeriet, S_0 = tilskud og M_0 = omkostninger til fiskeriforvaltning.³

Det maksimale samfundsøkonomiske afkast

Det maksimale samfundsøkonomiske afkast beregnes formelt i tre trin i en bio-økonomisk model, givet størrelsen af det samfundsøkonomiske afkast i udgangssituationen. Først identificeres i (3) – (7) totalomsætningen baseret på, at de biologiske muligheder udnyttes. Derefter findes omkostningerne i (8), og endelig beregnes det maksimale samfundsøkonomiske afkast i (9) – (11).

Totalomsætningen (T) for den enkelte fartøjsgruppe er givet ved:

$$T = \sum_{i=1}^n L_i p_i + OT \quad (3)$$

Priser beregnes i løbende prisniveau og antages konstante. Samtidig antages det, at

3. Årsagen til, at offentlige indtægter fra brugerbetaling er lagt til på højresiden i (2), er, at brugerbetaling indgår som minus i fartøjsgruppens omkostninger. Dette skal neutraliseres.

anden indtægt (OT) er konstant, herunder indtægt fra fangst af ikke-kvoterede fiskearter.⁴

Produktionsfunktionen for den enkelte fartøjsgruppes fiskeri på kvoterede bestande er givet ved den mængde, der bæredygtigt kan udtages af bestanden:

$$L_i = q_i, \text{ for } i = 1 \dots n \text{ og } H > 0 \quad (4a)$$

$$L_i = 0, \text{ for } i = 1 \dots n \text{ og } H = 0 \quad (4b)$$

hvor q_i = den mængde af fiskeart i , den givne fartøjsgruppe bæredygtigt kan udtage af bestanden årligt, forudsat at fartøjsgruppen fisker en konstant andel af bestanden. H = fiskeriindsats målt som antal havdage. Fiskeriet på de bestande, hvor der ikke eksisterer biologiske oplysninger, forudsættes at udvikle sig præcist som dem, hvor der eksisterer biologiske oplysninger.

Den biologiske produktionsfunktion er givet ved:

$$q_i = f(X_i(H_i)), \text{ for } i = 1 \dots n \quad (5)$$

hvor X_i og H_i er henholdsvis biomasse for og fiskeriindsats af en given fartøjsgruppe på fiskeart i . Den biologiske produktionsfunktion angiver således, at den mængde fisk, der bæredygtigt kan udtages af bestanden, afhænger af biomasse, som igen afhænger af fiskeriindsats. Dvs. q_i afhænger af både X_i og H_i , men da der er en indbyrdes afhængighed mellem X_i og H_i , modelleres q_i som alene afhængig af H_i .

Den biologiske produktionsfunktion kan nu udspecificeres i to. Først specificeres sammenhængen mellem q_i og X_i i (6) og derefter sammenhængen mellem X_i og H_i i (7). Sammenhængen mellem q_i og X_i kan beskrives med følgende funktion:⁵

$$q_i = aX_i e^{-(bX_i)} \quad (6)$$

Hvor a er en konstant og b en elasticitet i den biologiske produktionsfunktion. Parametrene kalibreres for bestande, hvor der eksisterer biologiske vurderinger, givet viden om den nuværende biomasse, den maksimale biologiske biomasse og den betragtede fartøjsgruppes faktiske fiskeri på de enkelte bestande. Det antages endvidere, at de analyserede fartøjsgrupper fisker en fast andel af fiskebestandene.

4. Antagelsen om at indtægter fra fangst af ikke-kvoterede fiskearter er konstant, kan være problematisk, da der kan være substitution mellem kvotebelagte og ikke-kvotebelagte bestande. Antagelsen er imidlertid af mindre betydning i nærværende artikel, da de betragtede fartøjsgruppers omsætning fra ikke-kvoterede fiskearter er beskednen.

5. Gompertz' (1825) »Law of mortality« inden for demografi udvikledes baseret på denne funktionelle form.

Sammenhængen mellem X_i og H_i specificeres ud fra, at ændringer i fartøjsgruppens fiskeriindsats anvendt til fiskeri efter fiskeart i antages at afhænge af ændringer i biomassen som angivet i (7):

$$\frac{X_{i,1}}{X_{i,0}} = \beta \left(\frac{H_{i,1}}{H_{i,0}} \right)^{-z}, \text{ for } i=1 \dots n \quad (7)$$

hvor $H_{i,0}$ og $H_{i,1}$ er antal havdage anvendt til fiskeri efter art i henholdsvis i udgangsperioden og i en ny periode, hvor $X_{i,0}$ og $X_{i,1}$ er biomassen for art i henholdsvis i udgangsperioden og i en ny periode, hvor β er en konstant, og hvor z er en elasticitet, der angiver, hvordan bestanden påvirkes af ændret fiskeriindsats. Den funktionelle form angiver, at udviklingen i biomassen bestemmes af udviklingen i fiskeriindsatsen, givet biomasse og fiskeriindsats i udgangssituationen. Dvs. ved en stor fiskeriindsats eksisterer en lille bestand og ved en lille fiskeriindsats opbygges på langt sigt en stor bestand. Som følge af at beregningen tager udgangspunkt i den nuværende biomasse og fiskeriindsats, beregnes den fremtidige fiskeriindsats under forudsætning af, at det nuværende niveau er bæredygtigt. Det antages endvidere, at $\beta = 1$ for at simplificere modellen, samt at $z = 1$, svarende til konstant skalaafkast. Hvis $z < 1$, er der faldende skalaafkast og hvis $z > 1$, er der stigende skalaafkast. Endelig antages da

$$H_1 = \sum_{i=1}^n H_{i,1} \text{ og } H_0 = \sum_{i=1}^n H_{i,0} \text{ at:}$$

$$\frac{H_{i,1}}{H_1} = \frac{H_{i,0}}{H_0}, \text{ for } i=1 \dots n \quad (8)$$

(8) indebærer, at fartøjsgruppens relative tidsforbrug på de forskellige arter i en ny situation er det samme som i udgangssituationen.

Under forudsætning af (8) kan X_i isoleres i (7) og indsættes i (6). q_i i (6) kan endvidere indsættes i (4a) givet at $H > 0$, hvorved fartøjsgruppens totalomsætning i en ny situation, hvor der er taget højde for bestandsdynamikken, fremkommer ved at indsætte L_i i (3).

Da beregningen foretages for det lange sigt, er alle omkostninger variable, og omkostningsfunktionen for den enkelte fartøjsgruppe er givet ved (9) i en ny situation under antagelse af stigende faktorpriser.

$$C_1 = A \left(VC_0 + \left(\frac{DS_0}{DS_1} \right) FC_0 \right), A = \alpha \left(\frac{H_1}{H_0} \right)^{(1+\nu)} \quad (9)$$

C_1 er de totale omkostninger med aflønning af kapital og arbejdskraft i alternativ anvendelse og FC i nærværende langsigtede model angiver de omkostninger, der knytter sig til det enkelte fartøj uanset fartøjets aktivitet, og som på kort sigt er faste. Disse inkluderer kapitalomkostninger, afskrivninger og fartøjsomkostninger, som er uafhængige af, hvor meget det enkelte fartøj fisker. Disse omkostninger er i nærværende langsigtede model variable. DS_0 er antallet af havdage pr. fartøj pr. år i den nuværende situation og DS_1 er antallet af havdage pr. fartøj pr. år i en ny situation. $\frac{DS_0}{DS_1}$ introduceres for at korrigere for kapacitetsudnyttelsen. A er en skaleringsfaktor for omkostningerne, der beskriver, hvor meget de ændrer sig, når den nuværende situation ændres til en ny. α er en konstant og $\alpha > 0$, da omkostningerne stiger med fiskeriindsatsen. α sættes for at simplificere til 1. Omkostningerne i en ny situation beregnes således med udgangspunkt i det nuværende omkostningsniveau, korrigeret for ændret fiskeriindsats i fartøjsgruppen. ν er en parameter for forskellen i fiskernes effektivitet i anvendelsen af input faktorer. Når $\nu > 0$ er nogle fiskere dygtigere end andre, hvilket medfører et økonomisk overskud (den infra-marginale rente jf. Copes 1972), hvilket indebærer, at omkostningsfunktionen bliver konveks.⁶ I nærværende artikel antages $\nu = 0,33$.⁷ Det antages, at omkostningerne pr. havdag er konstante.

Det maksimale samfundsøkonomiske afkast uden fradrag af offentlige nettoudgifter til fiskeriet (type A) er for de enkelte fartøjsgrupper givet ved:

$$R_1^A = (T_1 - C_1) \quad (10)$$

hvor T_1 er totalomsætningen i en ny situation og R_1^A = det samfundsøkonomiske afkast før fradrag af offentlige nettoudgifter til fiskeriet ud over andre erhverv i en ny situation. Det samfundsøkonomiske afkast efter fradrag af offentlige nettoudgifter til fiskeriet ud over andre erhverv (type B) beregnes ikke i en ny situation, da udviklingen i de offentlige nettoudgifter til fiskeriet ikke kendes eksakt. I det omfang reguleringssomkostningerne er større i en ny situation, vil det identificerede maksimale samfundsøkonomiske afkast være overvurderet.

6. Omkostningskurven (med aflønning af kapital og arbejdskraft i alternativ anvendelse) kan også i visse situationer få andre former end den valgte. Således kan omkostningskurven (C i figur 1) få en omvendt U -form, såfremt reguleringssomkostningerne inddrages, og disse samtidig opvejer øvrige omkostninger. Dette skal ses i sammenhæng med, at reguleringssomkostningerne normalt vil være højere ved forvaltning end ved frit fiskeri, hvor de kan være nul. I nærværende artikel identificeres omkostningsfunktionen dog uden reguleringssomkostninger, da den maksimale ressourcerente alene identificeres for type A.

7. At $\nu = 0,33$ er et skøn. Skønnet er dog baseret på en approksimation foretaget på grundlag af en undersøgelse af faktiske forskelle i effektivitet i fiskeriet efter torsk i Nordsøen, Lindebo m.fl. (2002). Fartøjers effektivitet måles ved en score baseret på »Data Envelopment Analysis«, som angiver, hvor mange produktionsfaktorer et givet fartøj bruger i forhold til det mest effektive.

I modellen bestemmes det maksimale samfundsøkonomiske afkast for de enkelte fartøjsgrupper ved en optimeringsberegning under antagelse af uændret flådestruktur og fiskerimønster, givet en antaget maksimal kapacitetsudnyttelse på 225 dage om året for små fartøjer og 300 for store. Kapacitetsudnyttelsen er skønnet på grundlag af det faktiske antal havdage, at større fartøjer kan være længere på havet end små, at den svenske regulering i dag umuliggør fuld udnyttelse af kapaciteten, at det norske fiskeri er sæsonbetonet, samt at muslinger ikke kan fiskes, hvis det er for varmt. Det maksimale samfundsøkonomiske afkast bestemmes ved (11):

$$\text{Max}_{H_1} \left(R_1^A \right), \text{ givet } DS_1 = 225, 300 \quad (11)$$

Hvilket kan omskrives ved at indsætte (10) i (11) og indsætte (9) for C_1 i (10). X_1 isoleres endvidere i (7) og indsættes i (6) under forudsætning af (8), som igen indsættes i (4) for $H > 0$. L_i indsættes i (3) for alle i og endelig indsættes T i (10). Herved fremkommer (12):

$$\text{Max}_{H_1} \left(\sum_{i=1}^n \left(a \frac{X_{i,0} H_{i,0}}{H_{i,1}} e^{-\left(b \frac{X_{i,0} H_{i,0}}{H_{i,1}} \right)} \right) p_i + OT - \left(\frac{H_1}{H_0} \right)^{(1+v)} \left(VC_0 + \left(\frac{DS_0}{DS_1} \right) FC_0 \right) \right), \quad (12)$$

givet $DS_1 = 225, 300$

(12) viser totalomsætning minus totalomkostninger med aflønning af kapital og arbejdskraft i alternativ anvendelse. I (12) er fiskeriindsatsen i en ny situation ($H_1 = \sum_{i=1}^n H_{i,1}$) eneste variable led. Denne bestemmes ved maksimering.

Fordeling af det samfundsøkonomiske afkast

Fordelingen af det samfundsøkonomiske afkast efter fradrag af offentlige nettoudgifter til fiskeriet (type B) identificeres ved at beregne andelen for kapitalen og arbejdskraften i fartøjsgruppen og den offentlige sektor, som angivet i (13) – (15).

$$w_L = \frac{(LI_F - T_F^I - LI_A + T_A^I)}{R_0^B} \quad (13)$$

$$w_K = \frac{(\pi - T_F^C - KI_A + T_A^C)}{R_0^B} \quad (14)$$

$$w_G = \frac{(T_F^I - T_A^I) + (T_F^C - T_A^C) + U + S + M}{R_0^B} \quad (15)$$

hvor w_L , w_K og w_G er henholdsvis arbejdskraftens, kapitalens og den offentlige sektors andel af det samfundsøkonomiske afkast i fartøjsgruppen, LI_F = total aflønning af arbejdskraft i fartøjsgruppen, LI_A = total aflønning af fartøjsgruppens arbejdskraft i alternativ anvendelse, π = overskud i fartøjsgruppen, KI_A = alternativ forrentning af den kapital, der er investeret i fartøjsgruppen i form af rente eller overskud, og hvor T_F^I , T_A^I , T_F^C og T_A^C er betalt skat fra henholdsvis arbejdskraften i fartøjsgruppen, fartøjsgruppens arbejdskraft i alternativ anvendelse, kapitalen investeret i fartøjsgruppen og fartøjsgruppens kapital investeret alternativt.

Does development aid help poor countries converge to our standard of living?

Tryggvi Thor Herbertsson

Askar Capital, Reykjavik, Iceland, E-mail: tthh@askar.is

Martin Paldam

School of Economics and Management, University of Aarhus, E-mail: mpaldam@econ.au.dk

SUMMARY: Aid flows are included in the standard convergence equation and estimated using cross-country, panel and GMM regression. Robustness of the result is tested by changing the model and by adding extra variables. The main results are: that absolute convergence and absolute aid effectiveness are both rejected, and conditional convergence is accepted. Aid has an activity effect in the short run, but conditional aid effectiveness is found to be dubious. Finally, we try to divide the countries into an A-group where aid is effective and a B-group where it harms. Several criteria for division are explored, but none are really successful – the most satisfactory is the one that divides countries according to income.

1. Introduction

Poverty in the less developed countries (LDCs) causes huge losses of welfare in the world. Thus, enormous welfare gains would result if poor countries were to converge to our standard of living. Many believe that development aid is an effective means of generating this convergence, see e.g. Sachs (2005). Since aid programs started in the mid 1960s, the average LDC has received about 7.5% of its aggregate GDP in aid per year.¹ Over 40 years this corresponds to three years of GDP, so it is substantial.

We thank Marias Gestsson for excellent research assistance. We also wish to thank Pia Wichmann Christensen, Chris Doucouliagos and Peter Sandholt Jensen for useful discussions. This paper has had a long gestation period, and we almost gave up due to the results, and when we finally finished the first version, Rajan and Subramanian (2005) came out with similar results. Also, Martin Paldam wants to thank the four students, who over the years have written their MA-thesis in the field. We refer to Kristiansen (2007) for unit-root and causality tests.

1. Aid is heavily biased toward small countries, and about 30% of the data are missing. Countries with no data probably receive below average, so aid to the aggregate LDC world is only about 2.5%.

However, cross-country growth rates are neither correlated to income levels nor to aid shares. The data reject both absolute effects: *There is neither absolute convergence nor absolute aid effectiveness*:² Poor countries do not converge, and aid flows seem to have no effect. Many researchers have found the two basic zero-effect results counterintuitive, and this has led to two bodies of empirical cross-country research, the convergence literature, CL, and the aid effectiveness literature, AEL. In both fields, the researchers have strived to demonstrate that by imposing more structure on the seemingly unrelated data, it is possible to make them tell a different – *conditional* – story. The two literatures use models that are curiously parallel, but they rarely refer to each other. The CL in particular rarely seems to refer to the AEL.

The main extra structure imposed on the basic models is to add sets of control variables. Obviously, it is possible to find control sets giving a wide range of results for both the convergence and the effectiveness effect. So, the control sets should be *reasonable*. The most reasonable sets aim at controlling for *country heterogeneity*. Further, the data may be divided into subsets with different explanations.

The two zero-correlation results have led to opposing policy conclusions: The lack of convergence has caused many to propose an increase in development aid, and the aid ineffectiveness result has caused many to doubt that aid works. So it is important to confront the results. This is precisely what we do in this paper, which took off from the idea that there might have been divergence without aid. Consequently, we start our quest from the idea that an answer to the question in the title could be found by merging the CL and the AEL models.

Table 1 introduces the models that define the terms as they are used in this article:³ Equations (1) defines *absolute* convergence and aid effectiveness, while equations (2) to (4) are *conditional* models.

The many models used are normally presented as being derived from *the* economic theory but they may also be a result of researchers mining the data. This is not problematic in the large CL, where the basic model has a clear link to neoclassical growth theory, and has been analyzed by researchers with many different priors. However, the AEL (of about 100 papers) is smaller, and the theory is less well-established. Furthermore, the AEL is affected by a major prior: The profession is reluctant to publish

2. Both literatures work with unweighted country observations. If countries are weighted by population size, China and India come to dominate. The two giants are still poor, and they have both high growth and small aid shares. A weighting thus causes absolute convergence and negative aid effectiveness.

3. The terminology used is in accordance with the one used in the CL, see e.g. Barro and Sala-i-Martin (2004). Unfortunately, the AEL uses the word *conditional* in a more restricted sense. Consequently, we use the term *conditional aid effectiveness* models for a broader set of models than the ones covered by the AEL termed *conditional models*. They are discussed in Section 2.5

Table 1. The equations analyzed in the paper.

Separate, but parallel equations	CL: Growth equation	AEL: Aid effectiveness equation
<i>Absolute</i> : Basic equation	$g_{it} = \alpha + \beta y_{it} + u_{it}$ (1a)	$g_{it} = \alpha + \mu h_{it} + u_{it}$ (1b)
<i>Conditional</i> : Controls	$g_{it} = \alpha + \beta y_{it} + \gamma_j x_{jit} + u_{it}$ (2a)*	$g_{it} = \alpha + \mu h_{it} + \gamma_j x_{jit} + u_{it}$ (2b)*
<i>Conditional</i> : Fixed effects	$g_{it} = \alpha_{it} + \beta y_{it} + u_{it}$ (3a)	$g_{it} = \alpha_{it} + \mu h_{it} + u_{it}$ (3b)
(1) or (3) with lagged income	$g_{it} = \alpha + \beta y_{it} + \delta g_{it-1} + u_{it}$ (4a)	$g_{it} = \alpha + \mu h_{it} + \delta g_{it-1} + u_{it}$ (4b)
Merged equation		
<i>Absolute</i> : Basic equation	$g_{it} = \alpha + \beta y_{it} + \mu h_{it} + u_{it}$ (1c)	
<i>Conditional</i> : Fixed effects	$g_{it} = \alpha_{it} + \beta y_{it} + \mu h_{it} + u_{it}$ (3c)	
<i>Conditional</i> : Fixed effects and one control	$g_{it} = \alpha_{it} + \beta y_{it} + \mu h_{it} + \gamma_j x_{jit-1} + u_{it}$ (3d)	
(1) of (3) with lagged income	$g_{it} = \alpha + \beta y_{it} + \mu h_{it} + \delta g_{it-1} + u_{it}$ (4c)	
Variables, coefficients, indices		
g_{it} Growth real per capita growth for $T = 5$ years	α, α_{it} Constant, fixed effects for countries and time	
y_{it} Income, ln gdp, i.e. GDP per capita, PPP prices	β Estimated coefficient of convergence	
h_{it} Aid share, ODA, in % of GDP	μ Estimated effect of aid	
x_{jit} vector of j controls, one control is not bolded	δ, γ_j coefficients to be estimated	
u_{it} Residuals	i, t, j Indices for countries, time and controls	

Note: The models with * are not estimated in the present paper. Equation (3d) is used in Section 5. Versions of (3) and (4) are estimated with aid unlagged and with aid lagged, the latter is termed (3bL), (4cL) etc.

negative results on aid effectiveness, as demonstrated by Doucouliagos and Paldam's meta analysis (2007a).⁴

While we use the standard data, ours is the first paper to systematically analyze the CL and AEL models in a parallel way to study what happens when they are merged. It should, however, be noted that we do not examine or test the leading pro-aid models. This has been done in Jensen and Paldam (2006) and Doucouliagos and Paldam (2006b).

Section 2 surveys the theoretical framework and main findings of the two literatures. Section 3 presents the data and descriptive statistics. Section 4 systematically analyzes the cross-country relations, using OLS regression techniques, while Section 5 presents panel data estimates and tests the robustness by inclusion of extra variables. Section 6 tries to divide the countries into an A-group where aid helps and a B-group where aid harms. Section 7 discusses the results, and finally Section 8 summarizes the results.

4. The study considers the distribution of *all* published estimates of aid effectiveness. Meta analysis has developed a test (FAT) to analyze such distributions for asymmetries. These tests detect a significant asymmetry looking as predicted by the reluctance hypothesis.

2. The CL and the AEL: Two bodies of literature

The models and variables discussed in the paper are listed in Table 1 for easy reference. The two sets of equations have great formal similarity. By far the largest literature is the CL. Here some agreement has been reached about the basic facts, so we are brief in our survey. In the AEL, little agreement has been reached. However, before we turn to the surveys, a general problem should be mentioned.

2.1. *The mining observation*

The AEL is the smaller literature of the two. The aid data starts in the mid 1960s, when aid started, and now amount to about 5000 annual observations. The models are estimated on data averaged over periods of four to ten years. This reduces the data to between 500 and 1200 observations. On subsets of these data 1025 regressions have been published, but many more have surely been run. Consequently, these data have been thoroughly mined.

The CL can use data as far back as they are available, but much more research has been done, so the mining done is probably even larger.

Test limits in econometrics are, by convention, the ones of *one* analysis run on virgin data.⁵ Mining reduces the risk of Type I errors (rejection of true model) and increases the risk of Type II errors (acceptance of false models). Two conclusions follow: Results that are not clearly visible in the raw data – or follow from basic models – should only be believed after independent replication; that is, by other authors on new data. Results hinging upon controls that are not strongly justified should be treated with suspicion.

2.2. *CL: Absolute convergence rejected. Conditional models find convergence*

Recent surveys of the CL can be found in Barro and Sala-i-Martin (2004, pp 1-84 and 511-66) and Aghion and Durlauf (2005, especially chapters 1, 7, 8 and 9).

One of the key theories of economics is the neo-classical growth model. Under rather general assumptions, it shows that economies accumulate capital so as to reach the same steady state income per capita. A robust (theoretical) prediction from the model is that eventually all countries will converge to the same level of income.⁶ The starting point for the convergence literature is the question: Does convergence actually happen?

Data for y and g cover many countries and five to twenty decades, and as mentioned they are basically uncorrelated. However, perhaps the lack of convergence is due to

5. Formulas have been proposed in the econometric literature adjusting the significance limits of the tests to the degrees of freedom left. They have not been applied in the EAL, and to our knowledge, neither in the CL.

6. Two mechanisms secure the catch-up: (1) the logic of the model itself, i.e., diminishing returns to the two inputs (labor and capital), (2) technological catch-up.

country differences, so that convergence occurs if country heterogeneity is controlled for. This point was made in the classical study of cross-country growth patterns by Barro (1991) as extended in Barro and Sala-i-Martin (1995, 2004).⁷ It uses a two step model:

Step 1 analyzes absolute convergence by equation (1a) $g_{it} = \alpha + \beta y_{it} + u_{it}$, which corresponds to the correlation between g and y . When it is estimated for rich countries only, see Baumol (1986), convergence does occur, but when the LDCs are included absolute convergence is rejected as $\beta \approx 0$.

Step 2 uses the two extended models (2a) or (3a) to control for country heterogeneity. Conditional convergence means that $\beta < 0$ in the extended model.

Version (2a) $g_{it} = \alpha + \beta y_{it} + \gamma_j' x_{jit} + u_{it}$. The j controls are country levels of, e.g., education, health, investment, governance, resources. A fairly broad range of the most credible x -sets turn β negative. It appears optimistic that 5-10 controls can account for all country differences, but it is appealing that the controls are concrete. However, each variable of the x -set normally has missing observations, so with a large x -set only a subset of the (g, y) -data can be used. Also, the choice of x -set gives the researcher considerable control over the estimate of β , causing a problem of moral hazard, as it allows him to tailor results to his priors.⁸ However, comprehensive robustness tests have been made (by Doppelhofer, Miller and Sala-i-Martin, 2004, and Sturm and Haan, 2005) showing that a little more than 10 variables do have a robust effect on growth, while another 5 to 10 are borderline robust.⁹

Version (3a) $g_{it} = \alpha_i + \beta y_{it} + u_{it}$. The constant is decomposed into a set of *fixed effects* for countries, α_i .¹⁰ This cannot be done in cross-country regressions, so panel techniques are used. Fixed effects for countries assume that country heterogeneity is constant for the period analyzed, so the models can be estimated for all data where observations for g and y are available. Fixed effects turn β negative.

Consequently, in the CL (2a) and (3a) tell almost the same story, though β is normally a little larger numerically in (3a) than in (2a).

2.3. AEL: Absolute aid effectiveness rejected. Conditional models disagree¹¹

The AEL is covered by a number of surveys too; the two most recent are McGillivray

7. Several alternative methods are used to study cross-country growth patterns, see Aghion and Durlauf (2005).

8. These models are estimated by two- or three-stage estimators to control for reverse causality. Hereby a set of first stage instruments is also included, so the choices of the researcher become even larger.

9. Aid has never made it to either list, nor is aid mentioned in Aghion and Durlauf's two volume survey (2005) covering the CL and everything else known about economic growth.

10. Also, a set of fixed effects for time periods α_t is often included to delete international economic fluctuations.

11. The aid effectiveness discussion is less known. It started with theoretical papers by Friedman (1958) and Bauer (1971), who argued from an (explicit) libertarian position that aid goes to governments, and thus

et al. (2006) and Doucouliagos and Paldam (2006b, 2007a).¹² Their conclusions are almost opposite:

McGillivray et al. is a qualitative study with assessed results. The key conclusion is that an upward kink appears in the results in the late 1990s. Before that it was unclear whether aid works, but now we know that it does. Doucouliagos and Paldam's surveys are quantitative (meta) studies which systematically compare all results. They show a steadily falling trend in the estimate for aid effectiveness, with no signs of a structural break. The AEL is a field where even the surveys disagree.

Step 1 considers absolute aid effectiveness, using the basic model (1b) $g_{it} = \alpha + \mu h_{it} + u_{it}$, where aid effectiveness means that $\mu > 0$. In most large data sets – see Tables 4 and 5 – absolute aid effectiveness is rejected as $\mu \approx 0$. This is uncontroversial. Step 2 is the two extended models (2b) and (3b) controlling for country heterogeneity:

Version (2b) $g_{it} = \alpha + \mu h_{it} + \gamma_j x_{jit} + u_{it}$. The j controls of the x -set are, once again, meant to control for country heterogeneity. In the results published $\mu > 0$, $\mu \approx 0$ and $\mu < 0$ in 38%, 56%, and 6% respectively.¹³ However, as mentioned the distribution is clearly downward censored by the reluctance bias, and no agreement has been reached as regards the right control set. Here it is worrying that the standard control sets that generate conditional convergence fail to generate conditional aid effectiveness.

Version (3b) $g_{it} = \alpha_i + \mu h_{it} + u_{it}$ uses fixed effects for countries. It is rarely used in the papers published, which is strange since fixed effects convert the cross-country data (in a basic way) into time series, precisely as demanded by the key policy question: What happens to development in the typical country if aid is either increased or decreased?

In choosing between version (2a) and (3a), we prefer the latter as it has six advantages: (1) Higher policy relevance; (2) simplicity; (3) it is void of moral hazard; (4) dummies are always available so the full (g_{it}, y_{it}) -set can be analyzed; (5) the dummies are truly exogenous; and finally (6) it leaves the variables from the standard x -set for robustness tests, which we carry out in Section 5.2.

Many aggregate time series contain cyclical components. When the series are aggregated to 5 years, the cyclical component may appear as residual autocorrelation, as

continued ...

encourages countries to pursue unsound socialist policies. The AEL, see Doucouliagos and Paldam (2006a), confirms that about 75% of the marginal activity financed by aid is public consumption. However, data for the degree of public ownership to trade and industry are not strongly correlated to the share of public consumption. The empirical AEL started with a critique of aid by Griffin (1970) and Weisskopf (1972), who argued from an (explicit) left wing position, and demonstrated that aid did not cause increased capital accumulation. Since then aid data have multiplied and almost 100 papers have been published.

12. Christensen, Doucouliagos and Paldam (2007a) is a master list of the AEL, while Christensen, Doucouliagos and Paldam (2007b) lists the even larger literature dealing with aid allocation, i.e. with the reverse causality.

13. The percentage numbers in the brackets are from Doucouliagos and Paldam (2007a).

found below. Therefore we include lagged growth, δg_{it-1} , in models (4a) and (4a). Sections 4 and 5 show that the term is often significant, but does not have much effect on the other coefficients estimated.

2.4. Third generation AEL-studies: The division hypothesis of an A- and a B-group

Another way to consider the absolute aid ineffectiveness result is to adopt the *division hypothesis*: where the countries are divided into an A-group of countries where aid works and a B-group where it harms.

This also happens if you toss a coin: In half the cases you win, and in the other half you loose. You can always find a division criterion ex post that divides a series of tosses into an A-group of heads and a B-group of tails. However, if the criterion is subjected to independent replication, it may fail. Thus, a credible division condition, z , is needed, and (as usual) it is only believable after independent replication.

A neat formalization results if z is scaled to have approximately as many positive as negative values. Then the A-group occurs for $z > 0$, and the B-group occurs for $z < 0$ and aid effectiveness can be estimated by one regression (5a).

$$g_{it} = \alpha + \sigma z_{it} + \mu h_{it} + \omega z_{it} h_{it} + u_{it} \quad (5a)$$

$$g_{it} = \alpha + \mu h_{it} + \omega h_{it}^2 + u_{it} \quad (5b)$$

The model (5b) is slightly different as $z = h$ is used as the interaction term.

By far the most influential conditional model is Burnside and Dollar's *Good Policy Model* (2000),¹⁴ where z is an index of good policy, which is scaled to divide the countries in two groups as mentioned. The original estimate was $\sigma \approx 1$, $\mu \approx 0$ and $\omega > 0$. The good policy index is almost an index of outcomes, so it is trivial that $\sigma > 0$, and $\mu \approx 0$ is not surprising either. However, it is important that $\omega > 1$, as it means that aid increases growth in the A-group of countries with good policies, while it decreases growth in the B-group of countries with bad policies. This result has appeared credible to many development practitioners, and it was popularized in World Bank (1998), *The Economist*, etc.¹⁵

The second most studied conditionality model is the *Medicine Model* (5b), where aid is interacted with itself. This idea has been pursued by several authors – most vi-

14. Burnside and Dollar use the EDA aid data from Chang, Fernandez-Arias and Serven (1998), where each aid loan/gift is cleaned for non-grant elements. The results are seemingly invariant to this refinement as the standard ODA and the new EDA data have a coefficient of correlation of 0.83.

15. It was also advocated in Paldam (1997a) and in several popular articles by the author till he attempted to replicate the model on new data; see Jensen and Paldam (2006), which also analyzes the Medicine Model.

gorously by Hansen and Tarp (2000). Their findings are that $\mu > 0$ and $\omega < 0$,¹⁶ so that while some aid increases growth, too much aid is harmful, just like medicine. Thus, the model predicts that an optimal dose of aid, h^* , exists. Tables 6a and 6b below include aid squared, with no effect.¹⁷

Both models can be extended with controls or fixed effects. As at the start of 2005 no less than 31 papers have proposed 10 conditioning variables.¹⁸ The pace of publications during the last five years predicts that many more studies in this family may appear during the next five years. For reasons given above, the key criterion for the credibility of an effect is that it has been independently verified by other researchers using new data. This has only been done for the two models mentioned. Here the replication failed, so none of the proposed conditioning factors have, as of now, been established.

2.5. Three types of possible biases in the AEL

In estimates of cross-country equations between macro variables, it is difficult to single out an equation where the explanatory variables are fully independent and exogenous. It is thus always possible to point to possible biases – this is also the case at present. We shall discuss the possible biases under three headings:

(B1) **Reverse causality:** Many studies adjust the estimates for reverse causality (by TSLS or GMM estimates), but this appears to have little effect on the results (see Doucouliagos and Paldam 2007a).¹⁹ The reason becomes apparent when we turn to the aid allocation literature. A priori the effect of growth on aid is unclear: (i) It is negative if aid is given to alleviate crisis. (ii) It is positive if aid is used to finance projects with high benefit/cost ratios, as countries with high growth generate more such projects. A total of 30 studies analyze the effect of recipient growth on aid allocation.²⁰ The 211 estimates published are generally small, but on average positive so (ii) dominates – see

16. In a model where aid has no effect, the addition of aid squared will – due to the correlation of aid and aid squared – cause the coefficients to aid and aid squared to move in the opposite direction, so that one becomes positive and the other negative. With a bit of luck and the right choice of control variables, both may turn significant. However, it appears to be equally easy to get either of the two sign combinations.

17. The two models are discussed and tested in considerable detail in Jensen and Paldam (2006) and in Doucouliagos and Paldam (2006c). Therefore, we treat the two models briefly at present.

18. Doucouliagos and Paldam (2006b) is a meta-study of the 31 papers.

19. In discussions of aid effectiveness, it is surprisingly common to meet the classical mistake of confusing flows (growth rates) and stocks (income levels). It is wrong to argue that a reverse causality bias exists in the growth-aid relation because we know that aid is negatively correlated to the income level. However, it is a good reason to control the growth-aid relation for the income level, as done below. Figure 1 shows how the relation between growth and aid goes to zero as the period, T , becomes small and increasingly negative as T rises.

20. See Doucouliagos and Paldam (2007c) on the effect of growth aid and Doucouliagos and Paldam (2007d) on the effect of country size and income on aid.

however Figure 2. To deal with the potential bias, we lag the aid variable relative to the g and y variables in Section 5, and we also control some of the estimates with GMM.

(B2) Not distinguishing between *activity and capacity effects*. We want aid to cause development, that is, growth. We know that aid leads to public spending (see note 11), so consequently aid must lead to economic activity in the short run. Activity has multiplier effects, so we can count on aid to cause an increase in y in the same year, and for perhaps two more years – this will appear as a short-run positive correlation between aid and growth as shown in Section 3. This is not a growth effect. It would appear even in the proverbial case where the aid is used to finance the digging of holes and filling them up again. It appears that some of the positive aid-to-growth effects reported in the AEL are activity effects.

(B3) *Omitted variable biases*: Many have been proposed, but only a couple have been confirmed. Two of these are: (B3.1) The big country bias: Large LDCs tend to grow faster than small ones and to receive less aid. (B3.2) The scaring-away bias: Bad rulers in particular tend to scare away both aid and investment, see Levy (1988). We have decided to disregard these biases, assuming that they cancel each other out. However, we control for the Gastil index for democracy in Section 5.

Our conclusions from the AEL review are: (R1) Absolute aid effectiveness is rejected. (R2) Weak and non-conclusive conditional relations from aid to growth have been found. (R3) Aid may work differently in different countries. That is, we may be able to identify an A-group of countries, where aid works, and a B-group, where it rather harms.

2.6. Merging the models – what may happen?

When the two equations (1a) and (1b) are merged, they become: (1c) $g_{it} = \alpha + \beta y_{it} + \mu h_{it} + u_{it}$, as shown in Table 1. The same happens to equations (2a) and (2b), etc. In each set of equations, ($\neq a$) and ($\neq b$) are *unmerged*, while equation ($\neq c$) is the *merged* version. We thus get an unmerged estimate of β and μ and a merged set of estimates of the same two variables.

The main purpose of our study is to systematically examine what happens to the two coefficients, β and μ , when the merged equation is estimated. Table 2 lists the five possibilities. One possible result is that aid and convergence is basically independent, but as aid shares are correlated with poverty, it is only possible if aid effectiveness is zero, $\mu \approx 0$. The other possibilities show that aid and convergence are interdependent. When we started, we hoped for possibility 1, where aid is effective and prevents divergence. As the reader will see, the results are closer to possibility 4.

3. Data and some descriptive statistics

This section describes the data and takes a first look at some descriptive statistics,

Table 2. The five possible effects of merging the two basic equations.

Possibility 1 Interpretation	Merger causes $\beta\downarrow$ and $\mu\uparrow$ Aid works. Less convergence with no aid
Possibility 2 Interpretation	Merger causes $\beta\uparrow$ and $\mu\uparrow$ Aid works. More convergence with no aid. Inconsistent
Possibility 3 Interpretation	Merger causes $\beta\uparrow$ and $\mu\downarrow$ Aid harms. More convergence with no aid
Possibility 4 Interpretation	Merger causes $\beta\downarrow$ and $\mu\downarrow$ Aid harms. More divergence with no aid. Inconsistent
Possibility 5 Interpretation	Merger has no effect on β and μ Convergence is independent of aid. Consistent if $\mu \approx 0$

Note: Two inconsistent possibilities are shaded in grey.

showing the structure of correlations. PWT is Penn World Tables and WDI is World Development Indicators. See the reference list for the web addresses of all sources.

3.1. Data: Three main time series and six control variables

Growth rate for GDP per capita, $g = (\ln(GDP_t) - \ln(GDP_{t-T}))/T$, where $T =$ averages over 5 years. In fixed domestic prices that reflect the trade-offs agents actually face. Source WDI.

Initial gdp, $y =$ logarithm of RGDPCH (real GDP per capita in 1985 international prices, chain index). Source PWT.²¹

Aid share, $h =$ ODA/GDP, both in current US \$. ODA (Official development assistance) is net disbursements of loans and grants made by official agencies of the members of DAC and some Arab countries to promote economic development and welfare in recipient economies listed as developing by DAC. Only loans with a grant element of more than 25% are included. ODA also includes technical cooperation and assistance. Source WDI.²²

Investment share: This is an average of real gross domestic investment, private and public, in proportion to GDP. Source PWT.

Inflation: Defined as $\text{infl}/(1+\text{infl})$ where infl is the average of the log difference of the GDP deflator. This transformation reflects the magnitude of the inflation distortion

21. See Nuxoll (1994) on the combination of WDI-data in local prices and Penn World Tables in PPP prices.

22. ODA refers to aid flows from official donors to LDCs and the transition economies of Eastern Europe and the former Soviet Union as well as to certain advanced developing countries and territories as determined by DAC. Official aid is provided under terms and conditions similar to those for ODA.

Table 3. Averages in 1971-2000 based on 72 aid recipients where full data set exists.

Three fractiles	Sorted according to growth		Sorted according to aid		
	Growth	Aid share	Three fractiles	Aid share	Growth
Slowest growth	-1.7%	9.9%	Highest aid	14.3%	-0.4%
Middle group	0.8%	5.9%	Middle group	4.7%	1.1%
Fastest growth	3.5%	3.9%	Lowest aid	0.7%	1.8%

Note: The averages are unweighted. Guinea-Bissau and Sao Tome and Principe are excluded. They have had moderate growth, but an average aid inflow of no less than 48% and 62% respectively.

in production, see Herbertsson (1999) and, equivalently, the implicit inflation tax rate. It attempts to capture the nonlinear relationship between growth and inflation: Growth is thus less sensitive to an increase in inflation from 500 to 600% per year than, say, an increase from 2 to 100% per year. The deflator is derived by dividing current price estimates of GDP at purchaser values (market prices) by constant price estimates. Source WDI.

External debt: The average of foreign debt divided by GDP at market prices. Foreign debt consists of the outstanding stock or recognized direct liabilities of the government to the rest of the world, generated in the past and scheduled to be extinguished by government operations in the future or to continue as perpetual debt. Source WDI.

Openness: The average of the sum of exports and imports of goods and services divided by GDP. Exports (imports) of goods and services represent the value of merchandise exports (imports) plus amounts receivable from (payable to) nonresidents for the provision of nonfactor services to residents. Nonfactor services include transportation travel, insurance, and other nonfactor services such as government transactions and various fees. Source PWT.

Economic freedom: The logarithm of economic freedom. Source Fraser Institute.

Political freedom: The Gastil index of political freedom. Source Freedom House.

3.2. Some correlations

Table 3 shows a negative connection between aid and growth. This can be interpreted in the two ways the data are sorted: (i) The countries with the slowest growth receive most aid; or (ii) the more aid countries receive the slower they grow. If argument (ii) is true, then aid is harmful. The top line in the table considers (72/3 =) 24 countries with average aid shares of 14.3% and an average growth of -0.4%, while the bottom line shows that the 24 countries that receive 0.7% in aid grow by 1.8%, no less than 2.2% faster per year.

Figure 1 gives a systematic presentation of the coefficients of correlations calculated for the periods $T = 5, 10, \dots, 45$ years. For $T = 5$ each of the two points are the

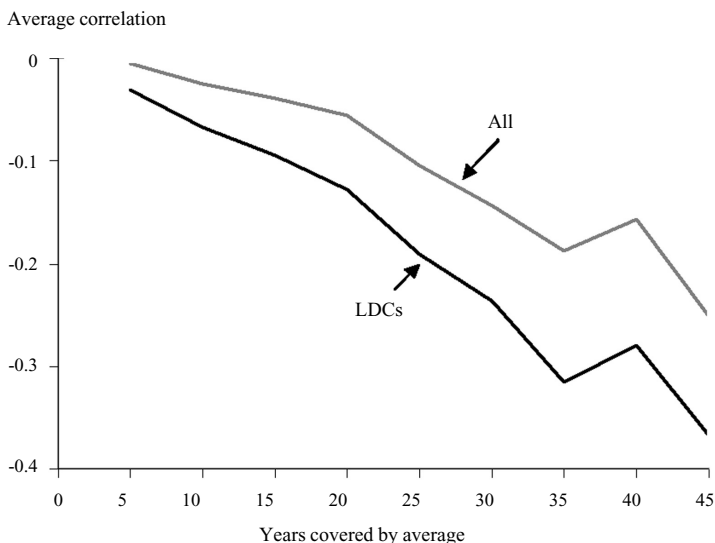


Figure 1. Correlations between average growth and the average aid share.

average of 9 correlations, where the first is for all observations with averages from 1960-65, the second is for all observations from 1965-70, etc. For $T = 10$, the points are the averages of the 8 correlations, where the first is from 1960-70, the next is from 1965-75, etc. This continues until $T = 45$ which contains only one correlation from 1960-2005.

For small T 's, the correlation is zero, but as T goes up, the correlation turns more and more negative as it increasingly come to reflect the relation between the aid share and the level of incomes.

3.3. Correlograms: A first look at the dynamics

Figure 2 shows average correlograms for the annual data for the same 72 countries analyzed in Table 3. For the data of country i , we have calculated the correlogram $c_{ij} = \text{corr}(h_{it}, g_{it} + j)$, where $j = -14, -9, \dots, +14$. This generated 72 correlograms, but we only present (2×3) averages, corresponding to the (2×3) data sets of Table 3. Each point is calculated from between 1'152 and 2'160 observations, with less at the two ends and most at the center. The correlations move by 0.2 to 0.3 around the zero axis, so even when most of the averages are in the range from -0.1 to +0.1, and of dubious significance, there are nevertheless some movements in the curves.²³

23. If data were independent between countries, we could use the significance limits of about ± 0.12 . However, some cross-country correlation exists, and the true significance levels are higher. They are hard to calculate, but they are likely to be between ± 0.15 ($df = 400$) and ± 0.20 ($df = 100$).

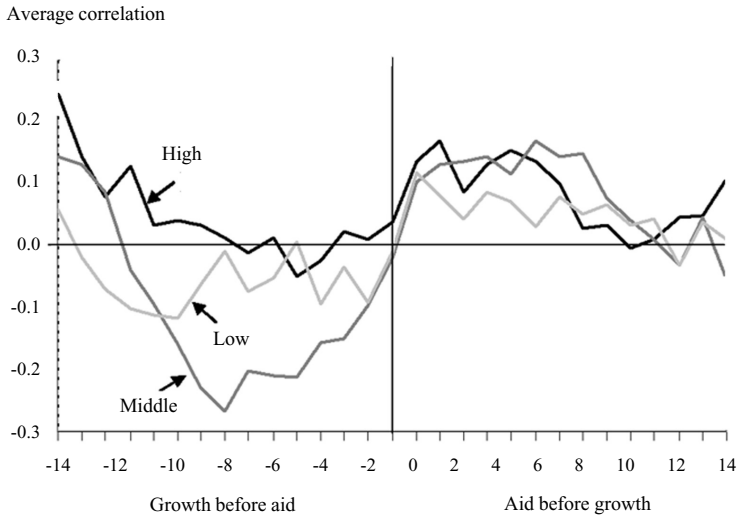


Figure 2a. Aid-growth correlograms for the 72 countries of Table 3, sorted by aid share.

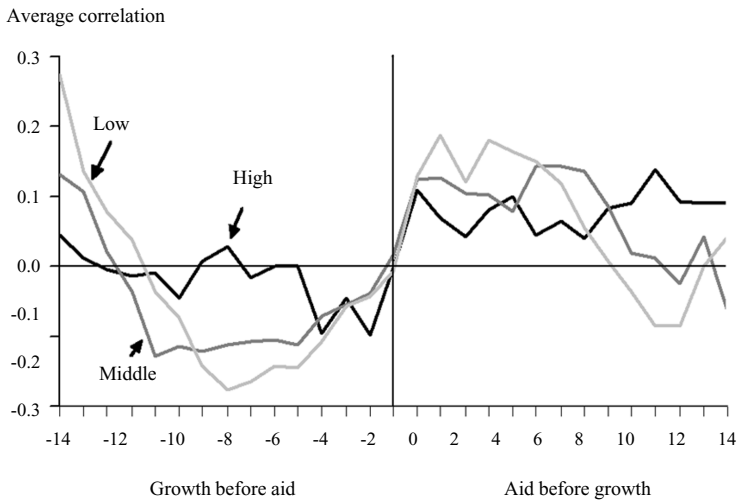


Figure 2b. Aid-growth correlograms for the 72 countries of Table 3, sorted by growth.

It is important to note that the figures provide no information about the catch-up of poor countries. They only illustrate the dynamics, and hence the causality, between aid and growth within the average country of each group. It is hence a starting point when

we address two biases: (B1) the reverse causality bias and (B2) the activity effect that should be separated from the growth effect.

For $j = 0$, we look at the simultaneous relationship between the two variables – it is almost zero. However, to both sides there are clear signs of some causal connection.

To the left is the *growth-before-aid* part of the graphs. This part analyzes the causal link from growth to aid. Here we see – on most of the 6 curves – a *negative hump* for $j = -10$ to -2 , so an economic crisis may give a little more aid.

To the right is the *aid-before-growth* part of the graphs. This part analyzes the causal link from aid to growth. For $j = 10$ to 14 , we see how aid in year t is correlated with growth 10 to 14 years later. With such a long lag, there should be very little to see, as is indeed the case. However, for $j = 1$ to 8 , we see the effect of aid in the form of a *positive hump* on curves. The positive humps are in the order of 0.05 to 0.15, so the growth hump due to aid is small. Also, Figure 2a shows that both humps are largest for countries receiving the least aid, and smallest for the countries receiving most aid. In Figure 2b, the hump is lowest for the countries with the lowest growth – countries that already receive high aid and have low growth benefit little from additional aid.

This interpretation of the two figures is problematic for two reasons. First, the largest positive growth effect (hump) is found in the countries receiving less than 1% in aid. This appears to be unreasonable. Second, it is well known that correlograms are sensitive to cyclical movements in the series. That is, some of the systematic movements may be cyclical movements in the growth rate and the level of aid. In other words, the positive humps on the right hand side of the two figures may be an upturn automatically following the downturn shown on left hand side.

An example demonstrates this problem: Imagine that a poor country has a civil war lasting some years, causing negative growth, and aid is stopped. Then peace returns, a lot of aid comes in, and there is also a resumption of normal activity, which results in high growth.

This will generate a cyclical pattern in both variables that makes it look as if aid is a much more powerful variable than it actually is. Much the same story can be told of droughts.

The possibility of cyclicity explains why we include the term δg_{it-1} in equation (4) of Table 1. With lags of five years between growth and past growth, cyclicity should cause a negative coefficient to past growth. We expect some multicollinearity between the effect of aid and the effect of past growth if past growth produces negative coefficients, but here the results are somewhat mixed.

3.4. On the property of the data and causality tests

Before we start on the regressions we should mention that many authors have discussed the statistical properties of the data and models we use.

Table 4a. Explaining growth by aid and income level, all countries Stacked OLS regressions with fixed effects for periods, $N = 970$.

Model	Income, β	Aid, μ	Aid lagged, μ	Growth lagged, δ	AR ²	MAR ² (aid)
1a	0.398 (4.1)				0.332	[0.011] ^(a)
1b		-0.047 (-4.2)			0.332	0.011
1b lag			-0.019 (-1.6)		0.322	0.001
1b both		-0.103 (-5.1)	0.070 (3.3)		0.339	0.018
1c	0.258 (2.3)	-0.031 (-2.4)			0.335	0.003
1c lag	0.424 (3.8)		0.006 (0.5)		0.332	-0.000
1c both	0.312 (2.8)	-0.090 (-4.4)	0.077 (3.7)		0.344	0.012
4a	0.056 (2.6)			0.184 (6.3)	0.358	[0.004] ^(a)
4b		-0.038 (-3.4)		0.189 (6.6)	0.361	0.007
4b lag			-0.015 (-1.3)	0.200 (7.0)	0.354	-0.000
4b both		-0.083 (-4.2)	0.056 (2.7)	0.181 (6.3)	0.365	0.011
4c	0.118 (1.0)	-0.031 (-2.4)		0.183 (6.3)	0.361	0.003
4c lag	0.256 (2.3)		-0.000 (-0.0)	0.184 (6.3)	0.357	-0.001
4c both	0.169 (1.5)	-0.077 (-3.8)	0.061 (2.9)	0.172 (5.9)	0.366	0.008

Note: AR² is the R² adjusted for degrees of freedom. AR² = 0.321 for the model with fixed effects for the 8 periods only. The MAR²(aid) is the marginal AR² due to the one or two aid variables as appropriate. The average MAR²(aid) is 0.006. ^(a) The MAR² for income.

Table 4b. Same as Table 4a but for LDCs only, $N = 749$.

Model	Income, β	Aid, μ	Aid lagged, μ	Growth lagged, δ	AR ²	MAR ² (aid)
1a	0.222 (4.9)				0.306	[0.022]
1b		-0.029 (-2.4)			0.289	0.005
1b lag			-0.004 (-0.3)		0.284	-0.001
1b both		-0.088 (-4.0)	0.075 (3.2)		0.298	0.013
1c	0.266 (5.7)	-0.046 (-3.8)			0.318	0.012
1c lag	0.235 (5.1)		-0.017 (-1.3)		0.307	0.001
1c both	0.266 (5.8)	-0.106 (-4.8)	0.075 (3.3)		0.327	0.021
4a	0.177 (3.9)			0.178 (5.5)	0.332	[0.013]
4b		-0.026 (-2.2)		0.199 (6.1)	0.323	0.004
4b lag			-0.006 (-0.5)	0.202 (6.2)	0.319	-0.001
4b both		-0.073 (-3.3)	0.059 (2.5)	0.189 (5.8)	0.328	0.008
4c	0.218 (4.7)	-0.041 (-3.4)		0.169 (5.2)	0.341	0.009
4c lag	0.190 (4.1)		-0.017 (-1.3)	0.178 (5.5)	0.333	0.001
4c both	0.222 (4.8)	-0.090 (-4.1)	0.062 (2.7)	0.158 (4.8)	0.347	0.015

Note: See Table 4a. AR² = 0.285 for the model with fixed effects for the 8 periods only. MAR²(lagged income) is 0.034. The average MAR²(aid) is 0.007.

It is all quite bulky to present, but fortunately Kristiansen (2007) contains careful tests on virtually the same data used in this paper.

We have found no indication that it is problematic to estimate the models we present from the diagnostic tests – as the matrices can be inverted, there is no singularity, and hence no unit roots.²⁴ Also, Kristiansen (2007; 50-54) contains a set of unit root tests that indicate that there are no problems. This is probably because we use (non-overlapping) 5-year time periods, which reduce the autocorrelation in the series.

The data permits us to run Granger causality tests, Kristiansen (2007; 54-58). They show precisely what could be expected from Figure 2: With a short time period ($T = 1$ to 4), there is causality both ways, but with increasing lags both causalities quickly vanish.

4. Cross-country estimates: The pattern of coefficient changes

In this section, we look for the contribution of aid to catch-up in a standard cross-country convergence equation. As in section 5, we use a basic time period of 5 years. This gives us 9 periods between 1960 and 2005. However, the first period is used for lags. All regressions are run for 1965 to 2005.

4.1. The basic cross-country regressions

Table 4 shows the basic cross-country regressions for the largest sample the data permits, 1965 to 2005, (Table 4a) and for LDCs alone (Table 4b). No control variables are included, and no countries have been deleted for any reason other than missing data. Each section contains stacked OLS-regressions with fixed effects for periods.

Thanks to the many observations included, most coefficients in both tables are significant; even when the power – measured by their marginal AR^2 -scores – of the income term and the aid term(s) are modest. In particular, the $MAR^2(\text{aid})$ in the 24 estimates presented are less than 0.007 on average.

However, the results do give a significant divergence in all but two of the regressions. The aid variables always give negative coefficients when only one is included. When both aid variables are included, their sum is negative in all four cases. So the results point to absolute aid ineffectiveness. Also, in most cases the merger of the two equations causes possibility 3 from Table 2 – the worst outcome. However, the movements in the coefficients are small indeed.

5. Panel estimates: Robustness to extra variables

Table 5 presents the same results as in Table 4, but for the panel version of the model.

24. Doucouliagos and Paldam's meta analysis (2006b, 2007a and b) include variables indicating the statistical techniques used. It is found that new data are far more important for the results than new estimators.

We have here used all available data between 1960 and 2005, divided in 9 periods of 5 years. We include fixed effects for countries and this changes the results, notably for the income and the lagged growth variables. For some of the regressions we also use Arellano and Bond GMM-estimates to control for simultaneity. In this section, the program is allowed to pick the highest number of observations possible for each regression. Section 5.3 contains a set of robustness tests.

5.1. Conditional convergence confirmed

When we control for fixed effects for countries, we always get conditional convergence, $\beta < 0$. All simple panel estimates of β are in the small range from -0.038 to -0.051, with t -ratios between 9 and 13. The β 's barely react to neither the merger between the two models, nor to the inclusion of lagged growth. In fact, when we try the six controls in 5.3, the estimate of β is still unchanged. Conditional convergence is thus confirmed as expected.

Tables 5a and 5b present the GMM-estimates using the Arellano-Bond estimator. These estimates actually decrease the estimated β 's. In fact, the estimates here appear unreasonably large (numerically). For reasons explained in Section 2.5, we prefer the OLS-estimates.

In Barro and Sala-i-Martin (2004; Chapters 11 and 12), it is argued that the estimates for β are in the range of -0.015 to -0.040. However, the meta-analysis by Abreu et al (2002) covering 48 studies shows a rather large range of estimates. It is difficult to imagine how fast countries would converge if they had precisely the same starting point and just differed by the income level.

5.2. Aid ineffectiveness remains

While the convergence coefficient, β , becomes significant, the aid effectiveness coefficient, μ , remains dubious. Of the 24 estimates, there is an almost equal number of positive and negative coefficients. And the GMM estimates are smaller than the ones of the panel estimates. There is no sign of a negative bias in the OLS-estimates. If there is a bias, it is upwards, as expected from the studies of the effect on growth of aid allocation.

We observe that μ has a systematic change of sign. When aid is unlagged, it is always insignificant and mostly negative, but when lagged, it is positive and significant in all four lagged regressions (3aL) and (4aL). At first glance, we thus find that aid works. However, as soon as we merge the equations – that is, control for development level – all four estimates of μ fall and become significantly negative or insignificant if lagged. This also happens in the GMM-estimates.

The coefficient of past growth plays a small role in the models, but it does suggest that there is a cyclical component in growth. This did not appear in the cross-

Table 5a. Panel estimates for all countries with fixed effects for countries and time periods.

Model	Start	Income, β	Aid, μ	Aid lagged, μ	Growth lagged, δ	N	NC
3a	60	-0.039 (-11.8)				1214	178
3b	60		0.017 (0.8)			1200	182
3bL	65			0.065 (3.4)		1038	179
3c	60	-0.043 (-12.6)	-0.029 (-1.4)			1144	174
3cL	65	-0.038 (-11.2)		0.021 (1.1)		1003	174
4a	65	-0.041 (-11.2)			0.003 (0.1)	1057	177
4b	65		0.014 (0.6)		-0.054 (-1.7)	1053	180
4bL	65			0.070 (3.4)	0.045 (1.6)	1018	179
4c	65	-0.046 (-12.0)	-0.044 (-1.9)		-0.037 (-1.2)	1013	174
4cL	65	-0.039 (-11.3)		0.021 (1.0)	0.079 (2.9)	985	174
GMM-estimates using the Arellano-Bond estimator:							
4a	70	-0.123 (-22.5)			0.006 (0.2)	878	173
4b	70		-0.056 (-1.8)		0.108 (3.0)	868	174
4bL	70			0.076 (2.7)	0.162 (4.1)	834	167
4c	70	-0.123 (-24.0)	-0.116 (-4.6)		-0.068 (-2.5)	835	170
4cL	70	-0.118 (-22.4)		-0.025 (-1.1)	0.006 (0.2)	807	164

Note: Calculated by STATA 9 that does not report R²'s.

Table 5b. Same estimates as in Table 5a, but for LDCs only.

Model	Start	Income, β	Aid, μ	Aid lagged, μ	Growth lagged, δ	N	NC
3a	60	-0.042 (-9.3)				912	137
3b	60		0.016 (0.6)			891	142
3bL	65			0.060 (3.0)		763	140
3c	60	-0.048 (-10.1)	-0.036 (-1.5)			853	135
3cL	65	-0.040 (-8.5)		0.018 (0.9)		738	135
4a	65	-0.045 (-8.9)			0.024 (9.3)	789	137
4b	65		0.009 (0.4)		-0.066 (-1.7)	782	141
4bL	65			0.065 (2.9)	0.054 (1.6)	749	140
4c	65	-0.051 (-9.6)	-0.050 (-1.9)		-0.024 (-0.7)	753	135
4cL	65	-0.042 (-8.9)		0.015 (0.7)	0.111 (3.4)	726	135
GMM-estimates using the Arellano-Bond estimator:							
4a	70	-0.124 (-17.7)			0.044 (1.2)	651	133
4b	70		-0.062 (-1.8)		0.081 (2.0)	638	135
4bL	70			0.064 (2.1)	0.156 (3.5)	606	129
4c	70	-0.126 (-19.0)	-0.119 (-4.2)		-0.042 (-1.4)	615	131
4cL	70	-0.121 (-17.7)		-0.028 (-1.1)	0.055 (1.6)	588	126

Note: See note to Table 5a.

country regressions (Table 4) because past growth worked as a proxy for country differences. When it is included, it nearly always reduces the coefficient to aid. If we accept this interpretation, we may go back to Figure 2 and interpret the growth effect we thought we found at lags $j = 1$ to 8 as a cyclical effect.

Our basic result is thus that aid has a negative rather than a positive effect on growth, but all in all this is dubious. As the numbers of observations are from 588 to 1214, even small effects should show up as significant. The results are much like the cross-country results as regards aid. Nothing is credible and significant even though the other two variables produce statistically significant results, and a very stable conditional convergence coefficient is found.

The two changes when the model is merged can be calculated for 12 (dependent) pairs. The values of $\Delta\beta$ change signs, and are small, with the average $\Delta\beta \approx -0.001$ (0.2). However, $\Delta\mu$ is negative in all cases, with an average value of $\Delta\mu \approx -0.059$ (3.3), where the parentheses contains t -ratios for cross-estimate stability. The size of $\Delta\mu$ is substantial relative to the size of the coefficients, and shows that the positive effect of aid lagged in the non-merged equation is an artifact. Thus, as in Section 4, we find no clear evidence of aid effectiveness. We cannot conclude that aid has prevented convergence, but only that if the level of GDP is not controlled for, the estimate is misspecified.

5.3. *Extra variables – some robustness tests*

Table 6a tests the basic equation (3c) of Table 5 for robustness by including a set of extra variables. Table 6b does the same for (3cL). We have included only variables with some basis in economic theory, so that each of the new variables included might tell a story.

The first line repeats the regression tested from Table 5a (but for a shorter period), and then six extra variables are included one at the time. The convergence coefficient stays as constant as one could wish. However, the coefficient to aid also remains largely stable. Only one of the extra variables changes the coefficient in size, but two of them destroy the significance, so the robustness of the aid effect is also reasonably good. The economic interpretation is straightforward:

Investment is known to be the most robust variable in growth regressions since Levine and Renelt (1992), and investment does get a large and significant coefficient in the estimate. However, it leaves the two other coefficients unchanged. The aid flows and domestic investments give rise to no multicollinearity. The next four variables are meant to catch aspects of domestic policies.

High inflation and high debt are strong signs of unsuccessful policies. However, while the inflation tax is an alternative to borrowing money, aid is not all gifts, so aid

Table 6a. Equation (3c) with one control variable, γx .

Aid unlagged	β	μ	γ	R ²	N	NC
Extra variable $x =$	Catch-up	Aid effect	Effect of x			
Equation (3c)	-0.047 (6.89)	-0.050 (1.56)		0.43	610	141
Investment	-0.044 (6.52)	-0.040 (1.27)	0.135 (3.48)	0.44	595	129
Inflation	-0.046 (7.15)	-0.053 (1.74)	-0.000 (0.04)	0.44	587	128
Debt burden	-0.060 (7.62)	-0.082 (2.41)	-0.001 (6.63)	0.46	547	126
Openness	-0.055 (7.99)	-0.065 (2.10)	0.050 (4.86)	0.46	595	129
Economic freedom	-0.054 (7.86)	-0.075 (1.71)	0.067 (7.66)	0.53	448	98
Gastil index	-0.049 (6.48)	-0.046 (1.38)	0.001 (0.08)	0.42	562	130
Aid squared	-0.048 (6.96)	-0.085 (1.76)	0.057 (0.97)	0.43	610	141

Note: This is done for the data from 1970-2000. Fixed effects for countries are included. Once more we have done everything for the LDCs alone and found the same pattern at a marginally lower level of significance.

Table 6b. Equation (3cL) - for aid lagged - with one control variable, γx .

Aid unlagged	β	μ	γ	R ²	N	NC
Extra variable $x =$	Catch-up	Aid effect	Effect of x			
Equation (3cL)	-0.046 (6.07)	0.037 (1.43)		0.54	499	141
Investment	-0.042 (5.60)	0.036 (1.41)	0.121 (2.77)	0.55	484	128
Inflation	-0.046 (6.03)	0.043 (1.69)	0.020 (2.01)	0.55	481	128
Debt burden	-0.047 (5.29)	0.043 (1.63)	0.001 (1.28)	0.53	450	126
Openness	-0.050 (6.75)	0.013 (0.51)	0.050 (4.84)	0.57	484	128
Economic freedom	-0.058 (7.10)	-0.066 (1.54)	0.059 (6.06)	0.56	391	98
Gastil index	-0.048 (5.55)	0.041 (1.51)	0.001 (0.10)	0.54	456	130
Aid squared	-0.046 (6.04)	0.038 (1.44)	-0.008 (0.54)	0.54	499	141

Note: See note to Table 6a.

adds to debt. These connections are quite visible in our regressions – there is an uncannily strong (and unreasonable) negative effect of aid when the debt burden is added. However, when the aid variable is lagged the effect disappears.

The effect of adding the economic freedom index decreases the effect of aid. This is likely to be an effect of multicollinearity, but we note that the economic freedom index has a rather strong effect on growth. Our experience is that it works better in log form, indicating that while it helps to go from a very highly restricted economy to a more liberal one, the additional effect of going all the way to laissez faire is not so large. Also, we have tried the Gastil index of democracy with little success. Finally we controlled for aid squared to test the Medicine Model from Section 2.4. It fails in both regressions.

6. The division hypothesis: Can an A- and a B-group be identified?

The literature survey above, and some of our findings, suggest a *division hypothesis*. Aid recipients may fall in two groups: The A-group where aid increases growth and the B-group where it harms growth. When the aid effectiveness relation is estimated on the data for both groups together, it becomes zero; but if we could delete the countries in the B-group, aid effectiveness would show up. This is what we try to do in the present section.

6.1. A line that should rise, as the countries are concentrated in the A-group

Below, we propose three criteria for sorting the countries into two such groups, and we use each criterion to make an experiment aimed at gradually deleting the potential B-group countries. Each experiment is represented by a line in Figure 3. To compare the three lines, we need a consistent sample, so this section looks at the 72 LDCs where the data set is complete.

We start each experiment by sorting the countries by the grouping criterion, and estimate the aid equation on all data. The starting estimate is therefore the same in all three division experiments. It is the regression:

$$g_{it} = \alpha - 0.083 y_{it} - 0.023 h_{it} + u_{it}, \text{ where } R^2 = 0.61, N = 198, NC = 72. \quad (3c)$$

(7.9) (0.4)

The estimate of the coefficient of h is -0.023, which is the point »•« at 0 on Figure 3 (for no countries excluded). As in Table 5 the starting point is close to zero.

The idea is to delete one country at a time, from the expected »bad« end, which should be the one of the B-group, and then to re-estimate the equation to see whether the coefficient of h rises. This gives observation 1, 2, ..., 40, as drawn. If the grouping criterion is correct, the sample should concentrate more and more on the A-group. This should make the curve rise. By using three division criteria, we should be able to tell which line rises the most, and this should indicate the best division criterion.

6.2. Three division criteria: by growth, by aid share and by income level

We have encountered two theories about the nature of the groups in the literature, and in addition we have noted that development theory contains a number of low level equilibrium trap models, see Azariadis and Stachurski (2005). We thus operate with three basic divisions:

By growth rate: We here delete the country with the lowest growth first, then the one with the second lowest growth, etc., and consequently concentrate more and more on the most successful countries. By the Good Policy Model, this should concentrate the

Aid effectiveness

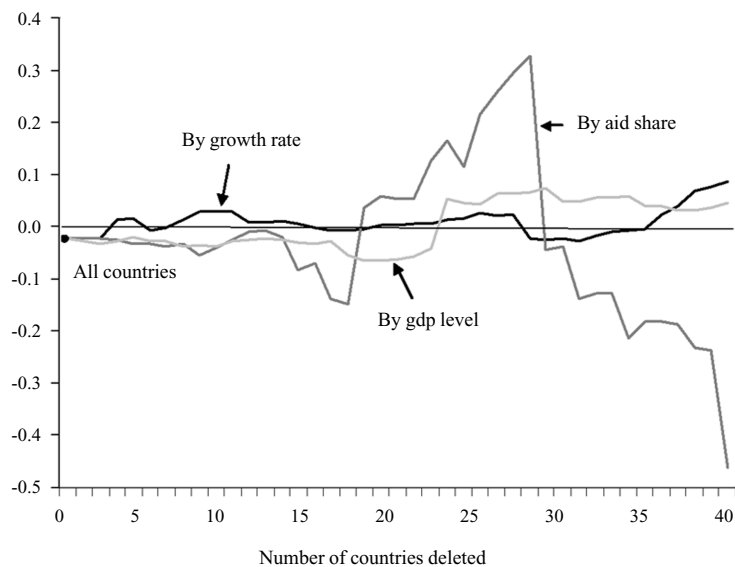


Figure 3. The three experiments: Looking for a rising curve.

countries more and more in the A-group, making the estimated aid effect rise. It is obvious that this experiment fails. There is no systematic rise in the relevant curve.

By aid share: Here we delete the country with the highest aid share first, and then the one with the second highest aid share, etc. By the Medicine Model, this should make the estimated aid effect rise.²⁵ When we look at the relevant curve it does move a great deal, and until 28 countries are deleted, it rises as it should, but then it turns around and falls. Thus, the experiment does not support the model.

By income level: Here the idea is that the poorest countries are in a low level trap, and if aid is not very large, it has no effect. However, once a country is out of the trap, aid works, so here we first delete the poorest country, then the second poorest, etc. Since the trap works until a certain threshold is reached, we expect a step up in aid effectiveness at a certain level. And this is very much what we see after the deletion of 23 countries. However, the step is not very high, so it is of dubious significance.

When everything is put together, we have to conclude that the evidence for the division hypothesis remains weak.

7. How »bad« are our bad results?

The above analysis tried to answer the following question: Does development aid

25. As the model has the terms $g = \mu h + \omega h^2$, the marginal effect of decreasing the aid share is $g' = \mu + 2\omega h$, which is linear. So by the model, the aid effectiveness should rise linearly.

help poor countries converge to our income level? The question is analyzed by standard cross-country and panel regression techniques. The answer is: The present aid flows have no clear effect on convergence.

The results of our study confirm the results of the recent meta-studies cited above so readers who are familiar with the literature should not be surprised. However we know that our results are controversial, see e.g., the more optimistic conclusions in the set of papers introduced by Hudson (2004) and the survey by McGillivray et al. (2006).

Our results are in line with Rajan and Subramanian (2005) and Easterly (2006). Many authors prefer more structured models, where the equation is controlled for half a dozen possible biases. However, it also appears that only few of the leading researchers are prepared to say that there exists clear macro evidence demonstrating that the present aid flows are an effective tool to close the gap between poor and rich countries.

This raises two questions: How can it be explained? Can the aid be reformed so as to generate catch-up? The next sections address these questions.

7.1. The micro-macro paradox and the negative growth externality of aid

If we accept the argument that the aid flows have little impact on the growth rate, this brings us to the micro-macro paradox. That is, the micro literature evaluating projects shows that about half of all aid projects succeed, the rest fail, but very few harm.²⁶ Thus, this literature suggests that the average project leads to some development. We can even assess how much.

Aid projects are often decided based on feasibility studies using social cost-benefit analysis which assess the contribution of the project to economic growth. The rule of thumb is that the cost-benefit ratio should be above 10% (0.1). When half the projects work, it corresponds to a realized social rate of return of at least 0.05. The average aid share is about 7.5%. Thus, aid should give $7.5 \cdot 0.05$ percentage points = 0.375 percentage points extra growth.²⁷ The average growth rate of the LDCs is about 1.5%. Consequently, no less than $100 \cdot 0.375 / 1.5\% = 25\%$ of the observed growth should be attributable to aid. This is substantial, and it should be easy to find by the methods used, but it does not show up. This is known as the *micro-macro paradox*.

It implies that even successful projects must have a negative externality on the real growth rate. Hence, there must be an invisible villain in the aid game. Three actors appear to be able to play that role.

(i) The Dutch Disease effect of aid.²⁸ In a quasistatic analysis, it is obvious that any transfer from abroad must appreciate the real exchange rate. The dynamic mechanism

26. See e.g., Cassen (1994), the annual IBRD-OED reports and Paldam (1997a).

27. These calculations can be endlessly refined, but the refinements we have considered go both ways.

28. The *disease* is the effect from a *booming aid sector* on the rest of the economy. The problem used to be termed the *transfer problem*. See Gylfason, Herbertsson and Zoega (1999).

Table 7. The status quo bias of aid.

-
- 1 A country is in disequilibrium, needing an adjustment
 - 2 The adjustment has short-run political costs
 - 3 The country receives an external donation, and does not need to adjust
 - 4 The disequilibrium grows, and so does the adjustment costs
 - 5 Go to 1
-

The country is hooked once the disequilibrium is so large that the cost of adjustment exceeds the limit the government can bear and stay in office.

whereby this is brought about depends upon the exchange rate regime, domestic policies, etc. The appreciation will inevitably harm the domestic tradables sector. If the projects generate little growth, and the tradables sector is the one that should generate the growth, a lopsided development might result. This mechanism can be dramatically observed in a few small countries receiving very much (resource rent or) aid, see Paldam (1997b) on Greenland and Hall and Herbertsson (2003) on Uganda.

(ii) The hidden cost syndrome. Studies of development often conclude that many LDCs suffer from a lack of executive capacity; i.e., both public and private sectors have too few competent personnel. Also, it is well known that aid projects use relatively much of this scarce resource. Hence, such projects may deprive other LDC activities of competent personnel, and thus make them fail. Consequently, aid projects may thus cause unrelated projects to fail.

(iii) The status quo bias of aid. Table 7 describes a likely mechanism. Many detailed studies of policymaking in countries receiving aid support the existence of a mechanism such as that described (see e.g., White, 1998).²⁹ However, aid is also sometimes used as a device for supporting reform, so perhaps the status quo bias has not been as strong in the recent past as it used to be.

7.3. Can aid become more effective and can the vicious circle be reversed?

In our opinion aid is caught in a *vicious circle*. It is caused by *the growing gulf between* promises and accomplishment in development aid. The gulf causes aid fatigue that comes in waves. When aid decreases, the need for bigger promises to everybody is necessary.

Aid is now rising, and promises have never been bigger. This has increased the number of goals of each development project, making projects less easy to monitor and evaluate, and consequently less efficient. This widens the gulf both ways.

29. The core of the book is four studies of countries that followed dirigist policies which failed and were reversed: Guinea-Bissau, Nicaragua, Tanzania and Zambia. They were for long supported by donors for political/ideological reasons. It is demonstrated that the support delayed adjustment. However, Burnside and Dollar (2000) claim that they could not find a connection between aid flows and economic policy changes.

Development aid is now supposed to abolish world poverty, bring about several types of sustainability, reduce discrimination against women and minorities of all kinds, build social capital, stop corruption and increase good governance, curb terrorism, prevent out-migration, increase the export of donor countries, make peace between warring states and tribes, reward friends, secure employment for aid workers, reconstruct countries after wars and natural disasters, etc. In addition, numerous NGOs are involved with still more diverse goals. There is hardly a problem in the world that aid is not supposed to cure or at least reduce.

All this is to be accomplished for about 0.3% of the aggregate GDP of the rich countries or about 2.5% of the aggregate GDP of the LDCs.

We suggest that aid could be made more effective, simply by reducing the number of goals and by making them (much) more coordinated, concrete and *realistic*. We thus propose aiming for simple, easily monitored, quantitative goals. The simple devise of first doing what can be done – i.e., start by harvesting the low-hanging fruit – may increase the efficiency of aid. If this could be shown to work, surely the willingness to give would increase. Maybe one could even hope that the vicious circle could be turned into a good one.

8. Conclusions

The purpose of this study was to explore convergence and aid effectiveness together, and to do so by fully transparent models. Everything is done as simply as possible; all observations are included, etc. The results are clear:

- (1) We found absolute divergence, and robust and significant conditional convergence. Both findings are very much as expected from the literature.
- (2) We found absolute aid ineffectiveness, and it remained ineffective in the conditional models. This is also as expected from the literature, though it has remained controversial despite 40 years of research.
- (3) Finally, as regards the interaction between the two models: Aid hardly has any effect on catch-up, but the level of income does reduce the effect of the aid effectiveness term.

The results (2) for aid effectiveness are controversial; not for reasons of economic research – they do confirm the results of almost 40 years of inconclusive research, but they have remained controversial simply because all of us want them to be different.

Perhaps we should simply note that the results are in accordance with the teaching of economics: Trade is better than aid.

Literature

- Abreu, M., Florax, R. J. G. M., Groot, H. L. F.d. 2005. A Meta-Analysis of b-convergence: the legendary 2%. *Journal of Economic Surveys* 19, 389-420.
- Aghion, P., Durlauf, S., eds. 2005. *Handbook of Economic Growth*. 2 vols. North-Holland, Amsterdam.
- Azariadis C., Stachurski, J., 2005. Poverty Traps? Cpt. 5, 295-384 in Aghion and Durlauf (2005).
- Barro, R. S. 1991. Economic growth in a cross section of countries. *Quarterly Journal of Economics* 106, 407-33.
- Barro, R. S., Sala-i-Martin, X., 1995. 2004. *Economic growth*. MIT Press, Cambridge, MA.
- Bauer, P. T. 1971. *Dissent on development: Studies and debates in development economics*. Weidenfeld and Nicolson, London.
- Baumol, W. J., 1986. Productivity growth, convergence, and welfare. *American Economic Review* 76, 1072-85.
- Burnside, C., Dollar, D. 2000. Aid, policies and growth. *American Economic Review* 90, 847-68.
- Cassen, R., et al. 1986, 1994. *Does aid work?* Clarendon, Oxford.
- Chang, C. C., Fernandez-Arias, E., Serven, L. 1998. *Measuring aid flows: A new approach*. World Bank. URL: <http://www.worldbank.org/research/growth/ddaid.htm>
- Christensen, P. W., Doucouliagos, H., Paldam, M. 2007a. *Master list of the AEL: The Aid Effectiveness Literature*. Posted on URL: <http://www.martin.paldam.dk>
- Christensen, P. W., Doucouliagos, H., Paldam, M. 2007b. *Master list of the AAL: The Aid Allocation Literature*. Posted on URL: <http://www.martin.paldam.dk>
- Doppelhofer, G., Miller, R. I., Sala-i-Martin, X. 2004. Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review* 94, 813-35.
- Doucouliagos, H., Paldam, M. 2006a. Aid effectiveness on accumulation. A meta study. *Kyklos* 59, 227-54.
- Doucouliagos, H., Paldam, M. 2006b. Conditional aid effectiveness. A meta study.³⁰
- Doucouliagos, H., Paldam, M. 2007a. Aid effectiveness on growth. A meta study. *European Journal of Political Economy* forthcoming.³⁰
- Doucouliagos, H., Paldam, M. 2007b. Aid effectiveness: The sad result of 35 years of research.³⁰
- Doucouliagos, H., Paldam, M. 2007c. Explaining aid allocation by growth. A meta study.³⁰
- Doucouliagos, H., Paldam, M. 2007d. A meta-analysis of development aid allocation: The effects of income and population size.
- Easterly, W. 2006. *The white man's burden. Why the West's efforts to aid the rest have done so much evil and so little good*. Oxford UP., Oxford, New York.
- Fraser Institute. Home of the Economic Freedom Index. URL: <http://www.fraserinstitute.ca/>
- Freedom House. Home of the Gastil Index. URL: <http://www.freedomhouse.org/>
- Friedman, M. 1958. Foreign economic aid: Means and objectives. *Yale Review* 47, 500-16.
- Griffin, K. 1970. Foreign capital, domestic savings and economic development. *Oxford Bulletin of Economics and Statistics* 32, 99-112.
- Gylfason, T., Herbertson, T. T., Zoega, G. 1999. A mixed blessing: Natural resources. *Macroeconomic Dynamics* 3, 204-25.
- Hall, A., Herbertsson, T. T. 2003. *Monetary transmissions and monitoring of the real economy in Uganda*. Report to the Governor of the Bank of Uganda.
- Hansen, H., Tarp, F., 2000. Aid effectiveness disputed. *Journal of International Development* 12, 375-98.
- Herbertsson, T. T., 1999. *Sources of economic*

30. Available from the working paper series of the Department of Economics at University of Aarhus or in the newest version from <http://www.martin.paldam.dk>.

- growth*. University of Iceland Press, Reykjavik.
- Hudson, J. 2004. Introduction: Aid and development. *Economic Journal* 114, 185-190. See also the three articles pp 191-271 in the same issue of the journal by Collier and Dollar, Dalgaard, Hansen and Tarp, and Hudson, Mosley and Verschoor.
- IBRD-OED. *Evaluation results*. (Annual) World Bank Operations Evaluation Department, Washington DC.
- Jensen, P. S., Paldam, M. 2006. Can the two new aid-growth models be replicated? *Public Choice* 127, 147-75.
- Kristiansen, J. K. L. 2007. Udviklingsbistand og Vækst – Empirisk identifikation af heterogene aid-effekter. MA thesis. University of Aarhus.
- Levine, R., Renelt, D.D. 1992. A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82, 942-63.
- Levy, V. 1988. Aid and growth in Sub-Saharan Africa: The recent experience. *European Economic Review* 32, 1777-95.
- McGillivray, M., Feeny, S., Hermes, N., Lensink, R. 2006. Controversies over the impact of development aid: It works; it doesn't; it can, but that depends ... *Journal of International Development* 18, 1031-50.
- Nuxoll, D. A. 1994. Differences in relative prices and international differences in growth rates. *American Economic Review* 84, 1423-36.
- Paldam, M. 1997a. Dansk u-landshjælp. Altruismens politiske økonomi. Aarhus Universitetsforlag, Aarhus Paldam, M. 1997b. Rent seeking and Dutch disease. An essay on Greenland. *European Journal of Political Economy* 13, 591-614.
- PWT, Penn World Tables, URL: <http://nber.harvard.edu/pwt56.html>.
- See also Summers, R., Heston, A., 1991. The Penn World Table (Mark 5). An Expanded Set of International Comparisons, 1950-1988. *Quarterly Journal of Economics* 105, 327-68.
- Rajan, R. G., Subramanian, A., 2005. *Aid and Growth: What Does the Cross-Country Evidence Really Show?* IMF WP No. 05/127. June (see also WP No 05/126).
- Sachs, J. 2005. The End of poverty. *How We Can Make It Happen in Our Time*. Penguin Books, London
- Sturm J.-E., Haan, J.d. 2005. Determinants of long-term growth: New results applying robust estimation and extreme bounds analysis. *Empirical Economics*, 30, 597-617.
- WDI, World Development Indicators, URL: <http://devdata.worldbank.org/dataonline/>
- Weisskopf, T. E. 1972. The impact of foreign capital inflow on domestic savings in underdeveloped countries. *Journal of International Economics* 2, 25-38.
- White, H., ed. 1998. *Aid and macroeconomic performance: Theory, empirical evidence and four country cases*. (International Finance and Development Series) Palgrave Macmillan, London.
- World Bank. 1998. Assessing aid: *What works, what doesn't, and why*. A Policy Research Report. Oxford UP New York.

Danish Transport Regulation and Labour Market Effects

Jens Hauch

Det Økonomiske Råds Sekretariat, E-mail: JH@dors.dk

Ninette Pilegaard

Danish Transport Research Institute, E-mail: np@dtf.dk

Thomas Bue Bjørner

AKF – Danish Institute of Governmental Research, E-mail: tbb@akf.dk

SUMMARY: We present a partial equilibrium model for road pricing in Denmark. The model analyses the joint problems of the taxation of the congestion externality and the effects on the labour supply. We combine two regulatory instruments, a toll ring and kilometre-based road pricing, with three ways of recycling the revenue (eased income taxes, increased subsidies for public transport and increased tax deductions for commuting). We find that the largest gain would arise from a combination of a toll ring and eased income taxes. Kilometre-based road pricing is less beneficial because the system costs are higher. Furthermore we find support for the hypothesis of weak double dividend but not strong double dividend. The model is inspired by Parry and Bento (2001) and Van Dender (2003) and extended in several ways, most important are the inclusion of several regions, more alternative modes of transportation, regionalised labour supply, modelling of a toll ring, peak/off-peak substitution and the inclusion of system costs.

1. Introduction

Due to high car taxes, transport congestion problems in Denmark seem to have been less severe than in a number of other European countries. However, traffic levels have been growing steadily over the past decade and there has been increased focus on congestion problems, especially in Copenhagen. For example, traffic volumes in Copen-

Thomas Bue Bjørner and Jens Hauch were both working at the Secretariat of the Danish Economic Council at the time the study was carried out. The authors are thankful for comments and discussions during the research process from Peter Birch Sørensen, Jørgen Birk Mortensen, Jan Rose Skaksen, Michael Rosholm and Peder Andersen as well as from several other former colleagues at the Secretariat. The responsibility for the article is, however, solely the authors.

Financial support from the Danish Social Science Research Council is gratefully acknowledged.

hagen have increased by 17% since 1990, while speed seems to have decreased at the same rate, The Danish Economic Council (2006). Congestion indicators also suggest a severe increase in congestion on the motorways around Copenhagen, The Danish Road Directorate (2005). Finally, a recent study by Nielsen (2005) shows that traffic congestion each day causes 120,000 hours of delay in and around Copenhagen.

From an economic point of view, congestion can be regarded as a classical externality: the individual traveller does not take into account in his trip decision the fact that his trip will reduce the speed of other travellers. The textbook solution is a Pigovian road-pricing tax that internalizes the externality in the trip decision of individuals. Until recently this was just a hypothetical text book case, but introduction of GPS technologies makes the necessary monitoring possible in practice. This technology is, however, still being developed and the system has only been used in demonstration projects. High and uncertain costs must be expected. In practice, toll rings have often been used instead of road pricing. London and Stockholm are recent examples of large cities that have introduced toll rings. The necessary technology is therefore available and well tested.

It has been pointed out by some economists that the expectations with respect to the benefits of road pricing or toll rings may be too high. In an early evaluation of the London scheme, Prud'homme and Bocarejo (2005) found that the benefits of the system were considerably lower than the cost (due especially to the high implementation and investment costs). Optimal regulation reduces but typically does not eliminate congestion.¹ Thus, given that private car travel is already heavily taxed in Denmark, the 120,000 daily hours lost in congestion noted above could, in principle, reflect an optimal regulation level in some but not all areas (or a level of congestion in some areas being too low).

Congestion regulation is complex, and several things should be taken into account when deciding on the level of regulation. First, there are large investment and operational costs associated with road pricing and toll ring schemes. Second, there are a number of other externalities associated with traffic (noise, air pollution, barrier effects etc) that should also be taken into account. Third, road pricing will have an effect on the income distribution between individuals and between regions. Fourth, a tax on transport may increase or reduce the impact of other distortions in the economy.

1. This is expressed in the following way by Arnott (2005): *»Traffic congestion is so high because of the spatial concentration of economic activities in cities. Everyone benefits from this spatial concentration through new and more varied products, lower prices for many consumer goods and higher economic growth, and city residents additionally through higher wages, ready access to experts, urban amenities and a richer set of social contacts. Traffic congestion is simply one of the costs we pay to enjoy these benefits. It is excessive congestion (due to under priced auto travel) that should be our principal concern.«*

There are basically two arguments for taxing transport activities: correction for external effects (Pigovian tax) and the need to raise public revenue (Ramsey tax). According to the Ramsey argument, goods and services that have low demand elasticities should be taxed relatively high, as this will create less distortion than taxing goods and services with high levels of elasticity of demand.

Taxation of commuter traffic may have a significant indirect effect on the labour market. About a third of all transportation is commuting or work related. A tax on commuting transportation will indirectly serve as an additional tax on the already highly taxed labour supply, and thus lead to lower labour supply. Leisure transport, on the other hand, is a complement to leisure time, which is untaxed. This suggests that commuting transport should be taxed less than leisure transport, Van Dender (2003) and Munk (2003). This is the efficiency rationale for the Danish commuting tax deduction.² An optimal level of road pricing will also depend on the will to modify other transport related taxes and subsidies. In the past, one of the economic efficiency arguments for subsidizing urban public transport has been that it was not technically feasible to tax private urban traffic at a higher rate than rural traffic. Given this earlier technical restriction on the available policy instruments, subsidies to urban public transport may have been a good second best instrument, Glaister and Lewis (1978). However, in the ideal situation, where the externalities of private transport can be regulated directly, public transport should be taxed (instead of subsidized) according to its marginal external effects (congestion and environmental effects).³

A number of earlier papers on the complex regulation of transport externalities have applied partial equilibrium models with a fairly detailed description of the transport system, but without explicitly including of the derived effects on the labour market, see for example De Borger et al. (1996) and Proost and van Dender (2001). Another more recent group of articles include the derived labour market effects, see Parry and Bento (2001), van Dender (2003) and Parry and Small (2005). These studies suggest that the use of road pricing revenue combined with the derived effects on the labour market have greater welfare implications than the benefits stemming from correction of the congestion externality. Parry and Bento find that the benefit of road pricing is doubled if the revenue is used to reduce the distortionary tax on labour. The models in the later studies are, however, very stylized, with only simple descriptions of the transport demand system.

2. However, in the long run a lower tax on commuting could lead to a non-optimal localisation pattern. Thus, there is a trade off between the short run flexibility of the labour market (low tax on commuting) and the desire for an optimal localisation pattern (high tax on commuting).

3. There may of course be income distribution arguments for subsidising public transport, but it should also be noted that subsidies to public transport may not be a very finely tuned instrument for achieving a given distribution target. There may also be economics of scale in public transport which is a further argument for subsidizing this.

The contribution of this present paper is to combine such types of labour supply models with a fairly detailed description of the transport demand system based on Danish data. In addition, we include three different regions in the model to allow for a calculation of the regional distribution impacts of different types of regulation. Partial models without labour supply effects have previously been used to compare the welfare effects of road pricing and toll rings. These studies have, however, not explicitly included the cost of the different systems. We explicitly include the costs of road pricing and toll rings in the model. As the system cost of a toll ring is lower than for road pricing, the net revenue from the toll rings tend to be larger than from the road pricing system. It is shown in Parry and Bento that the size of the net revenue cannot be ignored if the net revenue is used to reduce other distortions in the economy. On the one hand, a toll ring can be considered less efficient than road pricing, as the toll ring only indirectly targets the congestion externality. On the other hand, the lower cost of the toll ring may yield higher net revenue, which may be used to reduce other distortions.

The model used for this study – denoted ASTRA – is a partial equilibrium model with three regions (Copenhagen, Greater Copenhagen and the rest of Denmark). We only consider passenger transport. Consumers are allowed to substitute between different modes of transport and between travelling at peak or at off-peak hours. Congestion is explicitly modelled using a speed-flow function, while geographically differentiated marginal costs of other externalities are also included in the welfare calculation. Labour supply is determined in the model with substitution between consumption and leisure. Finally, the model distinguishes between transport for leisure and for commuting, where the latter is linked to labour supply. Thus labour supply is linked to commuting time via the speed-flow function.

The results of the analyses using the model suggest that there would be a small gain from a toll ring around Copenhagen, while road pricing appears to have a small negative impact on welfare due to the relatively high annualised cost. In the public debate on road pricing and toll rings, it is often argued that the revenue should be used to increase the subsidies to public transport. In contradiction to this point of view, the model shows that the welfare benefits of the toll ring would be substantially higher if the revenue was used to reduce income tax instead of subsidising public transport. Finally, without a toll ring – i.e. with a situation similar to the current regulation regime in Denmark – there would be a welfare gain from a further increase in the high taxes on car use, i.e. an increase in the tax on car use in both rural and urban areas.

The model is presented in the next section. Input data and model calibration are described in section 3, while results of analyses using the model are presented in section 4. Conclusions are offered in section 5.

2. The Model

We follow Parry and Bento (2001) and van Dender (2003) and model a partial equilibrium model with emphasis on the description of consumers' transport activities and their supply of labour. The model has three regions and it is calibrated to describe the Danish economy. There is a representative consumer in each region with a fixed residence location. The public sector collects different kinds of taxes, subsidises public transport, finances fixed government consumption and pays a lump sum transfer to all consumers. The public sector operates with a balanced budget. There is no explicit description of firms or international trade. Firms demand labour at a fixed wage and supply goods and commodities at fixed prices. Therefore changes in labour supply do not affect the prices of labour or commodities. Hence, the model is not a general equilibrium model.

In the modelling of the transport system, we include four modes of transportation (car, bus, train and cycling/walking), and there is a distinction between peak hour and off-peak travel, and between commuting and leisure transport. Congestion on roads is included by using an aggregated speed-flow function for each region. In addition to congestion we include geographically differentiated marginal external costs of externalities such as noise, air pollution and accidents. Labour supply is determined via substitution between leisure and consumption, subject to both financial and time budget constraints. Finally, we include the annualised costs of road pricing and toll rings. There is no description of freight transport.

In the following the model will be described briefly. A more detailed description of the model can be found in Hauch et al. (2006). Endogenous variables are in capital letters, while parameters and exogenous variables are in lower case letters.

The transport system

Consumers demand trips as a part of their demand for commodities. Transport demand is described by the number of trips purchased by the consumers, $Q_{i,p,od,m,t}$. The indices are defined in the following way: Index i represents the geographical location (residence) of the consumer: $i \in \{\text{Copenhagen, Greater Copenhagen, Rest of Denmark}\}$.⁴ Index p represents the purpose of the trip: $p \in \{\text{commuting (c), leisure (n)}\}$. Index m represents the transport mode: $m \in \{\text{car, bus, rail, light (bicycle/walking)}\}$. Index t represents time: $t \in \{\text{peak, off-peak}\}$. Index od represents the origin and destination of the trip.

The number of trips is transformed into traffic in each region in the following way: $cap_{i,p,od,m,t}$ is the number of passengers per conveyance and $dist_{i,ii,p,od,m,t}$ is the di-

4. Copenhagen (1) is the capital and is defined as the municipalities of Copenhagen and Frederiksberg. Greater Copenhagen (2) is defined as the counties of Copenhagen, Frederiksborg and Roskilde, and Rest of Denmark (3) is defined as the rest of Denmark. Greater Copenhagen surrounds Copenhagen which is explicitly modelled.

stance driven in region ii when a consumer living in region i purchases a trip with the origin-destination combination od (the set ii includes the same elements as the set i).

The amount (flow) of traffic in region ii , $F_{ii, m, t}$, is given as a simple summation of the kilometres driven in the region ii :

$$F_{ii, m, t} = \sum_{i, od, p} (Q_{i, p, od, m, t} \cdot dist_{i, ii, p, od, m, t} / cap_{i, p, od, m, t})$$

We furthermore define the road traffic, $FR_{ii, t}$, as the traffic flow on roads in each region (with buses weighting twice as much as cars):

$$FR_{ii, t} = F_{it, car, t} + 2 \cdot F_{it, bus, t}$$

The speed, $S_{ii, m, t}$ for cars and buses is assumed to depend (linearly) on the road traffic flow in the region:

$$S_{ii, car, t} = \alpha_{ii, car, t} - \beta_{ii, car, t} \cdot FR_{ii, t} \text{ and } S_{ii, bus, t} = \alpha_{ii, bus, t} - \beta_{ii, bus, t} \cdot FR_{ii, t}$$

It is assumed that traffic levels do not affect the speed of trains and light traffic (i.e. no congestion for these modes).

Transport has two types of cost, a monetary cost and a time cost. The time use per trip, $T_{i, p, od, m, t}$ for consumers living in region i , is composed of the time use in each of the regions ii in which the trip takes place, and these time uses depend upon the distance and speed in the regions ii . The time use per trip is given by:

$$T_{i, p, od, m, t} = \sum_{ii} \frac{dist_{i, ii, p, od, m, t}}{S_{ii, m, t}}$$

The private monetary cost, $P_{i, p, od, m, t}$, per trip is given by:

$$P_{i, p, od, m, t} = \sum ((c_{i, ii, p, od, m, t} + tax_{i, ii, p, od, m, t}) \cdot dist_{i, ii, p, od, m, t}) + toll_{i, p, od, m, t}$$

Here c is the factor cost per kilometre, while tax is the tax per kilometre (or if negative, the subsidy). The tax may depend on trip purpose if there is a tax deduction instrument for commuting trips (as is currently the case in Denmark). $toll$ is the fee for passing $toll$ rings.

In addition to congestion, traffic creates other externalities such as accidents and pollution. We assume that the external marginal costs of these are fixed and included in the problem of the social planner, but not in the optimisation of the trip decision of the agents.

The representative consumer

Besides transport, the representative consumer in each region consumes pure leisure N_i and a composite consumer good Q^0 . The utility function of the representative consumer is given by:

$$\Psi_i(Q^0, Q_{i,n,od,m,t}) + \Gamma_i(N_i) + \Omega_i(Q_{i,c,od,m,t}),$$

where Ψ represents the contribution to utility from consumption of the composite good and leisure trips. The additive separable term, Γ , represents the contribution to utility from pure leisure. This follows the specification of the Danish DREAM model, see Knudsen et al. (1998). Following van Dender (2003) and Parry and Bento (2001), commuting is included as an individual additive separable term in the utility function Ω . The commuting sub-utility function Ω allows commuting modes of transport to be imperfect substitutes.

The consumer has preferences for leisure trips to different regions and therefore substitutes imperfectly between leisure trips with different *od* combinations. In addition, he also chooses the transport mode and whether to leave at peak or off-peak hours for each trip. Similarly, the consumer has different preferences for working in each region, and substitutes imperfectly between labour supply to the three regions as well as choosing the mode of transport and whether to leave at peak or off-peak times for each commuting trip. The trips (leisure and commuting) are tied to utility units by a nested CES function, see Hauch et al. (2006).

Lower travel costs may improve the matching between worker and job, which will lead to improved efficiency. This is, however, not included in the model.

Note that the model focuses on the economic consequences of transport and labour-market-related policies. Consequently, the description of transportation-mode-choice is relatively simple, and there is no route-choice in the model.

Labour supply from the consumer in region i is given as L_i . The length of a work-day is assumed to be exogenously given, while the number of days worked is endogenous. The consumer can choose to supply his labour in each of the three regions. To supply a day of labour the worker always needs transport. Therefore, the labour/leisure decision is closely connected to the consumption of commuting transport (strict complementarity between labour and commuting). Thus, there is the following relation-

ship to the labour supply for the consumer living in region i , where h_i is the labour supply connected to a commute (i.e. the length of the work-day) for a consumer in region i :

$$L_i = \sum_{od, m, t} h_i \cdot Q_{i, c, od, m, t}$$

The income of a consumer is given as after-tax labour income and lump-sum transfers (Q_i). Net income is spent consuming transport and other consumption. tax^L is income tax, w_i is wage and P^0 and Q^0 are the quantity and price of a non-transport commodity. Note that the wage is determined by the localisation of the household. Different wages are included to take into account that the productivity of employees depends on where they are living. Living in Greater Copenhagen is generally considered attractive and housing prices are higher, which increases the share of high income (and therefore high productivity) consumers in this area.

$$(1 - tax_i^L) (w_i L_i + O_i) = P^0 Q^0 + \sum_{p, od, m, t} P_{i, p, od, m, t} \cdot Q_{i, p, od, m, t}$$

The consumer's time is restricted by an initial endowment, \bar{T} .

$$\bar{T} = N_i + L_i + \sum_{p, od, m, t} T_{i, p, od, m, t} \cdot Q_{i, p, od, m, t}$$

When the consumer spends time on transport (either leisure or commuting) this reduces his time of pure (utility-generating) leisure.

The consumer's problem is therefore to maximise his utility function subject to a financial and a time budget and subject to strict complementarity between labour supply and commuting.

The model captures linkages between the labour market and the transport sector by distinguishing between commuting and leisure trips. The model also includes a description of a wide range of choice situations which reflect the welfare effects of regulation (substitution between mode, peak and non peak travel time and destination). Nevertheless, there are also limitations that should be taken into account when interpreting the results: Firstly, as it is a partial model it does not include general equilibrium effects, which implies that feed back effects from the economy to the transport sector are not taken into account. This is especially important when larger changes are analysed. Secondly, firms demand for transport of persons and freight is not included. This implies

that not all gains are taken into account. Thirdly, the transport network in the three regions is represented by aggregated links, so there is no description of route choice. This implies, e.g., that detours caused by road pricing or toll rings are not taken into account.

3. Data

The model is calibrated to the year 2003 and data are from that year (primarily). The data for the model are obtained from a wide range of sources. The data are briefly described in the following. A more detailed description of the variables and parameters of the model can be found in Hauch et al. (2006).⁵

General economic variables such as labour supply, transfers, and income and commodity taxes were obtained at regional levels using the regionalized national account databases of AKF (Danish Institute of Governmental Research).

Data on the level of transport trips are calculated from the Danish TU data (collected by the Danish Transport Research Institute, DTU), which are interview-based trip diary data. The average number of trips by residents and the average travel distances of these trips were calculated, according to the following factors as defined in section 2: Residence (Copenhagen, Greater Copenhagen, Rest of Denmark), mode of transport (car, bus, train, cycling/walking), time (peak, off-peak), purpose of trip (commuting, leisure), origin-destination combination of trip.

The speed-flow functions, which are necessary to calculate the impact of changes in traffic levels on travel time, and ultimately on congestion externality costs, are based on information from the Danish OTM traffic model covering the areas of Copenhagen and Greater Copenhagen.⁶ A simple linear functional form is used, as this has a reasonably good fit between average speed and average traffic for relevant traffic levels.⁷ For »Rest of Denmark« it was assumed that traffic flow does not influence speed (i.e. no congestion externality).

Marginal cost estimates from other externalities than congestion (air pollution, noise, accidents etc.) were obtained from the Danish Ministry of Transport (2004).

5. The 2003 data have been used directly and have not been adjusted into a general equilibrium. Hence imbalances from 2003 are therefore included in the analyses, and therefore the results do not represent a long run general equilibrium.

6. This information was kindly provided by Otto Anker Nielsen, Jeppe Husted Rich and Stephen Hansen from the Institute of Traffic and Transport at the Technical University of Denmark. The slope of the speed-flow functions were adjusted subject to results reported in Nielsen (2005), where actual speeds of a large number of cars were monitored using GPS technology.

7. Note that the linear speed-flow relationship yields a marginal congestion cost that increases exponentially with traffic levels; see in general Newbery (1990) and Maddison et.al. (1996) or the Danish Economic Council (2006) for calculations of the marginal external congestion cost based on the speed-flow functions used for Copenhagen and Greater Copenhagen.

The annualised investment and operating costs of road pricing (GPS based) and toll rings have been calculated using information from a recent Danish study, Wrang et al. (2006). The annual cost of a road pricing system is DKK 510 million, which is considerably greater than the cost of the simpler toll ring technology of DKK 210 million per year.

The prices of and taxes on car use have been calculated by including variable costs (fuel, oil, tyres and repairs), annual costs (insurance and car ownership taxes) and annualised car purchase costs. The fixed car costs are included in order to reflect the long-run cost of car use in a simplified way. The interpretation is that changes in the demand for car trips will affect the level of car ownership. Thus, the current model does not distinguish between a change in the gasoline tax or the car ownership tax.⁸

Turning to the behavioural parameters, the labour supply elasticity used is 0.2. This is composed of a working-hour elasticity taken as 0.1, Frederiksen et al. (2001), and a participation elasticity of 0.1. The substitution elasticities between transportation and consumption on the one hand, and between car use and non-car use on the other, were chosen to reflect the price elasticities obtained by Fosgerau et al. (2004), Transportrådet (1999) and Børner (1994). The substitution between peak and off-peak travel time was chosen to allow for some (though modest) substitution based on results from the OTM traffic model on the traffic implications of road pricing and toll rings, see Københavns Kommune (2005). The substitution elasticity between workplace locations was determined ad hoc, see Hauch et al. (2006).

4. Model simulations

In this section we first discuss the regulatory instruments analysed, and then present the results. Next, sensitivity analyses are presented. Finally, a few supplementary analyses are presented.

The regulatory instruments

We analyse two types of road pricing systems, a kilometre-based system (Pigovian tax) and a toll ring. The greatest congestion problems in Denmark are in the capital area. We therefore focus on road pricing in Copenhagen and a toll ring around this area.

With a toll ring car drivers pay a fee each time the ring is passed, but driving is free both inside and outside the ring. With toll rings it is possible to differentiate the toll with respect to time of day and characteristics of the car or driver. A toll ring is inclu-

8. In principle the car ownership and use decision is more complex, see for example Børner (1997) and the model could be extended to allow for separate effects on car use and car ownership. However, this is beyond the scope of this paper.

ded in the model as an additional fee paid by the driver on trips passing the ring. The size of the fee depends on when the trip takes place.

With Pigovian road pricing a fee is levied for each kilometre driven. With such a system it is possible to differentiate the fee not only with respect to time of day and car/driver characteristics but also with respect to exactly where the driving takes place. A kilometre based system is included in the model as an additional fee paid by the driver for each kilometre driven. The size of the fee depends on where and when the trip takes place. In the model we focus on two instruments:

- A kilometre-based road pricing system in Copenhagen with a charge of DKK 4 per kilometre in peak hours and DKK 2 per kilometre in off-peak hours.
- A toll ring around Copenhagen with a fee of DKK 40 per passage in peak hours and DKK 20 in off-peak hours.

The toll levels chosen for the toll ring are approximately to twice the Stockholm tax level. Our simulations indicate that high taxes are needed to achieve a welfare gain, because of the high fixed costs that are necessary for running the systems. The level of the kilometre-based tax is chosen so that the gross revenues from the two systems are of approximately the same size. The main differences between the two systems in the model is that the fixed costs of the toll ring are smaller, and that the toll ring does not directly affect the cost of car trips that both start and end either inside or outside the ring. Thus the ring does not directly target the congestion externality, as only some of the relevant car trips are regulated.

We also look at three alternative ways to recycle the revenue. Recycling of the revenue is important, as it can contribute to increasing the labour supply and thereby production and total welfare. We assume that public consumption is unchanged, i.e. that all the net revenue is recycled to consumers. We analyse three types of recycling:

- Reductions in income taxes
- Subsidies to public transport in Copenhagen
- Commuting tax deductions for trips in Copenhagen

The income tax reductions are constructed so that the public balance towards the consumers in each region is unchanged, i.e. the regions where the consumers are most affected by the road pricing also experience the largest reductions in income tax. In practice this can be done by transferring the revenue to the local governments and in the model it is assumed that they reduce their income taxes correspondingly. This targeting

Table 1. Annual welfare gain of alternative road pricing systems in Copenhagen combined with different methods of recycling the revenue.

	Income tax reductions	Subsidises to public transport DKK, millions	Commuting tax deduction
Road pricing	-200	-338	-379
Toll ring	244	109	55

Source: Own calculations with ASTRA.

of the income tax reduction does not necessarily maximize the total welfare but it is politically tractable as in most cases all regions gain if there is a positive total welfare gain from the system.

Income tax reductions stimulate the labour supply. A welfare gain can be expected because the distortion on the labour market is reduced. Earmarking the revenue to subsidise public transport is widely debated and makes public transport more attractive, and can therefore potentially reduce congestion on the roads further. Seen from an economic point of view such earmarking should be avoided as it is most likely that it will not result in the optimal subsidy to public transport. Seen from a political point of view earmarking can be attractive as it may be more acceptable to the public. Commuting tax deductions reduce the effects of the road taxes on the total amount of transport but maintain an incentive to switch to non-car use modes of transport.

We take the existing tax system as given, and impose the regulation on the top of this. This implies that inefficiencies in the existing system are also present in the scenarios analysed. The results therefore do not correspond to an ideal solution but are probably closer to a realistic situation.

Results

Table 1 summarises the results from using the two instruments with the different recycling schemes.

It is seen from the table that there is a welfare reduction from using road pricing at the given level. The welfare reduction is smallest when revenue is used to reduce income tax. Thus, the gain from the road pricing regulation is smaller than the annualised fixed costs of the system.

There is an annual welfare gain from a toll ring equal to DKK 244 million if the revenue is used for reducing income taxes. The toll ring is to be preferred to a kilometre-based system because of the lower fixed costs.

The difference between the welfare gains from a kilometre-based system and from a

toll ring is larger than the difference in system costs. This may seem surprising, as road pricing can be targeted more precisely. However, the greater recycled revenue because of lower system costs from a toll ring stimulates the labour supply more via the larger reductions in income taxes. This increases the difference between the welfare effects of the two systems.

Income tax reductions are the most effective means of recycling the revenue, because of the stimulation of labour supply. Increased labour supply increases *ceteris paribus* the amount of transportation, and thus increases congestion and other externalities. These negative effects are, however, more than counter-balanced by the positive effects from increased income.

It is not efficient to use the revenue to subsidise public transport further. One reason is that labour supply is not stimulated as much as with reduced income taxes. Today public transport is heavily subsidised due to distributional concerns. Increasing these subsidies further will increase the present tax distortion. Similarly, increased commuting tax deductions are not the most effective means of recycling revenue either. Congestion is reduced by less, as the commuting tax deduction makes commuting more attractive; that is to say, the positive effects from the toll ring are reduced. The level of redistribution between consumers with high transport consumption and consumers with low transport consumption is, however, less with increased commuting tax deduction or subsidies to public transport than it is with income tax reductions. Subsidies to public transport are seen to be better than commuting tax deductions. One reason for this result is that a commuting tax deduction stimulates not only public transport but also private transport and therefore reduces the effect on congestion. Subsidies to public transport stimulate only public transport.

Table 2 shows the detailed consequences of a toll ring around Copenhagen combined with reduced income taxes.

The largest welfare gain is achieved by consumers living in Copenhagen. They are the most affected by the toll ring, as they experience the largest reduction in congestion and receive the greatest reduction in income tax. However, they also pay the highest total road taxes per capita. Equal reductions in the income tax across regions will change the distribution of income but will also result in a positive net benefit.

The labour supply increases both in Copenhagen and in Greater Copenhagen. The toll comprises a barrier to labour supply and reduces mobility. However, the effect of recycling the revenue to reduce income taxes increases labour supply, as does the increased speed of traffic. The total effect is an increased labour supply.

The total positive effect on labour supply is not uniform. Interregional commuting is reduced, but intraregional commuting is increased.

With the design of the income tax reduction all regions benefit from the toll ring. Our model does not, however, differentiate between consumers with different income

Table 2 Consequences of a toll ring around Copenhagen combined with reduced income taxes. Fee, DKK 40 per car in peak hours and DKK 20 in off-peak hours.

	Copenhagen	Greater Copenhagen	Rest of Denmark
	Change, DKK millions		
Welfare	139	93	11
Revenue from taxation	980	1,350	59
Change, %			
Consumption of transportation	-2.26	-0.49	-0.02
Other consumption	0.34	0.29	0.00
Leisure	0.03	0.01	0.00
Externalities	-3.30	-0.60	-0.01
	Change, % points		
Income tax	-1.0	-0.7	0.0
	Change, %		
Labour supply, from	0.22	0.16	0.00
Commuting:			
From Copenhagen, to	2.5	-6.1	-(^a)
From Greater Copenhagen area, to	-6.2	1.9	1.7
From Rest of Denmark, to	-(^a)	0.20	0.00
Speed (car)			
Peak hours	1.60	0.17	0.00
Off-peak	0.13	0.02	0.00
Traffic (car)			
Peak hours	-6.17	-0.63	0.00
Off-peak	-2.42	-0.38	-0.02

Note: (^a) Excluded from the model as the number of trips at these OD-combinations are several thousand times less than the demand level of the other trips. This creates numerical problems in the model. The demand levels for the excluded trips are so low that their exclusion has only a very limited influence on the result.

Source: Own calculations with ASTRA.

levels within the same region. It is most likely that not all consumers will gain from the toll ring.

The toll ring with a reduction in labour tax decreases road traffic by 6% in peak hours and 2% in off-peak hours. The traffic passing the toll ring will be reduced by 9%. The traffic reduction implies an increase in speed of 2% in peak hours, but there is almost no change in speed in off-peak hours. The consumption of transport commodities and the level of externalities will be reduced. There are two opposite impacts on leisure. On the one hand, increased labour supply decreases leisure. On the other hand, the lower level of congestion increases the amount of leisure. In total leisure is increased.

Table 3. Decomposition the welfare effects of the toll ring.

	Welfare change, DKK millions
(1) Toll ring with income tax redistribution of revenue (as in Table 1):	244
(2) As 1) but exclusion of environmental effects	180
(3) As 1) but exclusion of environmental and congestion effects	-321
(4) As 1) but with exclusion of environmental and congestion effects and administrative costs	-24
(5) As 1) but with lump sum redistribution of revenue	221

Source: Own calculations with ASTRA.

It should be noted that the 9% decrease in the traffic passing the ring is rather small compared to the effect found in London and Stockholm. The traffic passing the ring in Stockholm has been reduced by more than 20% with a toll that is only half the size of the Danish. The revenue in Stockholm is not returned to the consumers but used for financing the system, i.e. a higher reduction should be expected in Stockholm. The difference between the Stockholm results and ours are, however, so large that it indicates that the model may underestimate the effect on traffic.

Other types of transport are affected in several ways. There is a substitution away from cars towards the other types of transport – the substitution effect. The increased cost of car transport decreases the demand for car transport, which reduces the demand for all transport commodities – the income effect. The increased labour supply increases total demand, including demand for transportation. The demand for bus transport is furthermore affected by the reduced congestion resulting from the smaller number of cars on the road. It is assumed that buses are not taxed at the toll ring, and there will be an increase in the demand for bus transport.

Decomposition of results

A decomposition of the most important drivers behind the result is found in table 3.

It can be calculated from table 3 that the welfare effect of the improved environment is DKK 64 million per year (difference between (1) and (2)). The welfare effect of reduced congestion is DKK 501 million per year (difference between (2) and (3)). If it is assumed that there are no administrative costs and if there are no changes in externalities (environment and congestion) the welfare effect would be DKK -24 million per year, i.e. quite small. The value of redistributing the revenue through reduced income taxes instead of a lump sum is surprisingly small, only DKK 23 million per year (difference between (1) and (5)). On reason for this is that the increased labour supply stemming from the reduced income tax increases commuting (primarily in peak hours) which reduces the effect on congestion.

The most important drivers behind the result are therefore the welfare effect of reduced congestion and the administrative costs.

The results contribute to the discussion of double dividend. The double dividend hypothesis states that taxes on externalities can generate two types of welfare gains (dividends). The first gain is the environmental improvement – in our case both environmental improvement and reduced congestion. The other gain comes from increased economic efficiency as the revenue from the externality tax can be used to reduce other distortionary taxes such as income taxes, see e.g. Gørtz et al. (1999) for a general discussion on double dividend. Goulder (1994) distinguish between two types of double dividend a strong one and a weak one. There is a strong double dividend if there is a welfare gain from the tax reform (increased environmental taxation and reduced distortionary taxation) even excluding environmental benefits. There is a weak double dividend if the welfare gain from using green tax revenues to reduce existing distortionary taxes exceeds the welfare gain from lump sum redistribution of revenue.

There is a weak double dividend in ASTRA as the welfare gain from reduced income taxes (DKK 244 million per year) exceeds the gain from lump sum transfers (DKK 221 million per year). There is a welfare loss of DKK 321 million if environmental gains and congestion reductions are not taken into account. There is therefore no indication of a strong double dividend in the case with a toll ring and redistribution through reduced income taxes. In the same case we do not find evidence of a strong double dividend even if we exclude the administrative costs (the administrative costs of tax collection are often ignored in CGE analyses). Here we find a welfare loss of DKK 24 million (in the case of road pricing we find a loss of DKK 40 million).

Therefore the results supports the general learning from the literature that weak double dividends are common and strong double dividends are rare.

Sensitivity analyses and discussion

The empirical support for some of the elasticities in the demand system is relatively weak. In this section we discuss these weaknesses and quantify their importance.

A labour supply elasticity of 0.2 is assumed in the standard version of the model. This is composed of a working time elasticity of 0.1 and a participation elasticity of 0.1. It is, however, possible that the working time elasticity would not increase transportation correspondingly. This could be the case if the working time per day was increased or if the increased labour supply came from increased skills or increased effort.⁹ Changes in the consumer choice of place of residence could reduce transporta-

9. Note that taxes are reduced for consumers both inside and outside the working force. Consumers outside the working force will hardly increase their labour supply as a reaction to lower income taxes. Also for this reason a labour supply elasticity of 0.2 may exaggerate the expected effect on the labour supply.

Table 4. Welfare effects of a toll ring combined with reduced income taxes with alternative parameter specifications.

	Parameter value halved	Parameter value doubled
Welfare change, DKK millions		
Speed-flow function (dspeed/dflow)	29	564
Consumers' elasticity of substitution between:		
Transport – other consumption	210	322
Car use – non-car use	235	260
OD choice, commuting	140	561
OD choice, leisure	231	268
Peak hour travel – off peak travel	203	296

Source: Own calculations with ASTRA.

tion. Assuming a labour supply elasticity of 0.1 increases the welfare gain from a toll ring with reduced income taxes to DKK 290 million and does not change the ranking of the different types of regulation.

A wide range of other sensitivity analyses have also been carried out. Table 4 presents the calculated welfare gain with alternative parameter specifications.

It is clear from the table that changes in the assumed elasticities of substitution in the consumer demand system affect the resulting welfare changes from a toll ring. However, none of the alternative specifications change the overall result that there is a positive net benefit from a toll ring. This result also holds for significant changes in the speed-flow relationship. The welfare gain equals zero if the slope parameter in the speed flow relationship is 45% of the assumed level. As expected, greater flexibility in either demand or the speed-flow relationship would result in greater welfare gains from a toll ring.

The effects of regulation on the level of traffic depend on the chosen nesting structure in the demand system. The nesting structure in our model is similar to the structure used in the existing literature but is extended somewhat. This structure with very deep nests implies, however, that transportation types are generally complements. Other demand structures could result in greater substitution between types of transport. This could be an area for further work on the model.

The road tax levels are chosen on the basis of the Stockholm experiences. These, however, are not the levels that maximize welfare in the model. Model simulations indicate that the optimal toll is DKK 135 in peak and DKK 35 in off peak, which will generate a welfare gain of DKK 670 million per year. There are, however, costs that are not included in the model. A toll ring would create a more separated labour market with larger matching problems and less labour market competition. Commodity mar-

kets would also be more separated, possibly reducing competition. These negative effects would be expected to increase with increasing fees. Higher fees than used in tables 1-3 are therefore not necessarily appropriate even though a higher level is indicated by the model. Furthermore, the demand system in ASTRA is based upon nested CES functions. This implies that the results are mainly valid for minor changes in parameter values. The level of the optimal toll represents so large a change in parameter values that also for this reason, this result should not be viewed as valid.

Both road pricing and investments in infrastructure reduce congestion. Infrastructure investments reduce the marginal value of road pricing, and vice versa. This implies that the two instruments cannot be viewed separately but should be included simultaneously in a model analysis. This is theoretically possible in ASTRA, but it requires that the link between relevant potential infrastructure investments and the manner in which they affect the speed-flow relationship is known. This knowledge can be obtained with traffic models but requires several detailed analyses – analyses which have not yet been systematically carried out. The analyses made using ASTRA are for a fixed level of infrastructure. This means that the ASTRA results overestimate the potential gain from road pricing if the amount of infrastructure is less than the optimal, and underestimate the gain if the amount of infrastructure is greater than the optimal (De Borger and Wouters (1998)).

In the model it is assumed that household localisation is exogenous. This is not realistic in the long run. Larger welfare gains than found above must be expected with a long run endogenous household localisation as households will have improved possibilities of adapting to the new situation with a localisation near by their working place.

The modelling of regions is chosen to combine a focus on the Copenhagen area with availability of data. The chosen modelling of the regions therefore reflects the counties. It is a rough and not necessarily a suitable representation of the actual traffic pattern.

It is furthermore assumed that the length of a specific OD combination is exogenous. However, drivers may take detours to avoid passing the ring or entering a zone with high road pricing. This is not taken into account in the model. Inclusion of this would require a traffic model and is outside the scope of ASTRA. If these effects are significant they may reduce the benefits of toll rings (and road pricing).

The analysed road pricing system and toll ring is simplistic. More precise road pricing covering more than just one region or a system of several toll rings may increase the benefits.

Finally, firm's production functions are not modelled and general equilibrium effects are therefore excluded. A more detailed description of the effects on industries, trades and services and commodity markets would be an obvious next step. This de-

scription should also include the transport used by industries, trades and services. The exclusion of this type of transport probably reduces the welfare gain indicated by the model for both road pricing systems.

Supplementary analyses

In the long term, economic growth will increase the demand for transportation. Assume that economic growth increases labour productivity by 20% (modelled as an increase in the pre-tax wage of 20%). This increases the demand for transportation and consequently the need for regulation. Assume also that the fee at the toll ring increases by 20%. A toll ring combined with reduced income taxes will in this case give an annual welfare gain of DKK 372 million. Kilometre-based road pricing will now result in a welfare gain of DKK 50 million.

Technological development will probably reduce the costs of kilometre-based road pricing. The necessary GPS technology is already installed in several new cars and it is expected to become standard equipment in most cars in the future. The annual welfare gain from a kilometre-based system combined with income tax reductions would be DKK 361 million if it is assumed that the fixed costs of kilometre-based road pricing are halved.

It is clear from the above that the fixed costs associated with road pricing reduce the net benefit substantially. Increased petrol taxes could be an alternative that has practically no fixed costs. This instrument cannot be targeted towards congestion problems in specific areas, but it can reduce externalities in general and may create less distortion than income taxes.¹⁰ Analyses in Danish Economic Council (2006) indicate that the average taxation of cars is less than the marginal value of externalities in urban areas but greater in rural areas. Transport should be taxed both because of the externalities (a Pigou argument) and because of the public revenue (a Ramsey argument). Car transportation in urban areas is therefore under-taxed, but it is uncertain whether this is the case for rural areas.

Analyses with ASTRA indicate that an increase in the taxation of cars by DKK 0.50 per kilometre would result in a welfare loss of DKK 219 million per year if the revenue were used to reduce income taxes. Lower taxes would also produce a welfare loss, while a very small tax increase would result in a very small welfare gain. This indicates that the present tax level is close to optimal if a uniform kilometre tax is the only instrument.

We have also used the model to make cost effectiveness analyses of hypothetical infrastructure investments. It is possible to calculate the benefit from infrastructure investments that reduce average speed in a region by varying the a 's in the speed-flow

10. Cross-border trade in petrol might, however, reduce the gain from increased petrol taxes. This effect is not included in the model.

function. We do not know the costs of such infrastructure investments, and further work should be done before the model can be used for cost-benefit analyses of specific infrastructure investments. If an investment increases a by 1%, traffic will increase, and as a result the average speed will only increase by 0.9%. In total this will lead to an annual increase in welfare by DKK 701 million. Assuming that investment has a life-expectancy of 30 years and that the discount rate is 3%, there would be a positive welfare gain if it were possible to increase a by 1% by investments with a net present value of DKK 14 billion (DKK 11 billion with a rate of discount equal to 5%). This result can be compared with results from traffic models that calculate the effects from specific investments in infrastructure.

5. Conclusion and discussion

We have constructed a model for transport in Denmark with special focus on regulation of congestion and the derived labour market effects.

Our main finding is that there appears to be a welfare gain from a toll ring around Copenhagen. The welfare gain is highest if the revenue from the toll ring is used to reduce income taxes. This is because distortions in the labour market are reduced, which leads to increased labour supply. Alternative uses of the toll ring revenue such as increased income tax deduction for commuting costs or increased subsidies to public transport would also result in positive welfare effects, but at a lower level as compared to redistribution with lower income tax or lump sums.

A kilometre-based road pricing system produces a negative welfare effect in the model. The primary reason for this is the higher fixed costs associated with a kilometre-based road pricing system as compared to a toll ring.

We find that there is a weak double dividend in the taxation of congestion. That is, a redistribution of the revenue through a reduction of distortionary income taxes is better than redistribution through lump sum transfers. We do not find a strong double dividend as there is a welfare loss if the value of reduced environmental and congestion externalities is not taken into account.

The empirical basis for the estimation of parts of the demand system is somewhat uncertain, but a wide range of sensitivity analyses indicate that the above qualitative results are not contradicted if other levels of demand and congestion parameters are used.

The model could be extended in different ways. Freight transport (and other types of intermediate transport in the production of trades, goods and services) is not included in the model. Firms may also benefit from regulation of the congestion externality while consumers may benefit from the regulation of freight transport. Thus, inclusion

of firms (and their transport) is likely to increase the benefit of regulation (road pricing or toll ring).

The description of the transport network and transport behaviour is somewhat simplified. In real life car drivers may try to avoid a toll ring or road priced areas by making detours. This would reduce the expected benefits of both a toll ring and road pricing. In addition, we consider three (average) levels of congestion subject to the three zones in the model. In a model with more geographical areas (and more heterogeneous levels of congestion) it may be possible to fine tune regulation of congestion (e.g. to take into account the various levels of congestion within Copenhagen). Here road pricing appears to be a more flexible instrument and we would expect that welfare gains from road pricing regulation would be likely to be higher (toll rings could also be implemented more flexibly – with additional rings – but that would also increase the fixed costs).

Summarising, our results indicate that there are gains from regulation of congestion with a toll ring, but there are also a number of effects that must be investigated further, before a final conclusion can be reached. Over time increased traffic levels are likely to increase the gains from economic regulation of congestion and technological progress may also tend to favour real road pricing as compared to a toll ring.

Literature

- Arnott, R. 2005. City Tolls – One Element of an Effective Policy Cocktail. *CESifo DICE Report* 3/12, pp. 5-11.
- Bjørner, T. B. 1994. *Persontransport med bil*. AKF-memo. AKF Forlaget, København.
- Bjørner, T. B. 1997. En økonomisk model for bilejerskab og bilkørsel. *Nationaløkonomisk Tidsskrift*, 135 (3), pp. 233-50.
- Bjørner, T. B. og T. L. Lundhede. 2003. Prisen på stilhed 25 år senere. *Nationaløkonomisk Tidsskrift*, 141, pp. 279-99.
- Danish Economic Council. 2006. *Danish Economy, spring 2006*. Copenhagen.
- Danish Ministry of Transport. 2004. *External Costs of Transport. 2nd report – Marginal External Cost Matrices for Denmark*. Report conducted by COWI for the Danish Ministry of Transport, July 2004. Copenhagen.
- Danish Road Directorate. 2005. *Statsvejnettet – Oversigt over tilstand og udvikling. Report* no. 301. Copenhagen.
- De Borger, B., I. Mayeres, S. Proost and S. Wouters. 1996. Optimal Pricing of Urban Passenger Transport. *Journal of Transport Economics and Policy*, 30 (1), pp. 31-54.
- De Borger, B. and S. Wouters. 1998. Transport Externalities and Optimal Pricing and Supply Decisions in Urban Transportation: A Simulation Analysis for Belgium. *Regional Science and Urban Economics*, 28 (2), pp. 163-97.
- Fosgerau, M., M. Holmblad og N. Pilegaard. 2004. ART – En aggregeret prognosemodel for dansk vejtrafik. *Notat 5. Danmarks Transportforskning*. København.
- Frederiksen, A., E. K. Gravesen and N. Smith. 2001. Overtime Work, Dual Job Holding and Taxation. *IZA Discussion Paper* no. 323. The Institute for the Study of Labor, Bonn.
- Glaister, S. and D. Lewis. 1978. An Integrated Fares Policy for Transport in London. *Journal of Public Economics*. 9(3), pp. 341-55.
- Goulder, L. H. 1994. Environmental Taxation and the Double Dividend: A Reader's Guide. Paper presented at the 50th Congress of

- the International Institute of Public Finance. Harvard University, Cambridge MA, USA.
- Gørtz, M., J. V. Hansen and M. Larsen. 1999. CO₂-skatter, dobbelt dividende og konkurrence i energisektoren: Anvendelser af den danske AGL-model ECOSMEC. Det Økonomiske Råd. *Arbejdsrapport* 1999:1.
- Hauch, J. , T. B. Bjørner and N. Pilegaard. 2006. Transport Regulation Analysed in a Danish Equilibrium Model. The Secretariat of the Danish Economic Council. *Working paper* 2006:1.
- Knudsen M. B. L. H. Pedersen, T. W. Petersen, P. Stephensen and P. Trier. 1998. *Danish Rational Economic Agents Model – DREAM*, version 1.2. Copenhagen.
- Københavns Kommune. 2005. *Kørselsafgifter i København*. Copenhagen.
- Maddison, D., D. Pearce, O. Johansson, E. Calthrop, T. Litman and E. Verhoef. 1996. *The True Costs of Road Transport*. Earthscan, Newcastle.
- Munk, K. J. 2003. Computable General Equilibrium Models and Their Use for Transport Policy Analysis, *Report 4. Danmarks Transportforskning*, Lyngby.
- Newbery, D. M. 1990. Pricing and Congestion: Economic Principles Relevant to Pricing Roads. *Oxford Review of Economic Policy*, 6 (2), pp. 22-38.
- Nielsen, O.A. 2005. *Trængselsprojektet – Mødeløring af trængsel*. Center for Trafik og Transport, København.
- Parry, I. W. H. and A. Bento. 2001. Revenue Recycling and the Welfare Effects of Road Pricing. *Scandinavian Journal of Economics*, 103 (4), pp. 645-71.
- Parry, I. W. H. and K.A. Small. 2005. Does Britain or the United States Have the Right Gasoline Tax? *American Economic Review*, 95 (4), pp. 1276-89.
- Proost, S. and K. van Dender. 2001. The Welfare Impacts of Alternative Policies to Address Atmospheric Pollution in Urban Road Transport. *Regional Science and Urban Economics*, 31(4), pp. 383-411.
- Prud'homme, R. and J. Bocarejo. 2005. The London congestion charge: A tentative economic appraisal. *Transport Policy* 12, pp. 279-87.
- Transportrådet. 1999. *PETRA – Analysemodel for persontransport*. Notat nr. 6-99. Copenhagen.
- van Dender, K. 2003. Transport Taxes with Multiple Trip Purposes. *Scandinavian Journal of Economics*, 105 (2), pp. 295-310.
- Wrang, K., U. Nielsen and M. Kohl. 2006. *Kørselsafgifter i København – en samfundsøkonomisk analyse*. Report from the Environmental Assessment Institute (IMV), Copenhagen.

Bog anmeldelse

Torben Tranæs, red., Henrik Jacobsen Kleven, Claus Thustrup Kreiner, Niels-Kenneth Nielsen, Peder J. Pedersen: *Skat, arbejde og lighed – en undersøgelse af det danske skatte- og velfærdssystem*, Gyldendal 2006. Anmeldt af Jørgen Søndergaard.

Lad det være sagt med det samme: Denne bog fra Rockwool-fondens forskningsenhed indeholder mindst tre opsigtsvækkende ting.

Den første er, at Torben Tranæs i bogens indledningskapitel hævder, at kontanthjælpen udgør bundniveauet for lønninger på det danske arbejdsmarked, og at kontanthjælpen er på niveau med de politisk fastsatte mindstelønninger i USA. »Derfor kan mindstelønningerne i USA meget vel være en mindst lige så stor potentiel restriktion for beskæftigelse i USA som det sociale sikkerhedsnet er for beskæftigelsen i Danmark« (s. 26). Dette opsigtsvækkende udsagn underbygges gennem en sammenligning af beskæftigelsesfrekvenser for 25-54 årige mænd (der ikke sidder i fængsel). Forskellene er for tiden ganske rigtigt små. Og det får forfatteren til at konkludere, at en sænkning af sikkerhedsnettet heller ikke ville kunne inkludere særligt mange fra denne aldersgruppe (indvandrere undtaget).

Som læser kunne man godt ønske sig lidt flere »mellemregninger« i denne analyse: Kan der sættes lighedstegn mellem kontanthjælp og løngulv? Er værdien af fritid i gennemsnit nul (tilsyneladende ikke, for i samme kapitel hævdes det, at aktivering gennem at stjæle folks fritid er med til at øge beskæftigelsen!). Spiller det danske forhandlings-system slet ingen rolle – så vidt vides, er der ingen overenskomstaftalte lønninger på niveau med kontanthjælpen – de laveste overenskomstaftalte lønninger er snarere på niveau med dagpengene? Betyder det, at forfatteren

vurderer, at der findes et tilstrækkeligt udbud af jobs med lønninger på kontanthjælpsniveau i Danmark, således at de overenskomster, der dækker det store flertal af jobs på det danske arbejdsmarked, er uden betydning for eksklusion?

Den anden opsigtsvækkende ting er, at Niels-Kenneth Nielsen i to kapitler om indkomstfordelingen vælger at se bort fra ejerboliger og ejendomsværdibeskatning, »idet boligen for mange ikke er et investeringsobjekt« (s. 37). Negativ kapitalindkomst er følgelig ligeledes udeladt. Analysen fokuserer især på udviklingen i fordelingen af markedsinkomst og ser specifikt på udviklingen over en 20 årig periode i den andel af befolkning i aldersgruppen 25-59, der har en markedsinkomst under dagpengeniveau (eller under 60 pct. af medianindkomsten).

Også her kunne man ønske sig nogle »mellemregninger«. Hvor meget betyder det for de viste resultater, at ejerbolig og negativ kapitalindkomst er holdt uden for analysen. Er det et tilstrækkeligt argument, at folk ikke anser huskøb for at være en investering – har huskere så også et særligt forhold til friværdier, således at disse ikke betragtes som en del af forbrugsmulighederne? Forfatteren anfører herom, at forbrugsmuligheder baseret på stigende friværdier må anses for at være for skrøbelig en situation til, »at vi vil betragte personen som værende sikkert i havn som selvforsørgende« (s. 54).

Analyserne viser i øvrigt, at markedsinkomsterne blev mere ulige fordelt fra 1984 til 1994, hvorefter uligheden har været konstant. For ikke-vestlige indvandrere var stigningen særlig kraftig, og der er efter 1994 sket en vis forminskelse af uligheden. For de disponible indkomster er det lige omvendt: uligheden i indvandrerbefolkningen aftog frem til midten af 1990'erne, hvorefter den har været stigende igen. For den øvrige befolkning er der ingen klar trend i uligheden.

Peder J. Pedersen følger op ved at se på

indkomstmobiliteten blandt henholdsvis indvandrere og den øvrige befolkning og finder, at indvandrere ikke blot har en højere andel i lavindkomstgruppen, men at deres sandsynlighed for stige op i indkomstfordelingen er blevet mindre over tid sammenlignet med den øvrige befolkning, mens omvendt risikoen for at falde ned i indkomstfordelingen, for de, der ligger højere oppe, er øget for indvandrerne i forhold til den øvrige befolkningsrisiko.

Den tredje opsigtsvækkende ting er, at Kleven og Kreiner når frem til, at en marginal sænkning af marginalsattesatserne i den øvre del af indkomstfordelingen meget vel kan være fuldt selvfinansierende. De har kort sagt forsøgt at bestemme Lafferkurven for Danmark og fundet, at nogle af sattesatserne meget vel kan ligge på den gale side af Lafferkurvens toppunkt.

Krumtappen i sådanne analyser er naturligvis skattebasens elasticitet mht. sattesatserne – hvor meget vil skattegrundlaget vokse, hvis vi gennemfører en marginal sænkning af en sattesats eller ændrer marginalt på indkomstgrænsen for opkrævning af f.eks. topskat? Skattebasen kan vokse af to årsager: ved at flere deltager på arbejdsmarkedet, når en sattesats sænkes, og ved at de, der er på arbejdsmarkedet i forvejen, arbejder mere (eller bedre), så deres indkomst øges. Ved skatteændringer i toppen af indkomstfordelingen er det især den sidste effekt, der betyder noget. Der kommer ikke mange flere ind på arbejdsmarkedet ved en marginal sænkning af topskatten (det kan potentielt få nogle unge til at afslutte uddannelsen lidt hurtigere og nogle ældre til at udskyde tilbagetrækningen, men marginalsatten har begrænset effekt på sådanne beslutninger). Deltagelseseffekten kan til gengæld være væsentlig ved f.eks. ændringen i beskæftigelsesfradrag eller andre tiltag, der påvirker den effektive kompensationsgrad (efter skat) i overførselssystemet.

Problemet ved beregninger af selvfinansieringen er, at der ikke foreligger ret mange analyser af hverken den samlede elasticitet for skattebasen eller de to elementer i arbejds-

udbudselasticiteten. Forfatterne gennemfører ikke nye estimationer, men vælger elasticitetsparametrene på grundlag af foreliggende danske og udenlandske empiriske undersøgelser. De gør omhyggeligt rede for deres – i øvrigt ret forsigtige – valg.

Det opsigtsvækkende resultat, at lettelser i toppen af skattesystemet kan være selvfinansierende, bygger på en timeelasticitet (stigning i arbejdsindsats blandt de allerede beskæftigede) på 0,10.

Forfatterne illustrerer følsomheden i Lafferkurven over for denne elasticitet. Sættes den i stedet til 0,15 er skattelettelser til alle – bortset fra de beskæftigede i nederste kvartil – selvfinansierende. Sættes elasticiteten i stedet til 0,05 findes der ikke selvfinansierende skattelettelse. Her får man resultater, der ligner de beregninger, Finansministeriet og Det Økonomiske Råd tidligere har præsenteret.

I betragtning af hvor meget skattelettelser fylder i kampen om den politiske dagsorden, er det ikke nogen komfortabel situation for økonomistanden at være i. Antallet af milliarder, der skal findes til at finansiere en lempelse af topskatten (eller andre reduktioner i skatten på arbejde), afhænger af brøkdelen af en promise i arbejdsudbudselasticiteten. Det siger sig selv, at politikere vil foretrække elasticiteter i den høje ende (jf. de eksperimenter med selvfinansierende skattelettelser, der har været prøvet i andre lande). Med al respekt for det forskningsarbejde, der er gjort for at raffinere estimationerne af arbejdsudbuddet, er der vist ingen, der på det nuværende faglige grundlag med nogen grad af sikkerhed kan afgøre, om denne helt afgørende elasticitet er større eller mindre end 0,1 endside hvor meget.

En potentiel bias i den litteratur Kleven og Kreiner benytter er, at estimerede elasticiteter, der ikke er signifikant forskellige fra nul, måske ikke altid når frem til tidsskrifterne. Så én implikation af denne bogs analyser må være, at arbejdsudbudsforskerne må se at få turbo på yderligere estimationer – selv de mindste nuancer i resultaterne kan få vidtrækkende samfundsøkonomiske konsekvenser.

Hvem er målgruppen for denne bog? Forfatterne gør sig betydelige pædagogiske anstrengelser, men bogens stof er vist ikke for den forudsætningsløse læser. Dens publikum skal nok hovedsageligt findes blandt økonomer og økonomi-studerende. De får til gen-

gæld en uhyre velskrevet bog, der behandler sit stof med stort overblik og (i den sidste del) giver os et velvalgt ekstrakt af indsigterne fra de sidste 30 års internationale forskning i public finance. Den slags bøger kunne vi godt bruge nogle flere af.

Indkomst og dødelighed i en dansk befolkningsstikprøve

Kristian Schultz Hansen

E-mail: ksh@soci.au.dk

Ulrika Enemark

E-mail: ue@soci.au.dk

Anders Foldspang

Afdeling for Sundhedstjenesteforskning, Institut for Folkesundhed, Aarhus Universitet,

E-mail: af@mph.au.dk

SUMMARY: The objective of the present study was to estimate the association between income and mortality in a random sample of Danish citizens. This article investigates the validity of five theories on the association between income and mortality, by testing whether these theories could find support in analyses based on Danish data. From the adult population aged 18-64 of Central and Northern Jutland, a random sample of 102.251 individuals was drawn as of 1 January 2001. For each of these persons, relevant information, such as age, sex, place of living (municipality), family size and gross family income for the calendar year 2000, was extracted from administrative registers. Logistic regression analyses were performed for males and females separately, and a dichotomous variable (death occurring during 2003, yes/no) was used as the dependent variable, while age and various income measures served as explanatory variables. In accordance with the Absolute-Income Hypothesis, the analyses suggest a significant, negative association of the 1-year mortality risk with income. Furthermore, the present analyses lend some support both to the Relative-Position Hypothesis and the Deprivation Hypothesis. Moreover, neither the Relative-Income Hypothesis, nor the Income-Inequality Hypothesis could be supported.

Indledning

Indkomst og sundhed

Der foreligger en stor og stadig voksende litteratur vedrørende sammenhængen

Cand.stat, Ph.D., Morten Frydenberg, cand.merc.(mat), Ph.D., Lars Peter Østerdal, cand.polit, Ph.D., Kim Moesgaard Iburg samt to anonyme referees takkes for kommentarer til artiklen. Projektet er støttet af Statens Sundhedsvidenskabelige Forskningsråd (2008-00-0013) og Aarhus Universitets Forskningsfond (S-2005-SUN1-63).

mellem personers sundhed og indkomst. Den empiriske del af litteraturen finder ofte en stærk positiv association mellem individuel indkomst og høj sundhedsstatus, hvor sidstnævnte i mange studier repræsenteres ved et mål for dødelighed eller overlevelse. Empirien tyder med andre ord på, at dødsrisikoen på individniveau er faldende for stigende personlig indkomst, og omvendt at der er stigende overlevelsessandsynlighed med voksende indkomst. Der er tale om en relativt robust sammenhæng, der kan findes for den samlede dødelighed og for en række specifikke dødsårsager og desuden i forskellige lande og over tid, Lynch m.fl. (2004). Tilsvarende viser studier på mere aggregeret niveau, at der også ofte er en positiv sammenhæng mellem den gennemsnitlige indkomst og sundhed på landeniveau, Deaton (2003), Pritchett og Summers (1996). Et tilsvarende mønster kan findes i Danmark, Osler m.fl. (2002), Osler og Klebak (1998), idet der dog er en tendens til svagere association mellem indkomst og dødsrisiko hos kvinder end hos mænd, Munch og Svarer (2005).

Der hersker en vis uenighed om hvordan denne sammenhæng mellem indkomst og dødelighed opstår, og der er derfor foreslået en række forskellige mekanismer som forklaring. Et udbredt forslag omhandler de materielle muligheder, idet forbrug af visse goder kan forventes at have sundhedsfremmende effekt – såsom sund mad, motion, afstressende ferier, sund bolig og et passende forbrug af sundhedsydelser, Marmot (2002). Muligheden for at forbruge sådanne sundhedsfremmende goder vil afhænge af den enkeltes eller familiens finansielle ressourcer, såfremt der er en positiv sammenhæng mellem kvalitet og pris for de nævnte goder.

Ifølge en anden type forklaring – af psykologisk og sociologisk tilsnit – har borgerne i de rigeste lande opnået en vis materiel standard, således at der er brug for at identificere supplerende mekanismer til forklaring af den negative sammenhæng mellem indkomst og dødelighed, Marmot (2002). Det foreslås, at indkomsten – udover at give visse forbrugsmuligheder – også er en indikator for succes og dermed for den enkeltes sociale position. I forlængelse heraf påpeger flere studier, at følelsen af at have en lav position i samfundets hierarki er forbundet med dårligere helbred sammenlignet med personer med en højere position, Lynch m.fl. (2004). Jo højere indkomst, en person har opnået, jo mere succesfuld føler personen sig, og dette har igen positiv indflydelse på den personlige sundhed.

I modsætning til de to nævnte forklaringer, hvor indkomst formodes at skabe sundhed, er økonomer ofte mere tilbøjelige til at forestille sig den modsatte retning i forklaringskæden – nemlig at sunde personer skaber højere indkomst til sig selv. Eksempelvis betragtes sundhed for den enkelte person i Grossman-modellen som en beholdning af kapital, der kan opbygges eller vedligeholdes gennem investering ved regelmæssige besøg hos sundhedsvæsenets udbydere og gennem indtagelse af sund kost og ved at dyrke motion, Grossman (1972), Zweifel og Breyer (1997, kap. 3). Grossman's model

antager desuden, at der er en positiv sammenhæng mellem beholdningen af sundhed og indtjeningsvevnen, således at forbedret sundhed fører til stigende indkomst. Endelig kan der være tale om, at bagvedliggende faktorer bestemmer både indkomst og sundhedstilstand i samme retning, Fuchs (2004), Deaton (2003). Eksempelvis kan uddannelsesniveau påvirke sundheden positivt, samtidig med at den personlige indkomst stiger med uddannelsesvarigheden.

Det er oplagt, at man blandt en gruppes medlemmer både kan finde eksempler, hvor helbredet har bestemt indkomstniveauet, og hvor det modsatte gælder. Alligevel er den dominerende opfattelse blandt forskerne inden for emnet, at kausaliteten overvejende går fra indkomst til sundhed. Denne overbevisning stammer især fra studier af fødselskohorter, der er fulgt gennem længere tid, hvorved det har været muligt at observere, om indkomst eller sundhed kom først i årsagskæden. Det konkluderes ofte, at der er relativt få medlemmer af kohorten, hvor dårligt helbred kan konstateres at have ført til lav indkomst, således at dette kun i begrænset omfang kan forklare den ret generelle sammenhæng, som den positive association mellem indkomst og sundhed er, Marmot og Wilkinson (1999). I modsætning til dette finder en amerikansk undersøgelse blandt 70+ årige ingen association mellem udvalgte socioøkonomiske indikatorer inklusive indkomst og dødsrisiko efter kontrol for forudgående sygelighed, Adams m.fl. (2003). Dette leder forfatterne til at konkludere, at der ikke er en direkte kausal sammenhæng fra socioøkonomiske indikatorer til dødsrisiko. Det kan dog ikke afvises, at socioøkonomiske forhold kan have en indirekte indflydelse på dødsrisiko, da f.eks. indkomsten kan påvirke sygeligheden i foregående periode, Adda m.fl. (2003), Adams m.fl. (2003).

Den udbredte enighed om, at stigende indkomst har en stærk, negativ indflydelse på dødsrisikoen, forhindrer dog ikke en livlig debat om hvilke mekanismer, der kan medføre denne sammenhæng. Der findes en række forskellige teorier for sammenhængen mellem indkomst og sundhed/dødelighed, Wagstaff og van Doorslaer (2000), Lynch m.fl. (2004), hvoraf de fem mest anvendte kort beskrives nedenfor.

Teorier om indkomst og sundhed

Ifølge den *absolutte indkomst hypotese (AIH)* er det alene en persons absolutte indkomstniveau, der – alt andet lige – er bestemmende for en persons dødsrisiko, Wagstaff og van Doorslaer (2000), Gravelle (1998), således at den individuelle dødsrisiko falder med stigende indkomst. Dette er i overensstemmelse med en forklaring baseret på forbedrede muligheder for forbrug af sundhedsfremmende goder. Det antages endvidere, at den positive effekt af indkomstniveau bliver mindre og mindre ved stigende indkomst, således at den marginale reduktion i dødsrisiko vil være højere på et lavt indkomstniveau end på et højt indkomstniveau, Gravelle (1998). I modsætning til de efterfølgende hypoteser påvirkes en persons dødsrisiko ifølge AIH udelukkende af

egen indkomst og ikke eksempelvis af andres indkomst eller af andre forhold i det samfund, man er del af.

En negativ sammenhæng mellem stigende indkomst og dødsrisiko kan findes både i de rigeste lande og i lande med vækst i gennemsnitlig indkomst over en lang årrække, Fuchs (2004). Dette har ført til hypoteser om, at det absolutte indkomstniveau ikke er tilstrækkeligt til at forklare sammenhængen mellem indkomst og sundhed, Marmot m.fl. (1991), Wilkinson (1997). Den *relative indkomst hypotese (RIH)* foreslår derfor, at også en persons relative position i samfundet – målt ved den personlige indkomst – har indvirkning på helbredet og dermed på dødsrisikoen. Det antages, at en lav relativ personlig indkomst medfører øget individuel dødsrisiko, mens det omvendt er forbundet med bedre helbred at have en høj indkomst i forhold til andre.

Ifølge et forslag til en mekanisme, der kan frembringe en sådan sammenhæng, fører lav indkomst hos en person til en følelse af at tilhøre de nederste lag i samfundets hierarki, hvilket i sig selv kan medføre stress, usikkerhed, følelse af manglende kontrol med egen situation og tab af selvværd, Marmot m.fl. (1991). En del studier har sandsynliggjort, at eksempelvis stress kan igangsætte fysiologiske processer med skadelig indflydelse på helbredet, se Brunner and Marmot (1999). Endvidere viser andre studier, at følelsen af manglende kontrol med egen situation på arbejdspladsen har negativ indflydelse på selvvurderet sundhed, Borg og Kristensen (2000). En person med høj indkomst, med kontrol over egen situation og med høj social status, vil derimod ikke opleve tilsvarende negativ indflydelse på helbredet. RIH indebærer eksempelvis, at en person med uændret absolut indkomst over en periode kan opleve en ændring i dødsrisiko, nemlig hvis indkomsterne i samfundet ændrer sig generelt (ændret gennemsnitsindkomst).

En anden implikation af hypotesen er, at det vil være forbundet med dalende sundhedsniveau, hvis en person med en bestemt absolut indkomst flytter fra et område med lav gennemsnitsindkomst til et område med højere gennemsnitsindkomst. I et moderne, globaliseret samfund med livlig netværksdannelse kan det imidlertid være vanskeligt at afgøre, hvilken referencegruppe, der er vigtigst for en person. Der kan være tale om f.eks. arbejdskammerater, naboer, omgangskreds, beboere i nærområdet, personer fra TV, og kombinationer af disse, hvilket gør det vanskeligt at teste hypotesen empirisk, Deaton (2003), Lynch m.fl. (2004).

Endelig vil fordelingen af indkomsterne i samfundet (eller blandt en gruppe mennesker) ifølge *indkomstulighedshypotesen (IUH)* have betydning for den enkelte persons dødsrisiko, Lochner m.fl. (2001). En stigende indkomstulighed i en befolkningsgruppe forventes at medføre en højere dødsrisiko for alle medlemmer af gruppen, således at der er tale om en kollektiv og ikke blot en individuel effekt. Ifølge et forslag til en mekanisme hænger dette sammen med, at høj indkomstulighed i befolkningen i et om-

råde fører til lav social sammenhængskraft eller social kapital defineret som de karakteristika ved den sociale organisation (såsom tillid til andre, deltagelse i det civile liv o.s.v.), der muliggør et samarbejde til det fælles bedste, Lynch og Kaplan (1997). Det antages endvidere, at der er en positiv sammenhæng mellem niveauet for social kapital i et område og personernes sundhed og modsvarende en negativ sammenhæng mellem social kapital og individuel dødsrisiko. Ifølge et andet forslag til en mekanisme, der kan skabe en sammenhæng mellem indkomstulighed og sundhedsniveau, vil områder med høj ulighed ofte have særlige karakteristika hvad angår levekår, kultur og politisk liv. Det kan således tænkes – og nogle steder observeres – at områder med høj grad af accept af ulighed i mindre grad investerer i brede befolkningsgrupper gennem uddannelse, social- og sundhedsprogrammer, godt boligbyggeri, miljøbeskyttelse og andre programmer, der kan antages at have positiv indflydelse på sundheden, Lynch og Kaplan (1997), Kawachi m.fl. (1997). Indkomstulighedshypotesen forudsiger, at både rige og fattige personer i et område påvirkes negativt af en høj grad af indkomstulighed. En implikation af hypotesen er derfor, at en omfordeling af indkomst fra rige til fattige personer isoleret set vil medføre en forbedring af sundheden og dermed reduktion af dødsrisikoen for rige såvel som fattige, Mellor og Milyo (2002).

Som nævnt tidligere forestiller man sig under den absolutte indkomst hypotese, at effekten af ekstra indkomst på dødsrisikoen bliver mindre jo højere en persons indkomst er. Denne tankegang er trukket en smule skarpere op i en hypotese, som vi her vil betegne *lavindkomstgræsehypotesen (LIH)*,¹ der foreslår, at blot en person har opnået en indkomst over et vist kritisk niveau, så vil sundhedsniveauet være ens uanset indkomstniveau, Wagstaff og van Doorslaer (2000). Dødsrisikoen vil derfor alt andet lige være den samme for alle personer med indkomster over dette kritiske indkomstniveau. Et sådant kritisk niveau for indkomsten kan defineres både absolut og relativt. Personer med indkomst under lavindkomstgræsen vil ifølge hypotesen opleve en højere dødsrisiko end personer med indkomst over dette niveau. I øvrigt antages det for personer med indkomster under lavindkomstgræsen, at sammenhængen mellem indkomst og dødelighed er den samme som forudsat under den absolutte indkomsthypotese, d.v.s. faldende dødsrisiko for stigende indkomst, samt at den marginale reduktion i dødsrisiko vil være højere på et lavt indkomstniveau end på et højt indkomstniveau, Wagstaff og van Doorslaer (2000).

Som det sidste forslag til sammenhæng mellem indkomst og dødelighed er det ifølge den *relative positions hypotese (RPH)* ikke kun den absolutte indkomst, der har betydning, men også en persons relative placering i indkomstfordelingen, Wagstaff og van Doorslaer (2000). Det antages, at en højere plads i rangfølgen af indkomster er forbundet med bedre sundhed og lavere dødsrisiko end en lavere plads. Den enkeltes

1. Engelsk: Deprivation Hypothesis.

indkomst i rangfølgen af alle indkomster er en indikator for social position. I lighed med den relative indkomst hypotese spiller den sociale position en vigtig rolle for den relative positions hypotese, men de to teorier adskiller sig på forskellige måder. Under RIH har størrelsen af forskellen mellem en persons faktiske indkomst og sammenligningsindkomsten betydning for sundheden, og en stigning i en persons indkomst vil derfor forbedre sundheden. For RPH gælder imidlertid, at en stigning i personlig indkomst kun vil medføre et forbedret sundhedsniveau, hvis ændringen i indkomst også medfører en ændring i personens relative position i indkomstfordelingen. Eksempelvis vil en generel stigning i indkomsterne for samfundets rigeste personer eller et generelt fald i indkomsterne for de fattigste medføre ændringer i dødsrisikoen for disse grupper under RIH, mens dette ikke vil være tilfældet ifølge RPH, eftersom den relative position for de berørte grupper forbliver den samme.

Formål

Formålet med nærværende artikel er at estimere sammenhængen mellem indkomst og dødelighed i en dansk stikprøve under antagelse af, at årsagsretningen helt overvejende går fra indkomst til sundhed og dødelighed. Estimeringen gennemføres med undersøgelse af, hvilke af de beskrevne fem teorier om sammenhængen mellem indkomst og dødelighed, der kan finde støtte i danske data. Som følge af tilstedeværelsen af en række valide offentlige registre er der i Danmark ret unikke muligheder for at danne store datasæt med relevante oplysninger til empirisk vurdering af de fem beskrevne indkomstteorier, hvoraf AIH og IUH tidligere er testet i mange versioner og i mange lande, RIH kun et fåtal af gange, mens vi ikke har fundet offentliggjorte empiriske afprøvninger af LIH og RPH.

Materiale og metoder

Fra en database opstartet i 2004 med henblik på at kunne udføre forskning vedrørende sundhedstjenesten og forbruget af sundhedsydelser i befolkningen i Århus, Viborg, Ringkøbing og Nordjyllands amter er identificeret en række oplysninger, der kan anvendes til at analysere de fem nævnte hypoteser. Materialet i nærværende studie omfatter en 10% systematisk fødselsdatabaseret stikprøve af befolkningen i de nævnte amter udtaget pr. 1. januar 2001, i alt 52.154 mænd og 50.097 kvinder i alderen 18-64 år (tabel 1). Den summariske dødsrate for befolkningen i de fire amter adskiller sig ikke væsentligt fra hele landet, idet denne var henholdsvis 11,0 og 10,5 per 1000 personår i 2001, udregnet på basis af tal fra Danmarks Statistik (2002). Tilsvarende var den gennemsnitlige personindkomst i de fire amter kun lidt lavere end i landet som helhed, nemlig henholdsvis 204.500 og 212.300 kroner, Danmarks Statistik (2003). Da studiet er registerbaseret, er der ikke noget personbortfald.

Tabel 1. Stikprøvestørrelse og dødelighed (2003) fordelt på køn i en 10% stikprøve af befolkningen i Århus, Viborg, Ringkøbing og Nordjyllands amter, 2001.

Stikprøvestørrelse; dødsfald	Mænd	Kvinder
Personer i stikprøven pr. 1. januar 2001	52154	50097
Dødsfald i 2003	238	150

Baseret på en række befolkningsregistre foreligger individuelt koblede oplysninger om stikprøvestørrelsernes geografiske, demografiske og sociale karakteristika, blandt andet fødselsdato, dødsdato, familiestørrelse, bopælskommune og familiens bruttoindkomst. Personerne i stikprøven er fordelt på 88 bopælskommuner, hvoraf der er 118 personer i den mindste kommune og 19.255 personer i den største. I alt 31 kommuner har mindre end 500 personer i stikprøven og 0-7 dødsfald i 2003.

Med henblik på at undersøge de fem nævnte hypoteser vedrørende indkomst og sundhed er der udtrukket og udledt en række variable fra stikprøven. Som afhængig variabel og dermed som mål for sundhed anvendes en dikotom variabel, der antager værdien 1 for en person, der er død i løbet af kalenderåret 2003, og ellers 0. Indkomstindikatorer er udregnet for kalenderåret 2000 – og ikke 2001 eller 2002 – for at modvirke en eventuel bias som følge af, at indkomsten kan være unormalt lav umiddelbart før død (i kalenderåret 2003) som følge af sygdom. Personer, der er afgået ved døden inden 2003, d.v.s. i løbet af kalenderårene 2001 og 2002, er derfor udeladt af analysen svarende til i alt 688 personer. Derimod er fraflyttede i kalenderårene 2001 og 2002 ikke udeladt af analysen, idet databasen indeholder information om alle stikprøvestørrelsernes dødsdato op til slutningen af 2004 uanset bopæl. De anvendte indkomstmål afhænger af hypotesen og er nærmere beskrevet nedenfor. Endelig er alle analyser korrigeret for alder pr. 1. januar 2001.

Den statistiske analyse er foretaget ved hjælp af logistisk regression (Hosmer og Lemeshow 2000), og den betingede sandsynlighed for, at person i dør i løbet af 2003, kan derfor specificeres ved:

$$P_i(Død_{2003} | A_i, Z_i) = \frac{e^{\alpha + \beta A_i + \gamma Z_i}}{1 + e^{\alpha + \beta A_i + \gamma Z_i}} \quad (1)$$

– der kan transformeres til:

$$\ln \frac{P_i}{1 - P_i} = \alpha + \beta A_i + \gamma Z_i \quad (2)$$

Sandsynligheden P_i for at person i afgår ved døden i løbet af 2003 forklares ved alder A_i og indkomst Z_i , hvor sidstnævnte er en vektor af forklarende indkomstvariable,

som varierer mellem hypoteserne, mens α , β og γ er tilhørende koefficienter. Der testes generelt på et 5% signifikansniveau. Modeltilpasning er testet med anvendelse af Hosmer's og Lemeshow's χ^2 test, Hosmer og Lemeshow (2000).

Anvendte indkomstmål

Som mål for den absolutte indkomst Y_i for person i er her valgt at anvende den ækvivalente bruttoindkomst, der er defineret som en families samlede bruttoindkomst divideret med antal familiemedlemmer, hvor hver person dog er givet forskellige vægte (1 for første voksne, 0,7 for næste voksne og 0,5 for hvert barn). Da der er relativt få personer med ækvivalente bruttoindkomster over 600.000 kroner, hvoraf nogle af disse indkomster er meget afvigende, anvendes medianindkomsten for denne gruppe personer i stedet for deres faktiske indkomster. Medianindkomsten er for mænd i denne gruppe 716.500 kroner og 716.200 kroner for kvinder. På tilsvarende måde beregnes medianindkomsten for personer med indkomster under 0, således at -57.400 kroner anvendes for denne gruppe mænd og -42.400 kroner for kvinder.

Et første indblik i sammenhængen mellem ækvivalent bruttoindkomst og dødelighed (uden korrektion for alder) fås af figur 1, der er fremkommet ved først at ordne alle ækvivalente bruttoindkomster i stigende rækkefølge, dernæst inddele stikprøvememberne i et antal lige store grupper, og for disse grupper er beregnet dødsproportion og medianindkomst.

Indkomstdelen af regressionsligningen (1) under AIH kan skrives:

$$\gamma Z_i = \gamma_1 Y_i \quad (\text{Model 1})$$

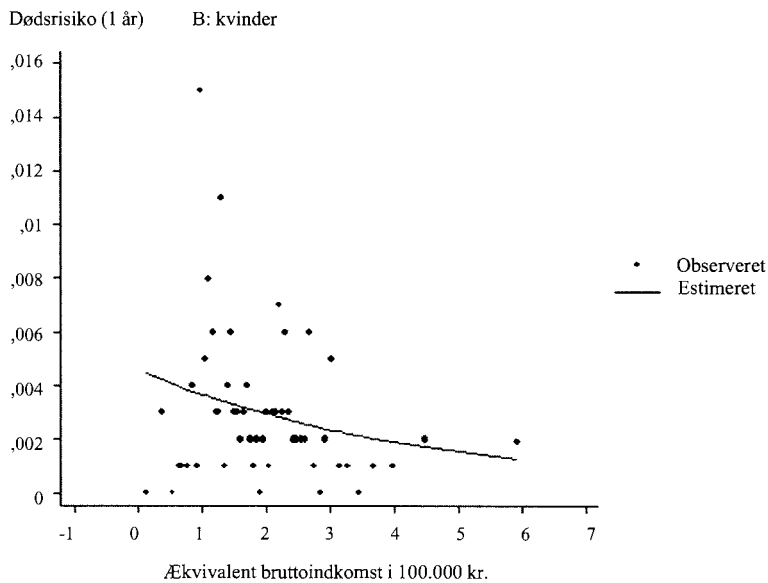
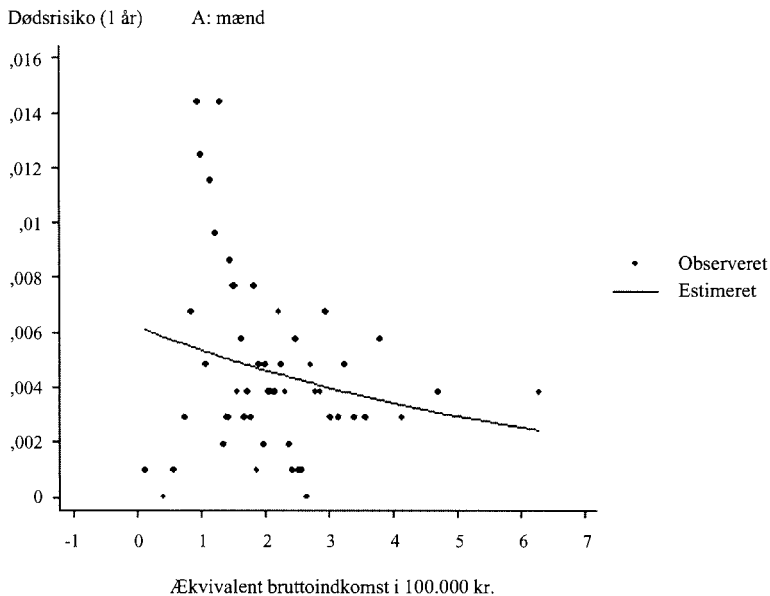
– hvor der gælder, at $\gamma_1 < 0$ ifølge hypotesen.

Nøgleproblemet i en test af RIH er at definere indkomstniveauet for den gruppe mennesker, der er reference for person i 's sociale status. I denne artikel har vi valgt at anvende gennemsnitsindkomsten i lokalsamfundet – nærmere bestemt bopælskommunen for den enkelte person. Der er derfor beregnet gennemsnit af ækvivalente bruttoindkomster \bar{Y}_i for personer, der bor i samme kommune. Størrelsen \bar{Y}_i varierer derfor efter, hvilken kommune person i bor i. I regressionsligningen (1) under hypotesen RIH indgår derfor:

$$\gamma Z_i = \gamma_1 Y_i + \gamma_2 \bar{Y}_i \quad (\text{Model 2})$$

– og det forventes, at $\gamma_2 > 0$ ifølge hypotesen.

Indkomstuligheden er operationaliseret for IUH ved at beregne Gini koefficienten G_i på baggrund af de ækvivalente bruttoindkomster for personer fordelt på bopæls-



Figur 1. Observerede og estimerede døds sandsynligheder (2003)[#] som funktion af ækvivalent bruttoindkomst i en 10% stikprøve af befolkningen i Århus, Viborg, Ringkøbing og Nordjyllands amter (2001).

[#] Baseret på logistisk regression uden alderskorrektion og defineret ved medianindkomst. Modeltilpasning: $p < 0,01$ for mænd og $p < 0,01$ for kvinder.

kommune. Størrelsen G_i afhænger derfor af, hvilken kommune person i bor i. Indkomstkomponten i regressionsligningen (1) kan derfor formuleres under hypotesen IUH:

$$\gamma Z_i = \gamma_1 Y_i + \gamma_3 G_i \quad (\text{Model 3A})$$

– og hvor hypotesen tilsiger, at $\gamma_3 > 0$.

Som et andet hyppigt anvendt mål for ulighed kan beregnes andelen L_i af de samlede ækvivalente bruttoindkomster, som indtjenes af kommunens 50% fattigste personer, hvor stigende L_i angiver faldende indkomstulighed i en kommune. Indkomstkomponten kan derfor skrives:

$$\gamma Z_i = \gamma_1 Y_i + \gamma_4 L_i \quad (\text{Model 3B})$$

– med forventning om, at $\gamma_4 < 0$ hvis IUH holder.

For at undersøge LIH har vi konstrueret en dikotom variabel F_i , som antager værdien 1, hvis person i har en ækvivalent bruttoindkomst over en kritisk indkomst Y^L på 119.000 kroner svarende til 60% af medianindkomsten i datasættet, mens værdien ellers er 0. Denne beregningsmåde anvendes ofte til at fastsætte en relativ lavindkomstgrænse eller en relativ fattigdomsgrænse, Keilman m.fl. (1997), Finansministeriet (2004). For personer med indkomst under lavindkomstgrænsen ($F_i = 0$) antages det, at dødsrisikoen er faldende for stigende indkomst i lighed med den absolutte indkomsthypotese. I modsætning hertil gælder for personer med indkomst over lavindkomstgrænsen ($F_i = 1$), at der ingen sammenhæng er mellem indkomst og dødsrisiko, d.v.s. at der er ens dødelighed for alle indkomstniveauer. Endelig antages det, at personer med indkomst over lavindkomstgrænsen ikke kan have højere dødsrisiko end personer med indkomst under denne grænse. Den laveste dødsrisiko for personer med indkomst under lavindkomstgrænsen opnås ifølge hypotesen netop ved grænseværdien Y^L , og dødsrisikoen for personer med indkomster over denne kritiske indkomstværdi kan derfor ikke overstige dødsrisikoen for indkomstniveau Y^L . Disse specifikke sammenhænge kan undersøges ved følgende indkomstkomponten i ligningen (1):

$$\gamma Z_i = \gamma_5 (Y_i - Y^L)(1 - F_i) + \gamma_6 (Y_i - Y^L) F_i + \gamma_7 F_i \quad (\text{Model 4})$$

hvor de enkelte parametre γ fortolkes på følgende måde: Hvis en person har en indkomst under lavindkomstgrænsen, således at $F_i = 0$, reduceres modellen til $\gamma Z_i = \gamma_5 Y_i - \gamma_5 Y^L$, hvorefter γ_5 angiver den marginale effekt af indkomst på dødelighed for personer med indkomster under lavindkomstgrænsen. Hvis en person i stedet har en

indkomst over lavindkomstgrænsen og dermed $F_i = 1$, reduceres Model 4 til $\gamma Z_i = \gamma_6 Y_i - \gamma_6 Y^L + \gamma_7$, således at γ_6 måler den marginale effekt af indkomst på dødelighed for personer med indkomster over lavindkomstgrænsen. Endelig kan en person have en indkomst præcis på lavindkomstgrænsen, d.v.s. at $Y_i = Y^L$ og $F_i = 1$, hvilket ved indsætning reducerer modellen til $\gamma Z_i = \gamma_7$, hvorved γ_7 måler et eventuelt spring i dødelighed ved den kritiske indkomstværdi. Såfremt LIH holder, vil følgende restriktioner på parametrene derfor gælde: $\gamma_5 < 0$, $\gamma_6 = 0$ og $\gamma_7 \leq 0$.

En indikator for den relative position er opnået ved først at ordne alle ækvivalente bruttoindkomster i stigende rækkefølge inden for hver kommune, hvorefter hvert indkomstniveau har fået tildelt et nummer mellem 1 og N_k (antal af personer i kommune k) svarende til placeringen i rækkefølgen. Inden for hver kommune er variabelen R_i dannet ved at normalisere disse numre til en skala mellem 0 og 100, hvor værdien 100 angiver den højeste position. Variabelen R_i angiver således den relative position af person i 's indkomst i indkomstrækkefølgen i person i 's bopælskommune, hvorefter modellen kan skrives:

$$\gamma Z_i = \gamma_1 Y_i + \gamma_8 R_i \quad (\text{Model 5})$$

– med forventet parameterværdi $\gamma_8 < 0$.

En beskrivelse af de anvendte forklarende variable findes i Tabel 2. Stikprøvens mandlige medlemmer havde i gennemsnit en signifikant højere ækvivalent bruttoindkomst i 2000 end kvinderne. Den gennemsnitlige ækvivalente bruttoindkomst på kommuneniveau varierede relativt lidt, idet gennemsnittet i den rigeste kommune kun var 44.000 kroner højere end i den fattigste kommune. Indkomstuligheden inden for kommuner målt på de to ulighedsmål G_i og L_i var generelt lav, og der var forholdsvis lille forskel på de mindst og de mest ulige kommuner (mindre end 10 procentpoint). Endvidere havde 82% af mændene og 79% af kvinderne i stikprøven en ækvivalent bruttoindkomst over lavindkomstgrænsen på 119.000 kroner. Endelig var mænds relative indkomstposition i kommunen højere end kvinders, hvilket afspejlede mændenes højere gennemsnitlige ækvivalente bruttoindkomst.

I overensstemmelse med foranstående gennemføres alle regressionsanalyser med den ækvivalente bruttoindkomst og alder som prædiktorer og separat for mænd og kvinder. Ved testning af hypoteser, der er mere komplicerede end AIH, er der desuden inddraget yderligere prædiktorer som specificeret ovenfor i modellerne 2-5.

Endelig er der udført en række følsomhedsanalyser. For det første er dødsfald en relativ sjælden hændelse i stikprøven, så de fem indkomstteorier er også testet med to alternative afhængige variable; død 2003-2004 og 2002-2004 (henholdsvis 751 og 1092 dødsfald). For det andet er gennemført regressionsanalyser med anvendelse af

Tabel 3. Prædiktions af dødelighed (2003)[#] ved hjælp af forskellige indkomstindikatorer opgjort i 100.000 kroner (2000), i en 10% stikprøve af befolkningen i Århus, Viborg, Ringkøbing og Nordjyllands amter (2001).

Prædiktorer	Mænd		Kvinder	
	OR	95% K.I	OR	95% K.I
<i>Absolut indkomst hypotese (Model 1)</i>				
Y_i – Ækvivalent bruttoindkomst	0,66 (a)	0,57 – 0,75	0,63 (a)	0,53 – 0,76
<i>Relativ indkomst hypotese (Model 2)</i>				
Y_i – Ækvivalent bruttoindkomst	0,65 (a)	0,57 – 0,75	0,64 (a)	0,53 – 0,76
\bar{Y}_i – Kommune gennemsnitsindkomst	1,07	0,23 – 4,89	0,32	0,04 – 2,52
<i>Indkomstulighedshypotesen (Model 3)</i>				
<i>Version A</i>				
Y_i – Ækvivalent bruttoindkomst	0,66 (a)	0,57 – 0,75	0,63 (a)	0,53 – 0,76
G_i – Gini koefficient i kommunen	1,00	0,95 – 1,05	0,99	0,93 – 1,06
<i>Version B</i>				
Y_i – Ækvivalent bruttoindkomst	0,66 (a)	0,57 – 0,75	0,63 (a)	0,53 – 0,76
L_i – Andel optjent af 50% fattigste	1,00	0,94 – 1,07	1,00	0,92 – 1,09
<i>Lavindkomstgrænsehypotesen (Model 4)</i>				
Y_i lavere end Y^L	6,13 (a)	1,40 – 26,77	3,60	0,69 – 18,78
Y_i højere end Y^L	0,71 (b)	0,60 – 0,84	0,65 (b)	0,51 – 0,82
F_i – Dikotom variabel for lavindkomst*	0,37 (a)	0,23 – 0,59	0,54 (a)	0,31 – 0,96
<i>Relativ positions hypotesen (Model 5)</i>				
Y_i – Ækvivalent bruttoindkomst	1,08	0,82 – 1,42	0,88	0,54 – 1,40
R_i – Relativ position i kommunen	0,98 (a)	0,97 – 0,99	0,99	0,97 – 1,01

Noter: # Prædiktions ved hjælp af multipel logistisk regression, med justering for alder i fyldte år. Koefficienter og odds ratioer knyttet til alder vises ikke i tabellen.

* Variablen antager værdien 1 hvis den ækvivalente bruttoindkomst > 119.000 kr. og 0 ellers

(a) $p < 0,05$ hvor $p = P(|U| > u) / 2$ og $U \sim N(0,1)$

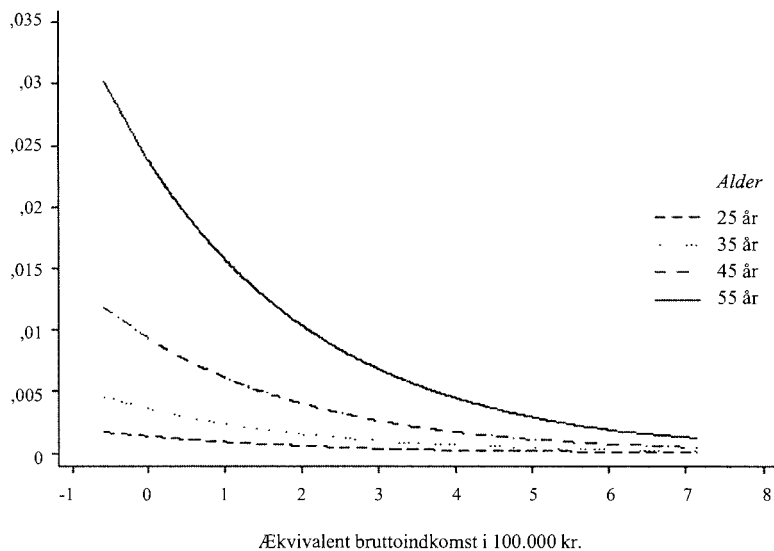
(b) $p < 0,05$ hvor $p = P(|U| > u)$ og $U \sim N(0,1)$

(2007) giver – med en enkelt undtagelse (se senere) – ikke anledning til at ændre på konklusionerne i det følgende.

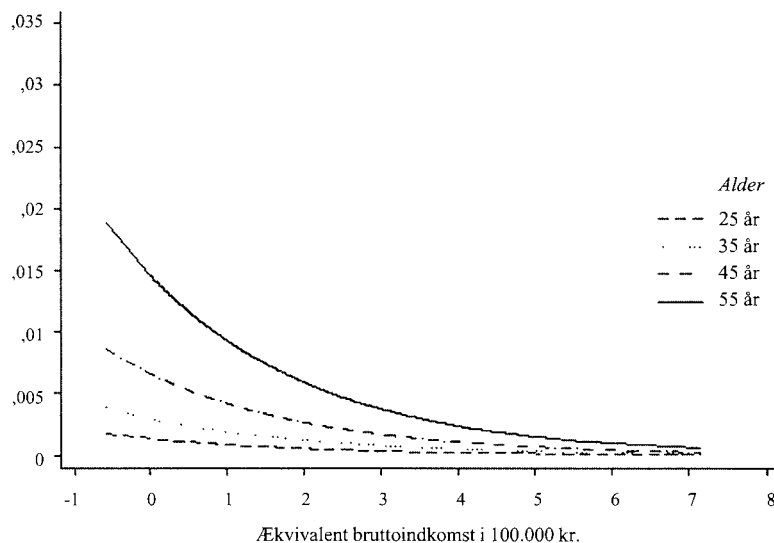
Den absolutte indkomst hypotese (AIH)

I den estimerede model for den absolutte indkomst er odds ratio for den ækvivalente bruttoindkomst signifikant lavere end 1 for både mænd og kvinder. Stigende indkomst er således associeret med faldende dødelighed i overensstemmelse med den absolutte indkomst hypotese. Mere præcist udtrykker de estimerede parametre, at en

Dødsrisiko (1 år) A: mænd



Dødsrisiko (1 år) B: kvinder



Figur 2. Estimerede dødssandsynligheder (2003)[#] som funktion af ækvivalent bruttoindkomst (2000) og alder i en 10% stikprøve af befolkningen i Århus, Viborg, Ringkøbing og Nordjyllands amter (2001).

[#] Baseret på logistisk regression. Modeltilpasning: $p > 0,60$ for mænd og $p > 0,05$ for kvinder.

stigning i indkomst på 100.000 kroner er forbundet med en dødsrisiko, som er 66% af risikoen for død før stigningen i indkomst for mænd og 63% for kvinder. Anvendes modellens estimerede parametre til at forudsige sammenhængen mellem indkomst og dødsrisiko for mænd og kvinder, fremkommer figur 2. Den 1-årige dødsrisiko falder med stigende indkomst, men den marginale reduktion i risiko er lavere jo højere indkomst, som også foreslået af AIH.

Den relative indkomst hypotese (RIH)

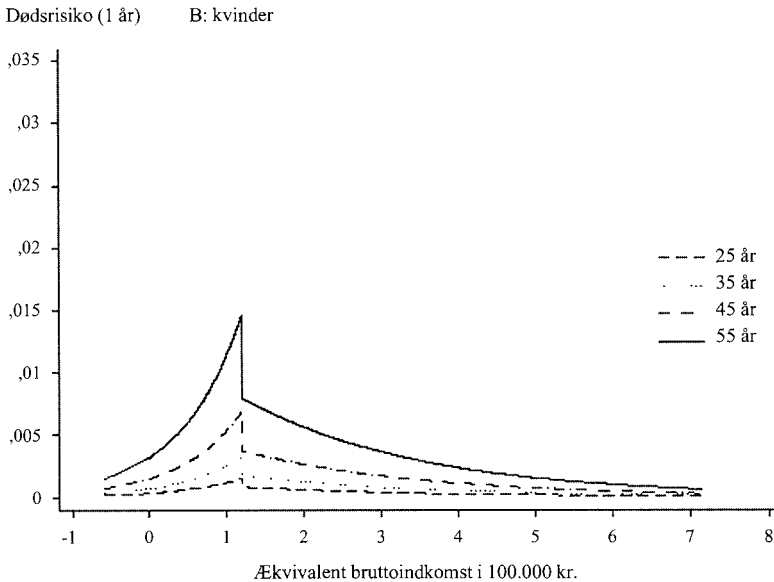
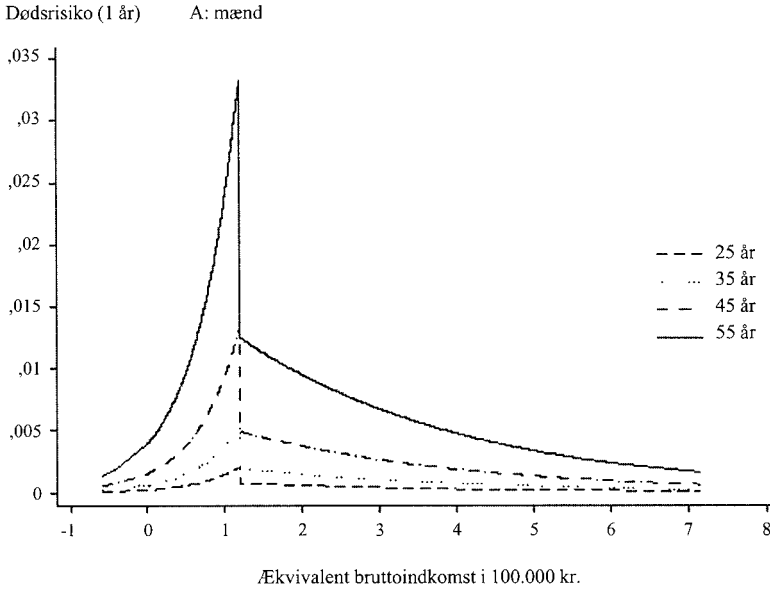
Model 2 tester om en persons relative indkomst indvirker på dødsrisikoen ud over effekten af den ækvivalente bruttoindkomst. Den relative indkomsthypotese undersøges i Model 2 ved anvendelse af den gennemsnitlige indkomst i kommuner, og OR er over 1 for mænd og under 1 for kvinder (tabel 3). For kvinderne er en stigende gennemsnitsindkomst i kommunen (svarende til faldende egen relativ indkomst) derfor associeret med en faldende dødsrisiko, hvilket er i modstrid med RIH. OR for gennemsnitsindkomst er dog insignifikant både for mænd og kvinder.

Indkomstulighedshypotesen (IUH)

Model 3A og 3B forsøger at teste indkomstulighedshypotesen ved brug af to ofte anvendte ulighedsmål som forklarende variable. OR for Gini koefficienten for en persons bopælskommune er omkring 1 for både mænd og kvinder, således at højere ulighed i en persons bopælskommune ikke er forbundet med højere dødsrisiko, hvilket er i modstrid med det forventelige ifølge indkomstulighedshypotesen. Nogenlunde de samme konklusioner kan knyttes til det andet ulighedsmål L_i , hvor en lav værdi er udtryk for stor ulighed og omvendt. Ifølge tabel 3 er OR omkring 1 for begge køn således, at andelen af de samlede indkomster optjent af de 50% fattigste ikke påvirker dødsrisikoen. Dette er i modstrid med IUH.

Lavindkomstgrænsehypotesen (LIH)

Det ses af Model 4 i tabel 3, at OR for ækvivalent indkomst under lavindkomstgrænsen Y^L er større end 1 for begge køn (men dog kun signifikant for mænd), hvilket ikke er forventeligt ifølge hypotesen. Endvidere er OR for indkomst højere end Y^L signifikant lavere end 1 for både mænd og kvinder, hvorimod hypotesen forudsiger, at OR skal være netop lig 1. Endelig er OR for et eventuelt spring i dødsrisiko ved lavindkomstgrænsen Y^L som forventet ifølge hypotesen d.v.s. mindre end eller lig med 1. Samlet set finder lavindkomstgrænsehypotesen således ikke støtte i datasættet ved den valgte kritiske indkomstgrænse på 119.000 kr. Det ses da også af kurverne over de estimerede dødsrisikoer (figur 3), at der er en voksende association mellem indkomst



Figur 3. Estimerede dødssandsynligheder (2003)[#] som funktion af ækvivalent bruttoindkomst (2000), bruttoindkomst under kritisk værdi (119.000 kr.; dikotom variabel), og alder, i en 10% stikprøve af befolkningen i Århus, Viborg, Ringkøbing og Nordjyllands amter (2001).

[#] Baseret på logistisk regression. Modeltilpasning: $p > 0,65$ for mænd og $p < 0,01$ for kvinder.

og dødelighed for personer med indkomst under 119.000 kr., hvilket er det modsatte af det forventelige ifølge LIH.

Den relative positions hypotese (RPH)

Den relative position er associeret med dødsrisikoen ifølge tabel 3. For både mænd og kvinder er OR for denne variabel mindre end 1, således at en stadig højere placering i indkomsthierarkiet er forbundet med lavere og lavere dødsrisiko som forudsagt af den relative positions hypotese. Denne association er dog kun signifikant for mænd. Sammenhængen for mænd mellem relativ position og dødsrisiko er stærk, idet variabelen R_i overtager forklaringskraft fra den ækvivalente indkomst Y_i . Hvis der foretages en justering af standardafvigelserne for estimerne med henblik på at tage hensyn til mulige korrelationer mellem beboere fra samme kommune, Stock og Watson (2007) medfører dette, at OR bliver signifikant mindre end 1 også for kvinder.

Følsomhedsanalyser

Anvendelse af en dikotom variabel for død med to eller tre års observationstid som afhængig variabel giver ikke anledning til kvalitative ændringer i konklusionerne tidligere draget på basis af tabel 3 undtagen vedrørende RPH, som ikke kan understøttes for mænd og ikke kan afvises for kvinder, når 3-års dødeligheden anvendes som afhængig variabel.

Tilføjelse af de tre socioøkonomiske indikatorer som forklarende variable inkluderet enkeltvis har som hovedregel ingen indflydelse på de hidtidige kvalitative konklusioner. Hvis eksempelvis de tre socioøkonomiske indikatorer inkluderes samtidig til prædiktion af 3-års dødeligheden blandt kvinder, overtager disse variable forklaringskraften fra indkomstvariablen under AIH.

Sammenhængen mellem dødsrisiko og ækvivalent bruttoindkomst kan beskrives bedre ved et polynomium af 2., 3. og 4. grad i indkomst målt ved et Hosmer-Lemeshow test for mænd end et førstegrads polynomium, mens dette ikke er tilfældet for kvinder.

Endelig giver udeladelse af de mindste kommuner, anvendelse af alternative indkomstulighedsmål eller udskiftning af den ækvivalente bruttoindkomst med personlig bruttoindkomst ikke anledning til at ændre på konklusionerne af de oprindelige analyser.

Diskussion

Den ækvivalente indkomst anvendt som mål for den absolutte indkomst er en særdeles god prædiktor for dødsrisiko i stikprøven ifølge de logistiske regressionsanalyser præsenteret i tabel 3 (Model 1). Denne sammenhæng, inklusive den faldende marginale effekt af indkomst, er i overensstemmelse med resultaterne i talrige andre

studier, f.eks. Backlund m.fl. (1996), McDonough m.fl. (1997), Fiscella og Franks (2000), Gerdtham og Johannesson (2004). Dog peger følsomhedsanalyserne på, at for mænd er sammenhængen mellem indkomst og dødsrisiko ikke negativ for alle indkomstintervaller, og at et polynomium i indkomst er en rimelig beskrivelse. Denne sammenhæng kan især forklares ved to forhold. For det første indeholder datasættet et antal studerende og selvstændige med meget lave indkomster og lav dødelighed samt personer under revalidering eller på kontanthjælp med højere indkomst og højere dødelighed. Samlet set medvirker disse grupper derfor til at skabe en tendens til stigende dødsrisiko for stigende indkomst blandt personer med indkomster under ca. 100.000 kroner. For det andet er antallet af personer med ækvivalent indkomst over 450.000 kroner relativt lille, og antallet falder hurtigt med stigende indkomst. Begrænses regressionsanalysen for mænd til indkomstintervallet 50.000-400.000 kroner, er der negativ sammenhæng mellem dødelighed og indkomst samt en faldende marginal reduktion i dødsrisiko for stigende indkomst som foreslået af AIH. Endvidere giver et polynomium i 2., 3. og 4. grad i indkomst for det nævnte indkomstinterval ikke en bedre beskrivelse end et førstegradspolynomium målt med et Hosmer-Lemeshow test.

RPH er ikke testet empirisk i andre lande, Wagstaff og van Doorslaer (2000) men understøttes for mænd af den danske stikprøve. AIH og RPH er imidlertid tæt forbundne, idet den absolutte indkomst ordner personer i samme rækkefølge (om end ikke med samme afstand mellem personerne) som positionsvariablen. På baggrund af observationen af at position overtager forklaringsgraden fra absolut indkomst i Model 5, må det konkluderes, at der er en sammenhæng mellem position og risiko for død. Men da absolut indkomst og position er associerede, og absolut indkomst samtidig ikke har nogen signifikant association med dødelighed i Model 5, kan det ikke vurderes, hvor meget ekstra forklaring, der kan tilskrives relativ position.

Den internationale litteratur indeholder kun få empiriske undersøgelser af RIH. Gerdtham og Johannesson (2004) anvender den gennemsnitlige indkomst i kommunen sammen med personens ækvivalente disponible indkomst inklusiv formueafkast til at teste hypotesen og konkluderer, at den relative indkomst ikke har nogen signifikant indflydelse på dødsrisikoen for et givent individuelt indkomstniveau i Sverige. I modsætning hertil finder Elstad m.fl. (2006) støtte til RIH i Norge, idet den relative indkomst målt som den procentvise afvigelse mellem en persons disponible indkomst og kommunens medianindkomst udviser en signifikant, negativ sammenhæng med dødsrisiko inden for de fleste indkomstgrupper. Det er dog usikkert, om den empiriske vurdering i den danske stikprøve kan begrunde en generel afvisning af RIHs relevans i Danmark. For det første anvendes i analyserne det kommunale indkomstgennemsnit som referencepunkt svarende til en hypotese om, at den enkelte person sammenligner sin situation med kommunens beboere som helhed, når egen sociale status vurderes.

Dette referencepunkt er muligvis for aggregeret – især i store kommuner. Et muligt alternativ i det givne datasæt kunne være at sammenligne med gennemsnitsindkomsten for personer med samme uddannelse eller type beskæftigelse. For det andet er der en forholdsvis begrænset variation i kommunernes gennemsnitsindkomst (tabel 2). Hvis alle personer i stikprøven relaterer sig til stort set samme kommunale gennemsnitsindkomst, vil det være vanskeligt at undersøge hypotesen om, at det er forbundet med højere dødsrisiko for en person med en given egenindkomst at bo i et område med høj gennemsnitsindkomst sammenlignet med et område med lav gennemsnitsindkomst. For det tredje lægger tidsforskydningen i variablene en begrænsning på den sammenhæng, som det er muligt at teste. Den ækvivalente bruttoindkomst og dermed også gennemsnitsindkomsten i kommunerne er udregnet for kalenderåret 2000, bopæl opgøres per 1. januar 2001, og status for død eller ikke død opgøres i 2003. Personerne i stikprøven kan dermed nå at flytte mellem kommuner med forskelligt gennemsnitligt indkomstniveau i perioden 2001-2003 og dermed gøre det svært at måle den umiddelbare effekt af et lavt eller højt kommunalt indkomstniveau. Det, der er testet her, er som nævnt sammenhængen mellem indkomst i 2000 og dødeligheden i 2003.

Der er betydelig uenighed i litteraturen om den teoretiske sammenhæng mellem ulighed og sundhed, og IUH har da også været flittigt debatteret og undersøgt, Daly m.fl. (1998). På individniveau har nogle forfattere, Lochner m.fl. (2001), Soobader og Leclere (1999), Kaplan m.fl. (1996), Dahl m.fl. (2006) empirisk fundet en voksende sammenhæng mellem ulighed og morbiditet/mortalitet, mens andre ikke finder nogen signifikant sammenhæng, Fiscella og Franks (1997), Daly m.fl. (1998), Osler m.fl. (2002), Gerdtham og Johannesson (2004). Enkelte empiriske studier understøtter UIH ved at påvise en positiv sammenhæng mellem sundhed og social kapital målt eksempelvis som graden af befolkningens deltagelse i valg, Islam m.fl. (2006) eller deltagelse i det civile liv, Kawachi m.fl. (1997), Mellor og Milyo (2005). I nærværende datasæt er der kun en lille variation i de fire anvendte ulighedsmål fra kommune til kommune, hvilket kan medvirke til at forklare, at IUH ikke finder støtte i den udførte analyse. Til sammenligning varierede Gini koefficienten i kommuner mellem 12 og 51% i et svensk studie, Gerdtham og Johannesson (2004) mod 24-34% i vort danske datasæt (24-30% hvis den mest ulige kommune udelades). Desuden kan der også her være et problem i forbindelse med bopælsskift mellem kommuner i perioden 2001-2003. Hvis bopælsskift foregår mellem kommuner med varierende ulighed, kan dette udvande eller forstærke en eventuel effekt af ulighed i datamaterialet.

LIH bliver ikke understøttet i nærværende studie med den valgte lavindkomstgrænse Y^L på 119.000 kroner, idet der er en voksende sammenhæng mellem indkomst og dødelighed for personer under lavindkomsten som det ses af figur 3 (dog kun signifikant for mænd). Dette paradoks i forhold til LIH kan imidlertid forklares med, at nær-

værende datasæt indeholder socioøkonomiske grupper med en gennemsnitsindkomst på langt under lavindkomstgrænsen kombineret med en meget lav dødsrisiko. Eksempler på disse er uddannelsessøgende samt et antal af stikprøvens selvstændige, der har meget lave og endda negative indkomster i 2000. Hos disse grupper indikerer ækvivalent bruttoindkomst ikke forbrugsmulighed. Samtidig findes andre grupper i stikprøven med højere dødelighed såsom personer på kontanthjælp eller under revalidering med en gennemsnitsindkomst lige omkring lavindkomstgrænsen, hvilket er betragteligt højere end indkomsten hos de uddannelsessøgende og de nævnte selvstændige. Hvis alle de nævnte socioøkonomiske grupper udelades blandt mænd (1974 personer), er der da heller ikke nogen signifikant voksende sammenhæng mellem indkomst og dødsrisiko for mænd under lavindkomsten på 119.000 kroner.

Det kan endvidere tænkes, at den kritiske indkomstgrænse Y^L er sat for lavt eller for højt. Hvis Model 4 estimeres med forskellige lavindkomstgrænser Y^L , kan LIH ikke afvises, når lavindkomstgrænsen for mænd er 200.000 kroner eller højere og 240.000 kroner eller højere for kvinder.

Endelig kan det overvejes, om der bør anvendes en testprocedure for hypotesen, hvor de tre restriktioner på parametrene γ_5 , γ_6 og γ_7 (og dermed også OR) skal overholdes simultant i modsætning til den anvendte procedure med tre uafhængige test. En sådan analyse ville medføre større sandsynlighed for afvisning af LIH.

I fremtidige regressionsanalyser kan det være relevant at teste de fem hypoteser ved hjælp af paneldata, hvor det kan forsøges at kontrollere for personernes generelle helbreds niveau i foregående år, hvilket vil kunne medvirke til at besvare spørgsmålet vedrørende kausalitet i sammenhængen mellem indkomst og dødelighed. Individuelt generelt helbreds niveau kan tænkes indiceret ved hjælp af oplysninger fra kontakt med sundhedsvæsenet, f.eks. sygehusdiagnoser.

Analyserne peger alt i alt på en markant og robust sammenhæng mellem indkomst og dødsrisiko. Hvis det kan sandsynliggøres, at resultaterne indikerer en kausal sammenhæng, vil en implikation være, at dødsrisikoen både på individniveau og samfunds niveau kan formindskes ved forskellige politikker, som påvirker indkomsterne og fordelingen af disse i persongrupper. Eksempelvis vil enhver stigning i indkomsten hos en eller flere personer medføre en formindskelse af dødsrisikoen for disse ifølge AIH. En yderligere implikation af denne hypotese er, at en omfordeling af indkomsterne i et samfund – bestående i, at personer med given indkomst overlader et beløb til personer med lavere indkomst – vil medføre en samlet nedgang i dødsrisikoen for samfundet som helhed. Dette følger af antagelsen om den faldende marginale reduktion i dødsrisiko for voksende indkomstniveau. Tilsvarende implicerer LIH, at omfordeling af indkomst fra personer med indkomster højere end den kritiske indkomstværdi til personer med indkomst under denne samlet set vil formindske dødsrisikoen i samfundet.

Man skal naturligvis være varsom med at designe politikker som de ovennævnte på det foreliggende ikke-eksperimentelle (observationelle) grundlag. Konkrete kausale sammenhænge er således ikke afdækket. Eksempelvis kan det for nogle grupper gælde, at indkomsten ikke er den vigtigste faktor til forklaring af niveauet for dødsrisikoen, men snarere at indkomsten korrelerer med andre forhold (arbejdsledelse, arbejdsmiljø, familiestruktur, uddannelse, sundhedsadfærd), der har mere direkte indflydelse på dødsrisikoen. For sådanne personer vil en overførsel af penge ikke nødvendigvis føre til lavere dødsrisiko, hvis de øvrige forhold forbliver uændrede. Som et yderligere eksempel synes det ikke at være nogen god ide at omfordele mere indkomst til studerende ud fra et ønske om reducere deres dødsrisiko, da denne allerede er meget lille.

De udførte analyser kan også bruges til at identificere risikogrupper med høj dødelighed, der så kan udvælges som målgruppe med henblik på særlig forebyggende og sundhedsfremmende indsats. Det første kriterium for udvælgelse kan så være indkomst, især de laveste indkomster (AIH), eller indkomster under en kritisk værdi (LIH). Et scenarium baseret på denne brug af prædiktionen kan være, at man i stedet for at henvende sig til hele befolkningen med f.eks. en adfærdsændrende kampagne kan nøjes med at målrette indsatsen til nogle relativt få udvalgte familier og enkeltpersoner.

Konklusion

I denne artikel har vi anvendt en stikprøve på mere end 100.000 beboere i den nordlige halvdel af Jylland til at vurdere validiteten af 5 forskellige teorier for sammenhængen på individniveau mellem indkomst og sundhed udtrykt ved 1-års dødsrisikoen. Resultaterne af den anvendte metode kan opsummeres ved at fremhæve især den absolute indkomst hypotese og den relative positions hypotese som teorier, der opnår stærk støtte i datamaterialet, mens også lavindkomstgræsehypotesen understøttes til en vis grad. Derimod finder hverken den relative indkomst hypotese eller indkomstulighedshypotesen støtte i den valgte form.

Litteratur

- Adams, P., M. D. Hurd, D. McFadden, A. Merrill og T. Ribeiro. 2003. Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics* 112, s. 3-56.
- Adda, J., T. Chandolal og M. Marmot. 2003. Socio-economic status and health: causality and pathways. *Journal of Econometrics* 112, s. 57-63.
- Backlund, E., P. D. Sorlie og N. J. Johnson. 2004. The shape of the relationship between income and mortality in the United States. Evidence from the National Longitudinal Mortality Study. *Annals of Epidemiology* 6, s. 12-20.
- Borg, V. og T. S. Kristensen. 2000. Social class and self-rated health: can the gradient be explained by differences in life style or work environment? *Social Science and*

- Medicine* 51, s. 1019-30.
- Brunner, E. og M. Marmot. 1999. Social organization, stress, and health, i Marmot, M. og R. G. Wilkinson, red. *Social determinants of health*, Oxford: Oxford University Press.
- Danmarks Statistik. 2002. *Statistisk Årbog 2002*, København: Danmarks Statistik.
- Danmarks Statistik. 2003. *Statistisk Årbog 2003*, København: Danmarks Statistik.
- Dahl, E., J. I. Elstad, D. Hofoss og M. Martin-Mollard. 2006. For whom is inequality most harmful? A multi-level analysis of income inequality and mortality in Norway. *Social Science and Medicine* 63, s. 2562-74.
- Daly, M. C., G. J. Duncan, G. A. Kaplan og J. W. Lynch. 1998. Macro-to-micro links in the relation between income inequality and mortality. *The Milbank Quarterly* 76, s. 315-39.
- Deaton, A. 2003. Health, inequality, and economic development. *Journal of Economic Literature* 41, s. 113-58.
- Elstad, J. I., E. Dahl, og D. Hofoss. 2006. Associations between relative income and mortality in Norway: a register-based study. *European Journal of Public Health* 16, s. 640-44.
- Finansministeriet. 2004. *Fordeling og incitamentter 2004*, Albertslund: Schultz Information.
- Fiscella, K., og P. Franks. 1997. Poverty or income inequality as predictor of mortality: longitudinal cohort study. *British Medical Journal* 314, s. 1724-27.
- Fiscella, K. og P. Franks. 2000. Individual income, income inequality health, and mortality: what are the relationships. *Health Services Research* 35, s. 307-18.
- Fuchs, V. R. 2004. Reflections on the socioeconomic correlates of health. *Journal of Health Economics* 23, s. 653-61.
- Gerdtham, U. G. og M. Johannesson. 2004. Absolute income, relative income, income inequality, and mortality. *Journal of Human Resources* 39, s. 228-47.
- Gravelle, H. 1998. How much of the relation between population mortality and unequal distribution of income is a statistical artefact? *British Medical Journal* 316, s. 382-85.
- Grossman, M. 1972. On the concept of health capital and the demand for health. *Journal of Political Economy* 80, s. 223-55.
- Hosmer, D. W. og S. Lemeshow. 2000. *Applied logistic regression*, New York: John Wiley & Sons, Inc.
- Islam, K. M., J. Merlo, I. Kawachi, M. Lindström, K. Burström og U. G. Gerdtham. 2006. Does it really matter where you live? A panel data multilevel analysis of Swedish municipality-level social capital on individual health-related quality of life. *Health Economics, Policy and Law* 1, s. 209-35.
- Juul, S. 2004. *Epidemiologi og evidens*, København: Munksgaard.
- Kaplan, G. A., E. R. Pamuk, L. W. Lynch, R. D. Cohen og J. L. Balfour. 1996. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *British Medical Journal* 312, s. 999-1003.
- Kawachi, I., B. P. Kennedy, K. Lochner og D. Prothrow-Stith. 1997. Social capital, income inequality, and mortality. *American Journal of Public Health* 87, s. 1491-98.
- Keilman, N., J. Lyngstad, H. Bojer og I. Thomsen, red. 1997. *Poverty and economic inequality in industrialized Western countries*, Oslo: Scandinavian University Press.
- Lochner, K., E. Pamuk, D. Makuc, B. Kennedy og I. Kawachi. 2001. State-level income inequality and individual mortality risk: a prospective, multilevel study. *American Journal of Public Health* 91, s. 385-91.
- Lynch, J., G. Davey Smith, S. Harper, M. Hillemeier, N. Ross, G. A. Kaplan og M. Wolfson. 2004. Is income inequality a determinant of population health? A systematic review. *The Milbank Quarterly* 82, s. 5-99.
- Lynch, J. W. og G. A. Kaplan. 1997. Understanding how inequality in the distribution of income affects health. *Journal of Health Psychology* 2, s. 297-314.
- Marmot, M. 2002. The influence of income on

- health: views of an epidemiologist. *Health Affairs* 21, s. 31-46.
- Marmot, M. G., G. Davey Smith, S. Stansfield, C. Patel, F. North, J. Head, I. White, E. Brunner og A. Feeny. 1991. Health inequalities among British civil servants: the Whitehall II Study. *Lancet* 337, s. 1387-93.
- Marmot, M. og R. G. Wilkinson, red. 1999. *Social determinants of health*, New York: Oxford University Press.
- McDonough, P., G. J. Duncan, D. Williams og J. House. 1997. Income dynamics and adult mortality in the United States, 1972 through 1989. *American Journal of Public Health* 87, s. 1476-83.
- Mellor, J. og J. Milyo. 2002. Income inequality and health status in the United States: evidence from the current population study. *Journal of Human Resources* 37, s. 510-39.
- Mellor, J. M. og J. Milyo. 2005. State social capital and individual health status. *Journal of Health Politics, Policy & Law* 30, s. 1101-30.
- Munch, J. R. og M. Svarer. 2005. Mortality and socio-economic differences in Denmark: a competing risks proportional hazard model. *Economics and Human Biology* 3, s. 17-32.
- Pritchett, L. H. og L. H. Summers. 1996. Wealthier is healthier. *Journal of Human Resources* 31, s. 841-68.
- Osler, M. og S. Klebak. 1998. Social differences in health in an affluent Danish county. *Scandinavian Journal of Social Medicine* 26, s. 289-92.
- Osler, M., E. Prescott, M. Grønbaek, U. Christensen, P. Due og G. Engholm. 2002. Income inequality, individual income, and mortality in Danish adults: analysis of pooled data from two cohort studies. *British Medical Journal* 324, s. 1-4.
- Soobader, M. J. og F. Leclere. 1999. Aggregation and the measurement of income inequality: effects on morbidity. *Social Science and Medicine* 48, s. 733-44.
- Stock, J. H. og M. W. Watson. 2007. *Introduction to econometrics*, Boston: Pearson Education, Inc.
- Wagstaff, A. og E. van Doorslaer. 2000. Income inequality and health: what does the literature tell us? *Annual Review of Public Health* 21, s. 543-67.
- Wilkinson, R. G. 1997. Socioeconomic determinants of health: health inequalities: relative or absolute material standards. *British Medical Journal* 314, s. 591-94.
- Zweifel, P. og F. Breyer. 1997. *Health economics*, New York: Oxford University Press.

Velfærdsvægte for forskellige indkomstgrupper i Danmark

Liv Bangsbo Petersen

Arbejdsmarkedsstyrelsen, E-mail: lpe@ams.dk

SUMMARY: This paper presents a new simple heuristic method of calculating revealed welfare weights for different income groups in Denmark. The welfare weights are estimated under the assumption that the existing Danish tax- and welfare system is optimal. The method is based on a model in Social Marginal Cost of Public Funds theory, which both incorporate hours-of-work, labour supply – and participation responses. The revealed welfare weights compare to objective standard criteria for acceptable weights; (1) decreasing in income and (2) non negative. The calculation of the Danish weights shows that in the presence of positive participation responses, the middle income groups have the largest priority at the expense of the other income groups (the welfare weights profile is inverse U-shaped). Furthermore the results showed that in different cases the welfare weights for the high-income groups become negative. Thus, the results contradict the two objective criteria for reasonable social preferences – or it contribute to create uncertainty about the optimality of the Danish tax-system.

1. Indledning

I de senere år har der været stor fokus på det danske skattesystem og konsekvensen af den måde, hvorpå det er opbygget. Debatten har i vid udstrækning handlet om, at incitamentet til at arbejde ikke har været stort nok. Der har både været argumenter for, at skattelettelse vil have en neglecibel effekt på folks lyst til at arbejde, idet folk vil »købe sig til« mere fritid – og argumenter for, at der vil forekomme en stor effekt på arbejdsudbuddet både for folk, som allerede er i arbejde (time-effekten) og for folk, som initialt er uden for arbejdsmarkedet (deltagelseseffekten). Der er dermed stor uenighed om, i hvor høj grad skattelettelse vil være selvfinansierende og om omfanget af dødvægtstabet ved beskatning i det eksisterende danske skattesystem.

En anden side af debatten om det danske skattesystem er måden, hvorpå individer

Denne artikel er baseret på min kandidatafhandling ved Københavns Universitet, Petersen (2007) og er udarbejdet i forbindelse med Zeuthen prisen. I den sammenhæng vil jeg gerne takke min vejleder Claus Thustrup Kreiner, som har været til stor hjælp i hele processen. Derudover vil jeg gerne takke Henrik Jacobsen Kleven og Claus Thustrup Kreiner for lån af data, som har muliggjort beregningerne af de danske velfærdsvægte. Eventuelle henvendelser kan ske på: livbangsbo@hotmail.com.

med forskellige indkomster er prioriteret i forhold til hinanden. I optimal indkomstbeskatningsteori fokuseres der udelukkende på, hvordan marginalskattebør se ud for de enkelte individer (indkomstgrupper). Den optimale marginalskatte rate findes ved, at samfundsplanlæggeren maksimerer en velfærdsfunktion, som bygger på en velfærdsprioritering for det enkelte individ. Den primære problematik i forhold til at sætte de optimale marginalskatte udmønter sig ved, at staten dels skal tage hensyn til omfordeling mellem indkomstgrupper, og dels skal forsøge at sætte skatterne, så de medfører et mindst muligt efficienstab. Der er dermed tale om det velkendte »trade off« mellem efficiens og lighed. Dette »trade off« er som udgangspunkt nødvendigt, da det ikke er muligt at identificere og beskatte medfødte evner. Samfundsplanlæggeren kan derfor som approksimation til »evnebeskatning« vælge at beskatte efter folks indkomst. Dette har imidlertid den negative effekt, at det forvrider folks arbejdsudbudsbeslutning.

Resultatet af, hvorledes de optimale marginalskatte rater kommer til at se ud for de forskellige indkomstgrupper, afhænger af, hvordan samfundsplanlæggerens underliggende sociale præferencer ser ud. Hvis samfundsplanlæggeren eksempelvis vægter en indkomstgruppe med 2, så betyder det, at der implicit i indretningen af skatte- og velfærdssystemet ligger en vurdering af, at det er dobbelt så godt at give én krone til denne indkomstgruppe sammenlignet med at fordele én krone jævnt ud til alle i samfundet. I den optimale indkomstbeskatningsteori foreligger der standardkriterier for, hvordan acceptable sociale velfærdsvægte bør se ud: (1) *de er faldende i indkomst* (2) *at der ikke er nogen indkomstgrupper, som vægtes negativt*. Udover disse objektive kriterier er der en bred vifte af muligheder for den relative prioritering mellem indkomstgrupperne og dermed stor mulighed for at opnå forskellige optimale marginalskatte rater.¹ Denne høje grad af subjektivitet i valget af sociale velfærdsvægte har den implikation, at teorien ikke bliver så anvendelig.

I stedet for at tage udgangspunkt i det gængse optimale indkomstbeskatningsproblem kan det være interessant at »vende problemet på hovedet« ved at tage udgangspunkt i de eksisterende skatte rater og derudfra forsøge at afsløre de velfærdsvægte for hver indkomstgruppe, som gør et skattesystem optimalt. Denne tilgangsvinkel har Bourguignon og Spadaro benyttet i en artikel fra 2006, hvor de ud fra optimal indkomstbeskatningsteori, udregner velfærdsvægtene for Frankrig. Til at udregne disse vægte benyttes en model, som kun medtager timearbejdsudbuddet (den intensive margin). Nyere litteratur viser dog, at der med stor sandsynlighed findes betydelige dis-

1. Der kan være meget stor forskel på, hvorledes individerne vægtes relativt, og det kan spænde lige fra Rawlske præferencer, der udelukkende prioriterer den dårligt stillede person til præferencer, som vægter alle individer lige meget uanset indkomst. De fleste velfærdsfunktioner anvendt i optimal beskatningsteori er funderet i sociale præferencer, som ligger et sted midt imellem de to ekstreme tilfælde.

krete effekter på folks beslutning om at deltage på arbejdsmarkedet (den ekstensive margin), hvorfor denne også bør inddrages.

Formålet med denne artikel er, at udlede en ny simpel metode til at beregne »afslørede velfærdsvægte for forskellige indkomstgrupper« samt at anvende metoden til at estimere velfærdsvægte for forskellige arbejdsudbudselasticitetsscenarioer på baggrund af en antagelse om, at det danske skattesystem er optimalt. Der benyttes en model, som udover timeeffekten også medtager diskrete deltagelseeffekter, hvilket bidrager til at gøre modellen så realistisk som mulig. I modsætning til Bourguignon og Spadaro, som benytter en metode, der oprinder fra den optimale beskatningsteori, så tager beregningerne i denne artikel udgangspunkt i Marginal Cost of Public Funds (MCF) teorien. Denne teori betragter den efficiensomkostning, der følger af de adfærdsrespons, som individerne foretager i forbindelse med en marginal skatteændring. Kleven og Kreiner (2006c) udleder en formel »Social Marginal Cost of Public Funds« (SMCF), som viser de sociale marginale omkostninger, der er forbundet med at opkræve én ekstra krone i skatteprovenu. Udover efficienseffekten, tager denne formel også højde for den omfordelingseffekt, som en given skattereform medfører. På den måde tager SMCF-målet både højde for efficiens og lighed.

I denne artikel benyttes en heuristisk metode til at finde de danske velfærdsvægte baseret på SMCF-teorien. Udgangspunktet er, at SMCF skal være ens ved alle skattereformer, for at et skattesystem kan være optimalt. For hvis det er tilfældet, at SMCF afviger ved forskellige reformer, er det muligt at opnå velfærdsgevinster ved at flytte rundt på skatterne, og skattesystemet kan derfor ikke være optimalt. Det datamateriale som anvendes i denne artikel, opdeler individerne i 100 indkomstpercentiler, hvilket betyder, at der kan udregnes en SMCF for hver percentil.² Fremgangsmåden for udledningen af de danske velfærdsvægte tager udgangspunkt i, at alle SMCF'er skal være lig med den SMCF, der findes ved en lump sum skat. Ved at antage dette er det muligt at lave en rekursiv udledning af de danske velfærdsvægte, hvor velfærdsvægten for percentil 100 initialt udledes, derefter for percentil 99 osv. Dette forklares nærmere i afsnit 3, hvor velfærdsvægtene efterfølgende udledes.

Udledningen af de danske velfærdsvægte i denne artikel er forankret i en helt anden teoretisk metode, end det er tilfældet for Bourguignon og Spadaros velfærdsvægte for Frankrig. På trods af forskelle i metodikken, anvender begge metoder data af effektive skattesatser, kendskab til indkomstfordelingen samt antagelse af arbejdsudbudselasticiteter. De to metoder adskiller sig yderligere ved, at modellen i denne artikel er baseret på diskrete værdier af data, hvilket medfører diskrete danske velfærdsvægte, mens data i Bourguignon og Spadaros metode er approksimeret kontinuerte funktioner,

2. Data er udarbejdet af Kleven og Kreiner i forbindelse med et forskningsprojekt for Rockwool Fonden, jf. yderligere afsnit 3.2.

hvilket bevirker en pæn kontinuert profil for de franske vægte.³ Ved at benytte den mere heuristiske fremgangsmåde, hvor data ikke har gennemgået en approksimerende udglatning, vil der fremkomme en noget mere volatil profil for de danske vægte. Denne relativt store variation i velfærdsvægtene, der forekommer indkomstgrupperne imellem, kan dog afslutningsvis udglattes ved hjælp af en tendenslinie. Man kan udtrykke det således, at Bourguignon og Spadaro udglatter de benyttede data initialt, mens metoden benyttet i denne artikel anvender de oprindelige data og udglatter til sidst.

Artiklen indeholder en række væsentlige resultater af de afslørede danske velfærdsvægte. For det første viser analyserne, under antagelse af positive deltagelseselasticiteter,⁴ at det danske samfund har implicitte præferencer, der primært tilgodeser den brede middelklasse frem for de lavestlønnede. Dette bryder med 1. kriterium af standard betragtningen, om hvorledes acceptable velfærdsvægte bør se ud. For det andet er der flere af resultaterne, der viser, at der er indkomstgrupper i det danske velfærdssystem, som vægtes negativt under antagelse af centralt skønnede arbejdsudbudselasticiteter.⁵ Dette bryder med 2. kriterium for acceptable velfærdsvægte. Endelig kan resultaterne bidrage til at så tvivl om, hvorvidt det danske skatte- og velfærdssystem er optimalt, idet resultaterne i denne artikel som nævnt bryder med standard betragtningen om, hvorledes acceptable velfærdsvægte bør se ud.

Artiklen er disponeret på følgende måde. Afsnit 2 gennemgår den model, der benyttes til at finde de danske velfærdsvægte. Afsnit 3 udleder og beregner velfærdsvægtene for Danmark. Velfærdsvægtene såvel som metoden sammenholdes med Bourguignon og Spadaros velfærdsvægte for Frankrig. Afsnit 4 diskuterer og konkluderer på resultaterne. Her vil bl.a. blive diskuteret sandsynligheden af en optimal skattestruktur ud fra de resultater, som er fundet i denne artikel.

2. SMCF teorien

I dette afsnit gennemgås den »intensive-ekstensive« model af Kleven og Kreiner (2006c), som ligger til grund for beregningerne af de danske velfærdsvægte.⁶ Derefter findes SMCF-formlen, som er udgangspunktet for den metode, der er valgt til at udlede velfærdsvægtene i afsnit 3.

Når deltagelsesbeslutningen skal modelleres, er det vigtigt at tage højde for det faktum, at når folk træder ind på arbejdsmarkedet, så sker det yderst sjældent, at folk kun

3. De benytter »Kernel estimation«.

4. Der er her tale om deltagelseselasticiteter som er faldende i indkomst, hvilket støttes af størstedelen af den empiriske litteratur, jf. afsnit 3.2.

5. Centralt skønnede elasticiteter refererer til de elasticiteter, som hyppigst findes i troværdige empiriske studier, jf. afsnit 3.2.

6. Ved yderligere information om modellen henvises til Petersen (2007).

udbyder deres arbejdskraft få antal timer. De fleste studier tyder på, at folk arbejder et betydeligt antal timer,⁷ hvilket nødvendiggør en diskret deltagelsesbeslutning.⁸ Dette gøres ved at indføre faste arbejdsomkostninger q (f.eks. transport, børnepasning, stress mv.), som har betydning for individernes beslutning om, hvorvidt de ønsker at deltage på arbejdsmarkedet eller ej. Modellen opdeler individerne i 100 lønpercentiler, hvilket har den egenskab, at der findes lige mange individer i hver percentil $i = 1, \dots, 100$. Heterogenitet i modellen findes i q -parameteren, lønninger og i præferencer.

For at finde det optimale arbejdsudbud på begge marginer løses forbrugerens problem:

$$\begin{aligned} \text{Max}_{c, h} v_i(c, h) - q \cdot 1(h > 0) & \quad (1) \\ \text{s.t. } c = w_i h - T(w_i h, z) & \end{aligned}$$

hvor v er en standard nyttefunktion, h er arbejdsudbuddet og c er forbrug. Sidste led $q \cdot 1(h > 0)$ er en indikatorfunktion, som er lig 1, hvis h er positiv og ellers 0. Skatter og transferinger er inkorporeret i $T(w_i h, z)$, hvor w_i er eksogen bruttoløn, og z_i er en parameter, der opfanger ændringer i skatte- og transfereringssystemet.⁹

For at finde arbejdsudbuddet skal forbrugerens problem først løses ved at finde det optimale udbud af arbejdstimer $h_i = h_i((1 - m_i)w_i, Y_i)$, hvor m_i er marginals-katten for gruppe i , og $(1 - m_i)w_i$ er den marginale løn efter skat og $Y_i \equiv m_i w_i h_i - T(w_i h_i, z)$ kaldes virtual income.¹⁰ Dernæst betragtes deltagelsesbeslutningen, hvor arbejdsmarkedsdeltagelse er betinget af, at nytten ved at være i arbejde er større end nytten ved at være arbejdsløs. Dette betyder, at den øvre grænse for de faste arbejdsomkostninger er givet ved $q \leq v_i(c_i, h_i) - v_i(c_i^0, 0) \equiv \bar{q}_i$, hvor $c_i^0 \equiv -T(0, z)$ er forbrugsniveauet for arbejdsløse i gruppe i . Det implicerer, at individer som har en fast arbejdsomkostning under \bar{q}_i vælger at udbyde deres arbejdskraft med h_i timer, mens individer med q større end \bar{q}_i vælger at være uden for arbejdsmarkedet. Timeudbuddet bestemmes af det enkelte individs skat på marginalen m_i , mens deltagelsesbeslutningen bestemmes af \bar{q}_i . Det er derfor nødvendigt at adskille de to arbejdsudbudsbeslutninger, idet de påvirkes af forskellige parametre. Denne opdeling bevirker, at der skal findes to forskellige arbejds-

7. Jf. bl.a. Cahue og Zyllerberg (2004) som finder, at der er størst tæthed for individer, der arbejder fuldtid og deltid.

8. Hvis deltagelsesbeslutningen i stedet for er kontinuert, betyder det, at individer træder ind på arbejdsmarkedet og arbejder ubetydeligt lidt. Denne adfærsrespons medfører ingen første ordens effekt på statens provenu, idet der er tale om individer, som initialt har optimeret deres arbejdsudbudsbeslutning.

9. T-funktionen gælder både for individer på og uden for arbejdsmarkedet, og er begrænset til at omfatte stykvis linearitet.

10. For mere information om »virtual income« samt håndteringen af »ikke-lineære budgetmængder« henvises til Cahue og Zyllerberg (2004) samt Kleven og Kreiner (2006c).

udbudselasticiteter. Den første er den velkendte ukompenserede timeelasticitet ε_i , som er givet ved: $\varepsilon_i = \frac{\partial h_i}{\partial \omega_i} \frac{\omega_i}{h_i}$, hvor $\omega_i \equiv (1 - m_i)w_i$. Den ukompenserede elasticitet kan ved hjælp af Slutsky-ligningen opdeles i en kompenseret elasticitet (substitutions-effekt) og en indkomsteffekt, og kan defineres som $\varepsilon_i = \varepsilon_i^c - \theta_i$.¹¹

Den ekstensive margin er karakteriseret ved, at der findes to deltagelseselasticiteter. Forskellen på de to elasticiteter er defineret ved, hvorvidt det er nettoindkomsten for beskæftigede der ændres (jf. η_i nedenfor), eller det er nettoindkomsten for de arbejdsløse (jf. η_i^0 nedenfor).

$$\eta_i \equiv \frac{c_i - c_i^0}{P_i} \frac{\partial P_i}{\partial (c_i - c_i^0)} \frac{\partial (c_i - c_i^0)}{\partial c_i} = \frac{(c_i - c_i^0) p_i(q_i)}{P_i(q_i)} \frac{\partial v_i(c_i, h_i)}{\partial c_i},$$

$$\eta_i^0 \equiv \frac{c_i - c_i^0}{P_i} \frac{\partial P_i}{\partial (c_i - c_i^0)} \frac{\partial (c_i - c_i^0)}{\partial c_i} = \frac{(c_i - c_i^0) p_i(q_i)}{P_i(q_i)} \frac{\partial v_i(c_i^0, 0)}{\partial c_i}.$$

Begge deltagelseselasticiteter måler den procentvise ændring i antal arbejdere i gruppe i som følge af en én-procent stigning i differencen mellem forbrug som arbejdende individ og forbrug som arbejdsløs $c_i - c_i^0$. Hvis fritid er et normalt gode, så er den marginale nytte ved forbrug (sidste led i ligningerne) større for en arbejdsløs end for en person i arbejde, hvilket implicerer, at $\eta_i^0 \geq \eta_i$. Forskellen mellem de to elasticiteter reflekterer en indkomsteffekt i deltagelsesbeslutningen.¹²

Kleven og Kreiner (2006c) udleder Social Marginal Cost of Public Funds (SMCF), som er en omfordelingsvægtet Marginal Cost of Public Funds. Der benyttes en additiv Bergson-Samuelson social velfærdsfunktion $\psi(\cdot)$ til inddragelse af omfordelingsaspektet for en given skattereform.

SMCF defineres som $SMCF = -\frac{1}{\lambda} \frac{dW/dz}{dR/dz}$, hvor en lille ændring i skattesystemet identificeres ved en ændring i z -parameteren (dz).¹³ dW og dR er ændringen i henholdsvis den aggregerede velfærd og det aggregerede skatteprovenu som følge af en ændring i skatten. λ betegner den gennemsnitlige sociale marginale nytte af indkomst for hele befolkningen. Denne parametrisering er nødvendig for at konvertere velfærdseffekterne fra nytteenheder over til mere håndgribelige indkomstenheder. For at finde frem til resultatet i ligning (2) nedenfor anvendes »envelope theorem«, således

11. Indkomsteffekten er positiv, såfremt fritid er antaget som et normalt gode. Substitutions- og indkomsteffekter er nærmere beskrevet i bl.a. Varian (1999).

12. Deltagelseselasticiteterne er positive, hvilket betyder, at en stigning i lønraten altid vil øge arbejdsudbuddet på den ekstensive margin.

13. For en mere detaljeret gennemgang af udledningen af SMCF henvises der til Petersen (2007).

at en direkte skattestigning på én kr. giver et velfærdstab for individerne svarende til præcis én kr, men der er ingen velfærdseffekt som følge af individernes adfærdssænderinger i forsøg på at undgå skattestigningen, idet individerne allerede befinder sig i nytteoptimum.

$$\frac{dW/dz}{\lambda} = - \sum_{i=1}^I g_i \frac{\partial T_i}{\partial z} P_i(\bar{q}_i) N_i \quad (2)$$

Her er g_i den gennemsnitlige sociale marginale nytte af indkomst blandt individerne på arbejdsmarkedet i gruppe i relativt til summen af de sociale marginalnytter af indkomst for alle individer. Velfærdsvægtene er skaleret således, at summen af velfærdsvægtene er én. Dette svarer til, at den samfundsmæssige værdi af at give én krone fordelt jævnt ud til alle er lig med én krone. $\frac{\partial T_i}{\partial z}$ betegner den mekaniske effekt, som følger af skattereformen, hvilken er defineret som den provenuforøgelse, en given skattestigning medfører uden nogen form for adfærdseffekter.¹⁴ Ligning (2) viser dermed, at velfærdseffekterne kan skrives som en vægtet sum af alle de mekaniske skattestigninger, som påføres hver indkomstgruppe. Det andet element i SMCF formlen er skattereformens effekt på statens aggregerede provenu, hvilket er givet ved:

$$R = \sum_{i=1}^I \left[T(w_i h_i, z) P_i(\bar{q}_i) + T(0, z)(1 - P_i(\bar{q}_i)) \right] N_i \quad (3)$$

Statens provenu er en sammensætning af nettoskatteindbetalingen for folk i arbejde (1. led) samt for de arbejdsløse (2. led).¹⁵ Ændringen i det aggregerede skatteprovenu, givet en lille skattestigning, kan opskrives som,

$$\frac{dR}{dz} = \sum_{i=1}^I \left[\frac{\partial a_i}{\partial z} - \frac{m_i}{1 - m_i} \left(\frac{\partial m_i}{\partial z} \cdot \varepsilon_i^c - \frac{\partial a_i}{\partial z} \cdot \theta_i \right) - \frac{a_i + b_i}{1 - a_i - b_i} \cdot \frac{\partial a_i}{\partial z} \cdot \eta_i \right] w_i L_i \quad (4)$$

hvor $a_i \equiv T_i / (w_i h_i)$ er den gennemsnitlige skatterate, $\frac{\partial a_i}{\partial z} \equiv \frac{\partial T_i}{\partial z} / (w_i h_i)$ er skattereformens indvirkning på ændringen i den gennemsnitlige skatterate, og $b_i \equiv -T_0 / (w_i h_i)$ er understøttelsesraten efter skat for individer udenfor arbejdsmarkedet. Ligning (4) har den fordel, at den opdeler arbejdsudbuddet på den intensive margin i en substitutions- og indkomsteffekt. Væsentligheden i at opdele den ukompenserede elasticitet i de to effekter består i, at de hver især afhænger af forskellige skattesatser,

14. Adfærdseffekterne er ækvivalente med det ændrede arbejdsudbud, som følger af skattereformen.

15. Skattebyrden er korrigeret for overførelsesindkomster.

hvor substitutionseffekten afhænger af marginalsatten, mens indkomsteffekten afhænger af gennemsnitsskatten. Adfærdsresponsen på den ekstensive margin afhænger, ligesom indkomsteffekten, af skattereformens effekt på gennemsnitsskatten. Men hvor både substitutions- og indkomsteffektens styrke afhænger af det initiale marginalskatte-niveau, så afhænger deltagelseeffekten af den initiale deltagelsesskat $a_i + b_i$. Ved at indsætte velfærdseffekten (2) og provenueffekten (4) i formlen for SMCF findes,

$$SMCF = \frac{\sum_{i=1}^I g_i s_i}{\sum_{i=1}^I \left[1 - \frac{m_i}{1 - m_i} (\phi_i \varepsilon_i^c - \theta_i) - \frac{a_i + b_i}{1 - a_i - b_i} \eta_i \right] s_i} \quad (5)$$

hvor $s_i \equiv \frac{\partial a_i}{\partial z} w_i L_i / (\sum_{j=1}^I \frac{\partial a_i}{\partial z} w_j L_j)$ viser, hvor meget percentil i skal betale af den samlede skattebyrde (den fortæller noget om egenskaberne ved indkomstfordelingen i forhold til en given reform), og $\phi_i \equiv \frac{\partial m_i / \partial z}{\partial a_i / \partial z}$ måler skattereformens progressivitet (hvor meget stiger marginalsatten ift. gennemsnitsskatten for individ i). Skattereformens fordelingsmæssige hensyn opfanges i tælleren, mens nævneren måler reformens efficienseffekt, hvilken er ækvivalent med adfærdsresponsen. De to mellemste led i nævneren reflekterer adfærdsresponsen på den intensive margin, hvilken afhænger af marginalsatten, skattereformens progressivitet samt substitutions- og indkomsteffekten for det individuelle timearbejdsudbud. I sidste led findes adfærdseffekten på den ekstensive margin, som afhænger af deltagelsesskatten og deltagelseselasticiteten. Det faktum, at arbejdsudbuddet på de to marginer afhænger af forskellige skattesatser, understreger nødvendigheden i at opdele dem, idet det ellers vil medføre unøjagtigheder i beregningerne.

3. De danske velfærdsvægte

Den heuristiske fremgangsmåde, som benyttes til at afsløre de sociale præferencer, der ligger til grund for det danske velfærdssystem, er forankret i antagelsen om, at for at en skattestruktur kan være optimal, så skal SMCF være den samme for alle skatter. Dette kan begrundes ud fra følgende; såfremt SMCF afviger ved forskellige skatteinstrumenter, er det muligt at øge den samlede velfærd ved at omrokere skatterne. Det kan eksemplificeres ved at antage, at der skal laves en reform i toppen af indkomstfordelingen, som giver én kr. til percentil 100. Dette skal finansieres ved, at der samlet opkræves én kr. i lump sum skat for hele befolkningen. Når $SMCF_{100}$ er større end SMCF ved en lump sum skat ($SMCF_{lumpsum}$), betyder det, at der samlet er gevinster ved

at gennemføre reformen ($SMCF_{100} - SMCF_{lumpsum} > 0$). Når $SMCF_{100} - SMCF_{lumpsum} < 0$, bør man foretage den »omvendte« reform, hvor der i stedet opkræves én krone i percentil 100, som derefter uddeles lump sum til alle individer, hvilket samlet medfører en velfærdsgevinst. Hvis det derimod er tilfældet, at $SMCF_{100}$ og $SMCF_{lumpsum}$ er identiske, er det ikke muligt at opnå samfundsmæssige gevinster ved at omrokere skatter mellem de to reformer, idet skattesystemet initialt befinder sig i optimum. Det er hele denne tankegang, som ligger til grund for udledningsmetoden, hvor alle $SMCF$ 'er for de 100 per-centiler skal være lig $SMCF_{lumpsum}$, hvilket efterfølgende giver mulighed for at isolere velfærdsvægten for den øverste percentil og derefter rekursivt udlede de resterende velfærdsvægte.¹⁶ Bourguignon og Spadaro (2006) beregner de franske velfærdsvægte ud fra en formel for den optimale marginalskat.¹⁷ I den optimale marginalskatteformel opereres der med kontinuerte marginalskatte og lønfordelingsfunktioner. For at opnå denne form for kontinuitet udglattes værdierne for marginalskatte og lønfordeling vha. »kernelteknik«, som har den fordel, at der kommer en pæn glat profil i de franske velfærdsvægte. Benyttes denne teknik, bør man være opmærksom på, at der foretages et subjektivt valg af, hvorledes den udglattende approksimation kommer til at se ud. Derudover dækker udglatningen over eventuelle variationer i de oprindelige data, hvilket i visse tilfælde kan føre til en noget unuanceret fortolkning af virkeligheden.

Til forskel fra Bourguignon og Spadaros metode udledes de sociale vægte her ud fra formelen for $SMCF$, som udelukkende anvender diskrete værdier. Det er hermed muligt, at operere med diskrete velfærdsvægte, som er baseret på reelle marginalskatte og indkomsttætheder og ikke blot en approksimation. Det viser sig imidlertid, at konsekvensen af at benytte diskrete værdier er en væsentlig variation i velfærdsvægtprofilen, hvorfor man med fordel afslutningsvist kan udglatte velfærdsvægtene med en tendenslinie.

3.1 Udledning af velfærdsvægtene

Den $SMCF$, som alle $SMCF$ værdier i et skattesystem bør være identisk med, er en lump sum beskatning, som betales af alle individer, uafhængigt af om de er på arbejds-

16. Rent teoretisk kan man antage, at alle $SMCF$ 'er skal være lig en hvilken som helst anden $SMCF$ end $SMCF_{lumpsum}$, f.eks $SMCF_k$, hvor reformen er målrettet percentil k ($k = 1, \dots, 100$). På den måde er det ligeledes muligt at se, om skattesystemet er optimalt. Det vil dog være problematisk i forhold til selve udledningen af velfærdsvægtene. Problemet består i, at alle velfærdsvægte fra percentil k forekommer i tælleren på $SMCF_k$ (jf. ligning (7) nedenfor), hvilket vil komplicere den rekursive metode, hvor velfærdsvægtene findes »bagfra« med start i percentil 100. Ved at benytte fremgangsmåden, hvor $SMCF$ skal være lig med $SMCF_{lumpsum}$, er det muligt at normalisere tælleren i $SMCF_{lumpsum}$ til 1, så man slipper af med velfærdsvægtene fra percentil k , hvilket muliggør den rekursive udledning (jf. udledningen af vægtene i næste afsnit)

17. Deres model oprinder fra Mirrlees' (1971) optimale beskatningsproblem.

markedet eller ej. Det betyder, at en lump sum skat ikke ændrer marginalsatten, hvilket har den implikation, at substitutionseffekten elimineres. Derved er det udelukkende indkomsteffekten, der er aktuel, og konsekvensen er, at arbejdsudbudsbeslutningen ikke forvrides. Udover den velkendte indkomsteffekt på den intensive margin, finder Kleven og Kreiner som tidligere nævnt, at det er nødvendigt på den ekstensive margin at sondre mellem ændringen af nettoindkomsten for beskæftigede og nettoindkomsten for arbejdsløse individer, hvilket fremgår af $\eta_i^0 - \eta_i$.¹⁸ Ved at foretage en normalisering i nævneren til 1 kan $SMCF_{lumpsum}$ skrives på følgende måde:¹⁹

$$SMCF_{(lumpsum)} = \frac{1}{\sum_{i=1}^{100} \left[1 + \frac{m_i}{1-m_i} \theta_i + \frac{a_i + b_i}{1-a_i-b} (\eta_i^0 - \eta_i) \right] n_i} \quad (6)$$

Såfremt der findes positive indkomsteffekter enten på timeudbud, deltagelse eller begge dele, bliver $SMCF$ mindre end 1. For at skattesystemet er optimalt, givet de aktuelle effektive skattesatser, så skal $SMCF_{lumpsum}$ være lig $SMCF$ for hver af de 100 indkomstpercentiler. Indkomstfordelingens opdeling i percentiler betyder, at hver indkomstpercentil har hver sin marginal- og deltagelsesskat. Med henblik på at reflektere hele indkomstfordelingen udregnes $SMCF$ ved at lave en lille skattereform, der kun hæver marginalsatten i én indkomstpercentil k , mens de resterende indkomstgruppers marginalsatter er uberørte.²⁰ $SMCF$ bliver følgende:

18. Den antages for at være ikke-negativ. Jf. yderligere i Kleven og Kreiner (2006b). For yderligere information om selve udledningsmetoden henvises der til Petersen (2007).

19. Den oprindelige formel er givet ved:

$$SMCF_{(lumpsum)} = \frac{\sum_{i=0}^{100} g_i n_i}{\sum_{i=1}^{100} \left[1 + \frac{m_i}{1-m_i} \theta_i + \frac{a_i + b_i}{1-a_i-b_i} (\eta_i^0 - \eta_i) \right] n_i}$$

Her er $n_i = \frac{N_i}{\sum_{i=0}^{100} N_i}$ hvor N_i er antallet af individer i gruppe i , og N_0 er antal arbejdsløse i befolkningen.

I stedet for at operere med en vilkårlig størrelse i tælleren kan der foretages en normalisering til 1. $\eta_i^0 - \eta_i$ er antaget til 0,03 i den numeriske analyse, jf. yderligere Petersen (2007).

20. Jf. Kleven og Kreiner (2006c).

$$SMCF_k = \frac{\sum_{i=1}^I g_i s_i}{1 - \frac{m_k}{1 - m_k} \varepsilon_k^c \phi_k s_k + \sum_{i=k}^I \frac{m_i}{1 - m_i} \theta_i s_i - \sum_{i=k}^I \frac{a_i + b_i}{1 - a_i - b_i} \eta_i s_i} \quad (7)$$

Tælleren viser skattereformens effekt på lighed, mens nævneren reflekterer effi-cienseffekten. Da der udelukkende er tale om, at det er individer i percentil k , som oplever en stigning i deres marginalskat, er det derfor kun den indkomstgruppe, som via 1. led i nævneren reducerer deres timearbejdsudbud som følge af substitutionseffekten. Derimod øger alle indkomstgrupper større eller lig med k deres timeudbud i nævnerens 2. led som følge af en højere gennemsnitsskat (indkomsteffekten). I lighed med indkomsteffekten, involverer deltagelseeffekten alle individer fra og med k , og konsekvensen er, at nogle individer vælger at træde ud af arbejdsmarkedet pga. de højere gennemsnitsskatter (nævnerens sidste led).²¹ I den øverste percentil er der, ud over velfærdsvægten g_{100} , kendskab til de resterende parametre, hvorved g_{100} efterfølgende kan udledes således at $SMCF_{i=100} = SMCF_{lumpsum}$.²² Når g_{100} er fundet, findes de resterende velfærdsvægte rekursivt ved at bevæge sig trinvist ned i indkomstfordelingen.²³ Den generelle formel for velfærdsvægtene $k = 1, \dots, 100$ bliver følgende:

$$g_k = \frac{1 - \frac{m_k}{1 - m_k} \varepsilon_k^c \phi_k s_k + \sum_{i=k}^{100} \left[\frac{m_i}{1 - m_i} \theta_i - \frac{a_i + b_i}{1 - a_i - b_i} \eta_i \right] s_{i(k)}}{\sum_{i=1}^{100} \left[1 + \frac{m_i}{1 - m_i} \theta_i + \frac{a_i + b_i}{1 - a_i - b_i} (\eta_i^0 - \eta_i) \right] n_i} - \sum_{i=k+1}^{100} g_i s_{i(k)} \quad (8)$$

Formlen skal forstås på følgende måde: k er den indkomstpercentil, hvor velfærdsvægten udledes, og s_k er den andel af skattereformen, som individer i percentil k skal

21. S_k er den andel, som individer fra percentil k skal betale af den samlede skattebyrde

$S_k = \frac{y_k - y_{-k}}{y_k - y_{-k} + (100 - k)(y_{-k+1} - y_{-k})}$, hvor y_k er gennemsnitsindkomsten i percentil k , og y_{-k} er den

nedre indkomstgrænse for percentil k . S_i er derimod den andel af den samlede skattebyrde, som alle indkomstgrupper større end k skal betale $S_i = \frac{y_{-k+1} - y_{-k}}{y_k - y_{-k} + (I - k)(y_{-k+1} - y_{-k})}$ (hvilken er den samme for alle grupper).

22. Der er her antaget et skøn for realistiske arbejdsudbudselasticiteter, jf. næste afsnit.

23. g_{100} indsættes efterfølgende i $SMCF_{i=99} = SMCF_{lumpsum}$ hvorefter g_{99} findes osv.

betale. Betegnelsen $s_{i(k)}$ er den andel, percentil i har af skattebyrden, som følge af deres ændrede gennemsnitskat, forårsaget af skattereformen i percentil k (den er ens for alle $i > k$). Ud fra ligning (8) kan man se, at velfærdsvægten for indkomstgruppe k afhænger negativt af de aggregerede adfærdseffekter ved en marginalskatteændring i k (de tre sidste led i tælleren). Derudover vil en højere $SMCF_{lumpsum}$ føre til en generel højere velfærdsvægt, idet gabet mellem SMCF ved en forvridende skat og SMCF ved en lump sum skat formindskes. Med andre ord; jo højere $SMCF_{lumpsum}$ bliver, jo mindre bliver efficiensgevinsten alt andet lige ved at gå fra den forvridende marginalskat over til lump sum beskatning, og jo højere vil velfærdsvægten være i indkomstgruppe k .²⁴ Derudover afhænger g_k negativt af summen af de velfærdsvægte, der er større end k , hvilket er med til at frembringe støj i resultaterne.

Den sociale velfærdsvægt for de arbejdsløse findes residualt. Ud fra tidligere normalisering er $\sum_{i=0}^{100} g_i n_i = 1$, og g_0 bliver følgende:

$$g_0 = \frac{1 - \delta \sum_{i=1}^{100} g_i \frac{1}{100}}{1 - \delta} \quad (9)$$

Her fremgår det, at udover summen af velfærdsvægtene for alle arbejdende individer, afhænger de lediges velfærdsvægt ligeledes af andelen af individer uden for arbejdsmarkedet $1 - \delta$, således at færre arbejdsløse, alt andet lige, medfører en højere velfærdsvægt for denne gruppe.²⁵

3.2 Data og elasticiteter

For at opnå stor grad af indkomstvarietet er det datamateriale, som ligger til grund for den numeriske analyse opdelt i 100 percentiler. Data er udarbejdet af Kleven og Kreiner (2006a) og er beregnet på Finansministeriets Lovmodel i forbindelse med udgivelsen af »Skat, arbejde og lighed« fra Rockwool Fondens Forskningsenhed.²⁶ Data-sættet omfatter 33,3 pct. af alle fuldtids- og deltidsbeskæftigede i Danmark, der er mellem 18 og 65 år, og som minimum tjener 35.600 kroner om året.²⁷

24. Det er vigtigt at huske på, at beregningerne af velfærdsvægtene er foretaget ud fra, hvordan skattesystemet beskatter hver enkel percentil i samspil med individernes fordeling af løn samt et skøn over arbejdsudbudselasticiteterne (samlet set adfærdssvarne).

25. δ er andelen af beskæftigede ifht. arbejdsdygtige og er i overensstemmelse med Kleven og Kreiners skøn sat til 83 pct.

26. Udregning af de effektive skattesatser og lønfordelingen er baseret på mikrodata fra 2003 og er fremskrevet til 2006 pris- og lønniveau (der er taget højde for lovgivning gældende i 2006). Alle figurer i artiklen bygger på data lånt af Kleven og Kreiner (2006b). Alle kildehenvisningerne i figurene vil blot henvise til data beregnet af Kleven og Kreiner (2006b).

27. I 2006 priser (eksklusiv studerende).

Effektive skattesatser

Når skattesatserne anvendes til at beregne effekterne på arbejdsudbud, er det væsentligt at benytte effektive skattesatser, idet disse medtager alle relevante elementer i forhold til incitamentet til at arbejde. Data indbefatter indkomstskatter, vareskatter, moms og andre afgifter, som kan have effekt på den reale forbrugsværdi af arbejdsindkomsten.²⁸ De effektive skattesatser afspejler indkomstens købekraft hos det enkelte individ, og det er vigtigt at benytte disse rater, da det er herudfra, at individerne træffer deres beslutninger.

De effektive marginalsatser ligger på ca. 60 pct. for den lavere ende af indkomstfordelingen, hvorefter de stiger gradvist til ca. 70 pct. for den øvre del af indkomsterne. Deltagelsesskatterne ligger på et generelt højere niveau, og profilen for disse er i modsætning til marginalsatserne faldende, idet overførslerne udfases i takt med, at indkomsten stiger. Deltagelsesskatten ligger på ca. 80 pct. for den nedre del af indkomstfordelingen og falder til ca. 70 pct. i toppen.²⁹

Lønfordelingen

Danmark er karakteriseret ved en venstreskæv lønfordeling, hvilket er ensbetydende med, at over halvdelen af befolkningen tjener mindre end den gennemsnitlige årlige arbejdsindkomst (inkl. pensionsopsparing), som var 316.326 kr. i år 2006. Den danske lønfordeling er yderligere forbundet med stor grad af tæthed sammenlignet med andre lande, og der er relativt få mennesker, som har høje indkomster. Egenskaberne ved lønfordelingen har store implikationer for størrelsen af efficiensomkostningerne ved beskatning og derved også, hvorledes velfærdsvægtene fordeles imellem individer. Hvis der er stor koncentration af individer tæt på den grænse, hvor der indføres en marginalskatteøgning, vil adfærdseffekterne have relativ stor effekt i forhold til det mekaniske provenu.³⁰ Således er det forholdsvis alvorligt at have høje skatter i percentiler, hvor der er stor koncentration af individer tæt på den nedre indkomstgrænse, idet efficiensomkostningerne er store.

28. Der er også taget højde for udfasning af offentlige overførsler som følge af højere arbejdsindkomst, da det er ækvivalent med en skat på arbejde (her kan bl.a. nævnes dagpenge, friplads i daginstitution, boligydelse mv.). Desuden gælder der også specielle regler for ægtepar, hvor der skal tages højde for samspilseffekter i tildeling af overførsler og skattepligtig indkomst.

29. Marginal- og deltagelsesskatterne (samt time- og deltagelseselasticiteterne) varierer fra individ til individ inden for hver indkomstpercentil. Her benyttes den gennemsnitlige skattesats for hver percentil (gns. skøn over elasticiteterne), og der ses derfor bort fra variationen inden for indkomstgrupperne. Forskellige sociale karakteristika så som køn, børn, alder, bopæl mv. kan spille ind på den effektive skattesats, jf. Kleven og Kreiner (2006a). En interessant vinkel i forhold til fremtidige undersøgelser, som artiklen ikke berører, er hvorledes forskelle i sociale karakteristika kan føre til stor variation i velfærdsvægtene inden for hver indkomstpercentil.

30. At der betragtes en skatteøgning, er blot et led i metoden til at finde vægtene, hvilket også fremgik af afsnit 2.1.1.

Arbejdsudbudselasticiteter

I den empiriske litteratur finder de mest realistiske studier en arbejdsudbudselasticitet på den intensive margin (timeelasticitet/ukompenseret elasticitet) liggende mellem $0 - 0,3$.³¹ Undersøgelserne tyder på, at elasticiteterne er relativt homogene på tværs af indkomster og køn. I forhold til deltagelseeffekter finder en stor del af studierne betydelige effekter specielt i bunden af indkomstfordelingen. Resultaterne er karakteriseret ved stor variation i forhold til sociale karakteristika, hvor der specielt blandt de lavestlønnede, enlige mødre, indvandrere og lavtuddannede er en tendens til høje deltagelseselasticiteter. De fleste empiriske studier finder deltagelseselasticiteter mellem $0,2 - 0,8$ for specielt udvalgte grupper.³² For de højere indkomstgrupper findes der generelt ingen signifikante deltagelseeffekter, hvorfor en faldende profil for deltagelseselasticiteter bedst afspejler litteraturens empiriske estimater.³³

Nyere undersøgelser tyder imidlertid på, at det ikke blot er arbejdsudbuddet, der har betydning for størrelsen af adfærdsrespons, men også folks reaktioner i form af mobilitet på arbejdsmarkedet, incitament til uddannelse, løn udbetalt i frynsegoder, effort på jobbet osv. Alle disse adfærdsrespons, som følger af en skatteændring, opfanges i ændringen i individernes skattepligtige indkomst. Det er dermed alle disse adfærdsrespons, der samlet måler dødvægtstabet ved beskatning, og derved det mål, som man bør benytte, når man skal beregne adfærdsomkostningerne ved det danske skattesystem (elasticitet af skattepligtig indkomst). De fleste studier finder betydelige positive elasticiteter, og en stor del af resultaterne ligger omkring $0,4$.³⁴

Elasticitetsscenerier

Med udgangspunkt i ovenstående empiriske undersøgelser vælges følgende elasticitetsscenerier:

Scenario 1 (S1): Bench mark tilfælde hvor de kompenserede og ukompenserede timeelasticiteter er $0,1$ (ingen indkomsteffekter) for alle indkomstgrupper og ingen del-

31. Jf. bl.a. Blundell m.fl. (1998) for England, Triest (1990) for USA og Frederiksen m.fl. (2001) for Danmark.

32. Eissa og Liebman (1996) finder bl.a. deltagelseeffekter på ca. $0,6$ for enlige mødre med lav uddannelse ved den amerikanske skattereform i 1986. Meyer og Rosenbaum (2001) finder elasticiteter omkring $0,7$ for enlige amerikanske mødre i perioden 1985-1997. Gravensen (1996) finder elasticiteter mellem $0,2$ og $0,7$ for enlige kvinder som følge af den danske skattereform i 1987.

33. Her er der tale om, at den empiriske litteratur ikke finder væsentlige deltagelseselasticiteter på nationalt plan. Det kan evt. tænkes, at der kan være nogle internationale effekter.

34. Feldstein (1995) er nok det mest kendte studium med meget markante resultater (elasticiteter for USA mellem $1,5$ og $3,0$). Jf. yderligere Goolsbee (2000) (USA), Gruber og Saez (2002) (USA), Ljunge og Ragan (2005) (Sverige). Jf. desuden Kleven og Kreiner (2006a).

tagelseeffekter.³⁵ Der tilsættes derefter en deltagelseselasticitet på 0,2 for percentil 1, som derefter falder gradvist ($S1(D)$).³⁶

Scenario 2 (S2): Her er der i forhold til $S1(D)$ medtaget en indkomsteffekt, således at timeelasticiteten er 0,05, hvor den kompenserede timeelasticitet er på 0,1, og indkomsteffekten er på 0,05. Der er medtaget en deltagelseselasticitet på 0,2 i indkomstpercentil 1. I analysen af velfærdsvægtene vil der kortfattet kommenteres på effekten af homogen deltagelse på tværs af indkomst $S2(H)$, hvor deltagelseeffekten er 0,2 for alle percentiler.

Scenario 3 (S3): Her øges deltagelseselasticiteten initialt til 0,4. Timeelasticiteten er som i S2.

Scenario 4 (S4): I dette scenario hæves den ukompenserede timeelasticitet for alle indkomstpercentiler til 0,1.³⁷ Den kompenserede elasticitet er nu 0,15, dvs. indkomsteffekten er stadig 0,05. Deltagelseselasticiteten er som i S2.

3.3 Resultater af velfærdsvægtene

Ved at anvende denne metode, hvor de effektive skattesatser er diskrete, og hvor indkomsterne er opdelt i percentiler, er konsekvensen, at velfærdsvægtene er forbundet med stor grad af negativ autokorrelation på grund af støj. Det viser sig på den måde, at den estimerede velfærdsvægt afhænger negativt af den foregående vægt. I områder, hvor velfærdsvægtene har numeriske store værdier, vil der forekomme større svingninger end i percentiler med mindre værdier. Ved at tage gennemsnittet af to efterfølgende velfærdsvægte ($g_i^* = \frac{g_i + g_{i+1}}{2}$) mindskes autokorrelationen markant.³⁸

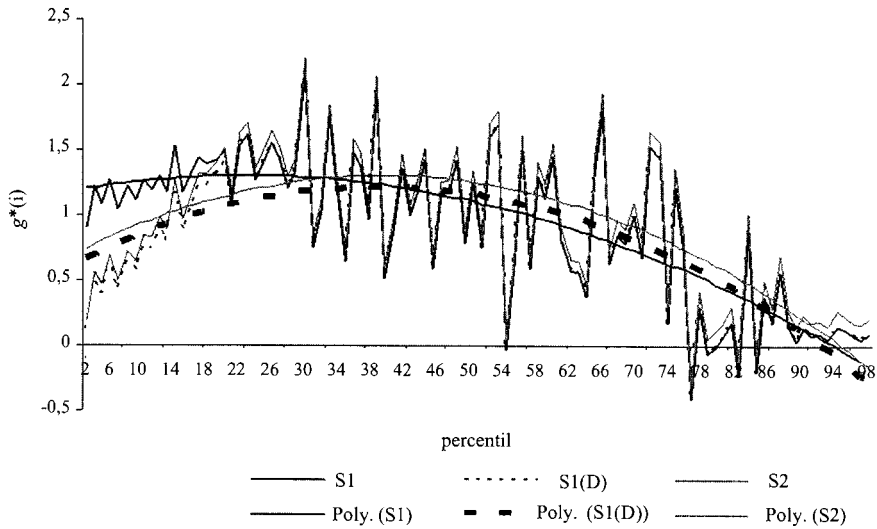
På figur 1 kan man se velfærdsvægtprofilen for elasticitetsscenerierne $S1$, $S1(D)$ og $S2$, hvor $S1$ svarer til lavelasticitetssceneriet i Bourguignon og Spadaro (2006). De tynde kurver svarer til de beregnede velfærdsvægte g_i^* , mens de tykke kurver er de udglattende tendenslinier. Det viser sig, at det er måden, hvorpå individerne er fordelt på de forskellige indkomstintervaller, som er den primære årsag til eksistensen af svingninger i velfærdsvægtene. Der forekommer stor variation i størrelsen af intervallerne, idet der som bekendt findes nøjagtigt lige mange individer i hvert indkomstinterval. I midten af indkomstfordelingen, hvor der er stor grad af indkomsttæthed, er intervallerne meget små i forhold til enderne, hvor intervallerne er noget større pga. større indkomstspredning. De store svingninger, der findes i midten af g_i^* -profilen, skyldes dels, at store værdier af velfærdsvægtene er mere volatile (den førnævnte negative autokor-

35. Dette scenario er ofte benyttet i litteraturen og anvendes også i Bourguignon og Spadaros (2006) analyse af de franske velfærdsvægte.

36. Medmindre at andet er nævnt, antages det, at den pålydende deltagelseselasticitet altid er knyttet til percentil 1, hvorefter den falder med 0,01 i hver percentil indtil den rammer 0.

37. Denne timerespons er analog til Kleven og Kreiners »centrale skøn« i Kleven og Kreiner (2006a).

38. De oprindelige velfærdsvægte g_i findes i Petersen (2007).



Figur 1. Velfærdsvægte for S1, S1(D) og S2.³⁹

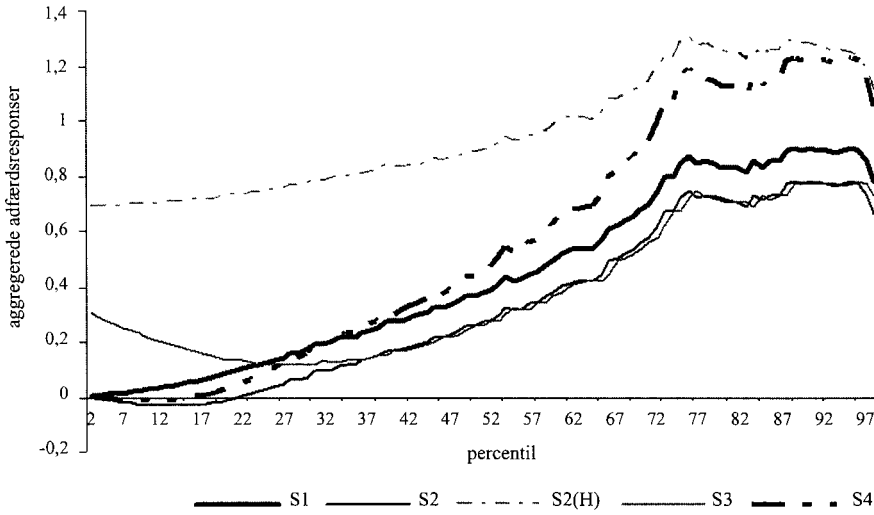
Kilde: Egne beregninger ud fra Data af Kleven og Kreiner (2006b).

relation), og dels at indkomstpercentiler med stor grad af tæthed er forbundet med større variation.⁴⁰

I figur 1 er profilen for velfærdsvægtene i S1 (de sorte kurver) nogenlunde konstant på et niveau omkring 1,2 – 1,3 i den nedre lønfordeling, hvorefter profilen falder og ender på vægte omkring 0,1 for de øverste indkomster. Det svarer til, at de rigeste individer vægtes ca. 12 gange mindre end lavindkomstgrupperne. Velfærdsvægten for de arbejdsløse er på 1,47. Tendensen i velfærdsvægtene for de højeste indkomster skyldes primært profilen i adfærdsresponserne, som kan ses på figur 2, idet de er negativt korrelerede (jf. ligning (8)). Adfærdsresponserne viser i hvor høj grad, individerne reagerer på en given skatteændring, hvilket er ækvivalent med selvfinanseringsgraden, således at en adfærdsrespons på 1 svarer til, at skatteændringen netop er fuldt ud selvfinansierende (en adfærdsrespons på f. eks. 1,2 svarer til at reformen er »120 pct. finansieret« osv). Figur 2 er ligeledes med til at give en god forståelse af udviklingen i vægtene for de resterende scenarier.

39. Alle graferne medtager kun velfærdsvægte fra og med percentil 2. g_1 er fravalgt, idet individer i denne indkomstgruppe formentlig reelt ikke er på arbejdsmarkedet, idet indkomstintervallet er 0-54.000 kr. om året. Derudover divergerede g_1^* generelt fra de andre vægte, idet indkomstintervallet for denne gruppe er relativt større, end det er gældende for de efterfølgende percentiler, jf. yderligere Petersen (2007). Resultater af g_0 findes løbende i teksten.

40. I Petersen (2007) Appendix 6.3 kan man tydeligt se sammenhængen mellem svingninger i velfærdsvægtene (differensen mellem g_{i+1}^* og g_i^*) og størrelsen af det dertil hørende indkomstinterval.



Figur 2. Aggregerede adfærdsresponsers for S1-S4.

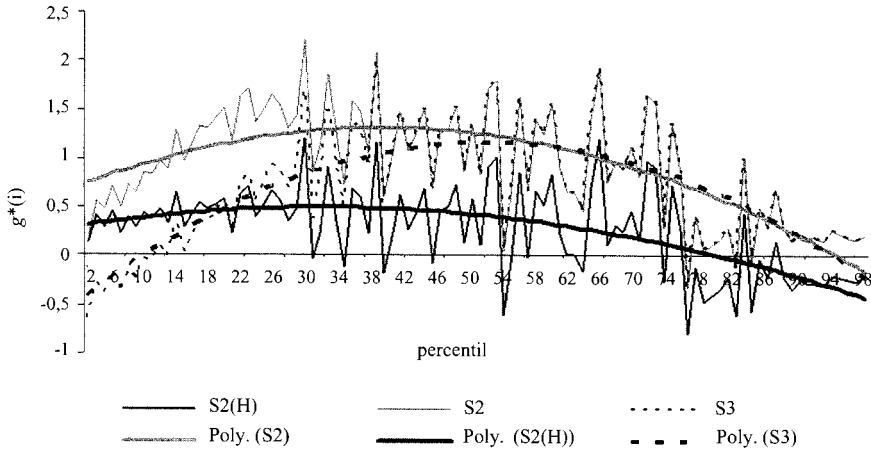
Kilde: Egne beregninger ud fra Data af Kleven og Kreiner (2006b).

Tilbage til figur 1, hvor der i de nedre percentiler i S1 forekommer en meget svag stigende profil i g_i^* på trods af, at der ikke er nogen deltagelseselasticiteter. Denne svage tendens kan tilskrives egenskaber ved indkomstfordelingen og ikke profilen for marginalskatte. Dette illustreres tydeligst ved at betragte et hypotetisk eksempel, hvor der antages homogene marginalskatte for alle indkomstgrupper, hvilket ikke eliminerer de svagt stigende vægte i den nedre lønfordeling.^{41, 42}

Til sammenligning med de franske velfærdsvægte har de danske vægte samme konkave profil, dog på forskellige niveauer. Niveauforskellen kan formentlig skyldes, at Danmark generelt opererer med langt højere marginalskatte, og at marginalskattepro-

41. Det findes ikke relevant at vise grafen for S1 med homogene skatter, idet profilen for velfærdsvægtene i den nedre lønfordeling, i det store hele, er identisk med S1. Formålet med at nævne eksemplet er blot at vise, at denne svagt stigende profil i velfærdsvægtene ikke kan tilskrives heterogene marginalskatte. Når der så samtidig er homogene elasticiteter, må profilen i g_i^* tilskrives egenskaber ved lønfordelingen. Her er det blot antaget, at marginalskatte er konstante på percentil 1 niveau. Det er blot et illustrativt eksempel, hvor der ikke er taget hensyn til provenueffekter.

42. Det er desuden værd at bemærke, at der ud over de generelle svingninger forekommer et skift i velfærdsvægtene omkring percentil 80. Dette skift, som går igen i alle elasticitetsscenerier, skyldes formentlig, at der forekommer et lidt større spring i S_k parameteren, end det hidtil har været tilfældet. Konsekvensen bliver, at når der i ligning (8) divideres med S_k , vil en relativ større S_k værdi (i forhold til værdierne i de lavere indkomstpercentiler) føre til en relativ mindre vægtning. Omkring percentil 54-55 sker der et mindre skift, hvilket kan henledes til et mindre spring i adfærdsresponsene, jf. figur 2. Metoden er generelt meget sensitiv over for variationer.



Figur 3. Velfærdsvægte for S2, S2(H) og S3.

Kilde: Egne beregninger ud fra Data af Kleven og Kreiner (2006b).

filen begynder at stige længere nede i indkomstfordelingen, end det er tilfældet for Frankrig. Samtidig er lønfordelingen i Danmark karakteriseret ved, at individer i toppen af indkomstfordelingen ligger tæt på den nedre grænse (y_{-k}), hvilket implicerer, at adfærdsresponserne er meget høje.⁴³

På figur 1 viser de stiplede kurver S1(D), hvor der er medtaget deltagelseselasticiteter. Det ses tydeligt, at profilen for velfærdsvægtene stiger mærkbart i bunden af indkomstfordelingen. Der er dermed tale om, at velfærdsvægtene får en invers U-formet profil selv med relativt beskedne elasticitetsværdier. Velfærdsvægten for de arbejdsløse er 1,92. I S2 er der i forhold til S1(D) medtaget indkomsteffekter på 0,05, hvilket kan ses ved de grå grafer i figur 1. Grafen for S2 ligger for alle percentiler en anelse over velfærdsvægtene for S1(D). Dette er intuitivt, idet indkomsteffekten trækker i retning af større arbejdsudbud, hvilket mindsker effekten på adfærdsresponserne ved en skattestigning. Velfærdsvægten for de arbejdsløse er nu faldet til 1,79. Velfærdsvægten for percentil 2 g_2 starter på ca. 0,2 og profilen er undervejs som i S1(D), og g_{100} er ligeledes omkring 0,2.

43. De adfærdsrespons, som følger af en given skattereform, bliver relativt høje ift. det mekaniske provenu, som skabes. Der er i Petersen (2007) foretaget en sammenligning af, hvordan individerne er placeret ift. nedre grænse for både Frankrig og Danmark. Det viste sig, at individerne i Danmark generelt lå tættere på den nedre grænse for de øverste indkomster, hvilket alt andet lige betyder, at adfærdseffekterne er relativt alvorlige i Danmark, jf. yderligere Petersen(2007). Derudover er det vigtigt at være påpasselig med validiteten i sammenligningerne af niveauforskelle mellem Danmark og Frankrig, idet velfærdsvægtene er beregnet på forskelligt metodegrundlag.

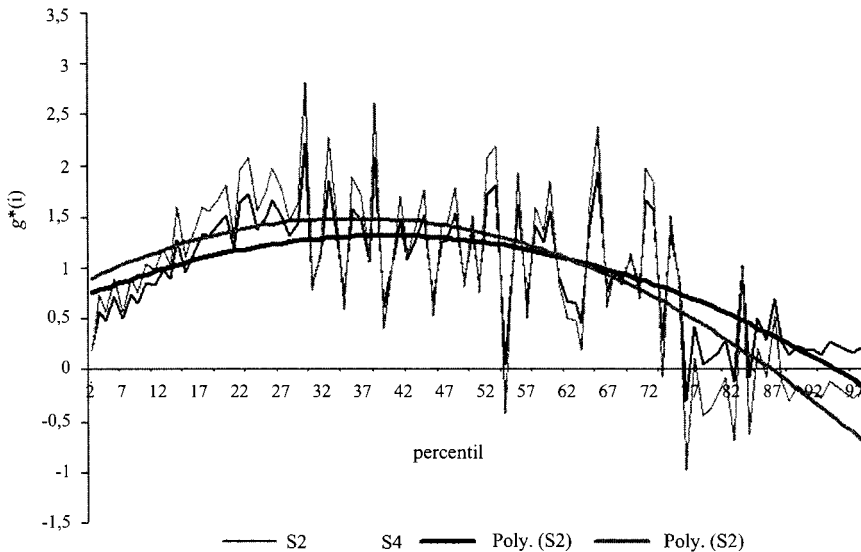
På figur 3 kan man se de danske velfærdsvægte med forskellige deltagesscenarier (S2, S2(H) og S3).

Den stigende profil i bunden af indkomstfordelingen for S2 skyldes primært, at deltagelseeffekterne er antaget faldende i indkomst. Dette kan bedst illustreres ved at betragte det specialtilfælde af S2, hvor deltagelseselasticiteten er antaget homogen på 0,2 for alle indkomster (S2(H), som ses ved de sorte linier i fig. 3). Det er nu tydeligt, at den stigende tendens er forsvundet, og selve profilen har ændret sig markant. Ved homogene deltagelseeffekter ligger velfærdsvægtene for alle percentiler på et markant lavere niveau, og profilen er mindre konkav. Derudover bliver velfærdsvægtene for de 25 øverste indkomster negative i S2(H). Det lavere niveau i vægtene for S2(D) giver udslag i en væsentlig højere vægt for de arbejdsløse på 4,62 set ift. de beskæftigede. Den store forskel i vægtene for S2 og S2(H) skyldes det faktum, at når der også medtages deltagelseeffekter for de højeste indkomster, så har små skattestigninger i de nederste percentiler store implikationer, da det kan betyde, at højtlønnede træder ud af arbejdsmarkedet. Dermed går den danske statskasse glip af et stort skatteprovenu, hvilket gør adfærdsresponsen relativt mere alvorlige, jf. figur 2.

Hvis deltagelseselasticiteten øges til 0,4 for percentil 1 (S3), bliver velfærdsvægtprofilen endnu mere invers U-formet, hvilket fremgår af de stiplede linier i figur 3. Det skyldes dels, at deltagelseselasticiteterne er blevet større og dels, at der er flere indkomstintervaller, som har en positiv deltagelseselasticitet.⁴⁴ Velfærdsvægten for percentil 2 er på -0,60, hvorefter den stiger og bliver positiv i percentil 12. Vægtene i den øvre fordeling svarer til dem i S2. Dette resultat er meget markant, idet individer med meget lav løn vægtes ekstremt lidt i forhold til individer med indkomster i midten af indkomstfordelingen. De laveste indkomster vægtes endda negativt. Det viser sig rent faktisk, at i dette tilfælde er velfærdsvægtene for de lavestlønnede mindre end vægtene for de rigeste indkomstgrupper. Denne væsentlig lavere prioritering af de lavestlønnede på arbejdsmarkedet slår igennem ved at de arbejdsløse nu vægtes med 3,03 i forhold til 1,79 i S2.

På figur 4 kan man se velfærdsvægtene for S4, hvor der er sket en stigning i timeelasticiteten. Denne forøgelse sammenholdes med den lavere timeelasticitet i S2. Det fremgår tydeligt af figuren, at en højere timeelasticitet øger velfærdsvægtene for den nedre del af indkomstfordelingen, hvilket sker på bekostning af vægtene for de ca. 40 øverste indkomstpercentiler, som ligger på et generelt lavere niveau, med negative vægte for de aller højeste indkomster. Profilen er stadig invers U-formet i S4 således, at det igen er de mellemste indkomster, der prioriteres mest. De arbejdsløse vægtes med 1,81 i S4.

44. Deltagelseselasticiteten falder altid med 0,01 fra hver indkomstpercentil. Så jo højere effekten er for percentil 1, jo flere indkomstpercentiler berøres af positive responser inden den til sidst rammer 0.



Figur 4. Velfærdsvægte for S4 og S2.

Kilde: Egne beregninger ud fra Data af Kleven og Kreiner (2006b).

4. Konklusion og diskussion

Formålet med artiklen er primært at udlede en ny simpel metode til at beregne velfærdsvægte ved forskellige elasticitetsscenerier for forskellige indkomstgrupper i Danmark, hvilket sker under antagelse af, at det eksisterende skattesystem er optimalt. Det bliver dermed muligt at afsløre, om de danske velfærdsvægte lever op til de to standardkriterier for accepterede velfærdsvægte i økonomisk teori, hvor vægtene er: (1) *faldende i indkomst* og (2) *ikke negative*. Ved at sammenligne resultaterne af de danske velfærdsvægte i forhold til de to objektive kriterier, er det muligt at diskutere sandsynligheden for optimalitet i det danske skattesystem.

Metoden til at beregne de danske velfærdsvægte, som blev udledt i artiklen, er en heuristisk fremgangsmåde, som tager udgangspunkt i SMCF-teorien. Initialt er de fundne velfærdsvægte forbundet med stor variation, men der fremkommer en klar profil gennem støjen, og ved afslutningsvis at anvende en udglattende tendenslinje, bliver metoden et brugbart redskab til at beregne velfærdsvægte for forskellige indkomstgrupper.

Beregningerne af velfærdsvægtene viser, at ved at medtage deltagelseselasticiteter aftagende i indkomst, så får velfærdsvægtene en inverteret U-formet profil, hvilket betyder, at det er mellemindkomsterne, der vægtes mest. Dette er imidlertid i konflikt

med 1. kriterium, som definerede, at vægtene skal være faldende i indkomst.⁴⁵ Analysen viser desuden, at jo større deltagelseselasticiteter for lavindkomsterne, jo mere konkav (invers U- formet) bliver profilen for velfærdsvægtene.

I forhold til optimal beskatningsteori synes det umiddelbart forkert, at mellemindkomsterne vægtes højere end de laveste indkomster. En måde at forklare dette atypiske forløb på kan være at betragte en mere »politisk økonomisk tilgangsvinkel«, hvor politik ikke blot er ideologi men også et middel til magt (f.eks. opnå flertal). I politisk økonomisk litteratur finder man teorien om »Median Voter,« hvor politikken bør tilrettelægges således, at den tilgodeser de midterste vælgere. Her kan de midterste vælgere symboliseres ved de midterste indkomster, som prioriteres mest på bekostning af de fattigste og de rigeste i beskæftigelse.⁴⁶ På den anden side er der også den mulighed, at skattesystemet blot har udviklet sig uden tanke om valgbarhed, og at den brede middelklasse er et produkt af det skattesystem, som eksisterer og ikke omvendt. Samlet set er det svært at konkludere, om de ikke-faldende velfærdsvægte er et bevidst valg – og derfor optimalt – eller om det oprindeligt ikke var hensigten at vægte mellemindkomsterne højest, og valget derfor er inoptimalt.

Resultaterne viser desuden, at såfremt deltagelseselasticiteten antages homogen på tværs af indkomst $S_2(H)$, forsvinder den inverse U-formede profil, og velfærdsvægtene for de øverste percentiler bliver negative, hvilket er i strid med 2. kriterium. Dette eksempel, hvor der også er deltagelseeffekter i toppen af indkomstfordelingen, kan henledes til globaliseringsdebatten, hvor der i visse kredse argumenteres kraftigt for, at et lavere skattetryk i Danmark vil øge chancerne for at rekruttere arbejdskraft fra udlandet.⁴⁷

Arbejdsudbuddet på den intensive margin har også stor betydning for fordelingen af velfærdsvægtene. Ved at øge timeelasticiteten fra S_2 til S_4 , vægtes lav- og mellemindkomsterne højere, mens de højeste indkomster vægtes mindre, endog negativt, hvilket er strid med 2. kriterium.

Der kan være to grunde til, at der fremkommer negative velfærdsvægte. For det første kan samfundsplanlæggeren have »negative« præferencer for de pågældende individer (»ikke-paretianske præferencer«). For det andet kan det eksisterende skattesystem være inoptimalt, således at der kan være plads til Paretoforbedringer.

Hvis det antages, at Samfundsplanlæggeren har »ikke-Paretianske præferencer«, så opnås en højere aggregeret velfærd, hvis indkomsten reduceres for de grupper, som

45. For deltagelseselasticiteter, der starter på 0,4, er der endda tale om negative velfærdsvægte for de laveste indkomster, hvilket også er i modstrid med 2. kriterium, hvilket der vendes tilbage til senere.

46. Ved at sammenligne »median voter« med middelindkomstgruppen antages dermed, at vælgerne primært ønsker, at deres egen indkomstgruppe vægtes så højt som muligt.

47. Her kan bl.a. nævnes Dansk Industri, de Konservative og Cepos mv. Det er her væsentligt at bemærke, at der vil være andre faktorer end blot deltagelseeffekten, som vil spille ind i forhold til rekruttering af arbejdskraft fra udlandet (bl.a. den skattebyrde, som den potentielle udenlandske arbejdstager initialt har mv.).

vægtes negativt. I S2(H) og S4 hvor det er de højeste løngrupper der vægtes negativt, kan de negative vægte henledes til, at lighed er et mål i sig selv. For S3, hvor de negative velfærdsvægte findes i bunden af indkomstfordelingen, giver det ingen mening at operere med »ikke-paretianske præferencer.«

Den mest realistiske årsag til forekomsten af de negative velfærdsvægte er formentlig, at skattesystemet ikke er optimalt. Denne inoptimalitet kan begrundes ud fra, at der kan foretages Paretoforbedringer. Herved forstås, at det er muligt at sænke skatten for individer med negative velfærdsvægte og derefter uddele provenuet lump sum til alle individer. På den måde opnås højere samlet velfærd, uden at nogen stilles dårligere end udgangspunktet. At der er mulighed for Paretoforbedringer i toppen af indkomstfordelingen, ses tydeligt ud fra de aggregerede adfærdsresponses, som er større end 1 (svarende til at skatterne er mere end selvfinansierende). Forklaringen på, hvorfor der forekommer negative velfærdsvægte for de laveste indkomster i S3, er imidlertid ikke helt så indlysende, som for de høje indkomster. Men forklaringen skal findes i profilen for de aggregerede adfærdsresponses (figur 2), som for S3 er faldende for de nederste percentiler (den mest almindelige tendens er stigende). Når profilen for adfærdsresponserne falder i indkomst, er det muligt at omrokere skattebetaling fra en indkomstgruppe i over til en højere indkomstgruppe $i + 1$ uden at forværre situationen for nogen. Derved stilles den gruppe, som opnår en lavere skattebetaling bedre, uden at nogen stilles værre.^{48, 49}

Samlet set kan det konkluderes, at hvis det danske skattesystem er optimalt, så har det danske samfund underliggende sociale præferencer, der hovedsageligt tilgodeser den brede middelklasse frem for de lavestlønnede på arbejdsmarkedet. Derudover er der i visse scenarier tale om, at nogle indkomstgrupper vægtes negativt. Der foreligger også den mulighed, at elasticiteterne generelt er sat for højt eller med en forkert profil, hvilket er med til at skabe usikkerhed om nøjagtigheden i velfærdsvægtene.⁵⁰ Den sidste mulighed er, at der findes inoptimalitet i det danske skattesystem. Dette kan begrundes ud fra eksistensen af individer, som vægtes negativt sammen med det faktum, at der er en stigende velfærdsvægtprofil i bunden af indkomstfordelingen, når der medtages heterogene deltagelseselasticiteter. Der er hermed tale om, at der i alle elasticitetsscenarier med undtagelse af bench mark tilfældet S1, er brud på mindst ét af de

48. Det er vigtigt, at den skattebyrde som $i + 1$ påføres (marginalskatten stiger), modsvarer den mindre skattebyrde, som er forårsaget af en skattesænkning længere nede i percentil i (gennemsnitsskatten falder), idet den samlede skattebyrde skal være uændret for dem, som oplever en marginalskatteændring (balancerede effekter). Dette er uddybet mere i Petersen (2007).

49. At adfærdsresponserne skal være faldende, er en nødvendig men ikke tilstrækkelig betingelse for, at der forekommer negative velfærdsvægte jf. yderligere Petersen (2007).

50. Elasticitetsscenarierne i denne artikel er generelt sat lavt, og der er lagt stor vægt på, at de stemmer overens med troværdige empiriske studier. Men der er som bekendt ingen, der kender de eksakte værdier, hvorfor analyser som omfatter arbejdsudbudselasticiteter, altid er behæftet med usikkerhed.

objektive kriterier for, hvorledes velfærdsvægtene bør se ud. Afslutningsvis kan det derfor konkluderes, at udledningen af de sociale velfærdsvægte bidrager til at så tvivl om, hvorvidt det danske skattesystem er optimalt.

Litteratur

- Blundell, R. W., A. Duncan og C. Meghir. 1998. Estimating labor supply responses using Tax Reforms. *Econometrica* 66, 827-61.
- Bourgunion og Spadaro. 2006. Tax-Benefit revealed Social Preferences. *Joint Research unit 8545 CNRS- EHESS-ENPC-ENS*.
- Cahue, P. og A. Zylberberg. 2004. Labor Economics. *The MIT Press*.
- Dahlby, B. 1998. Progressive Taxation and the Social Marginal Cost of public Funds. *Journal of Public Economics* 67, 105-122.
- Edgeworth, F. Y. 1897. The pure theory of taxation III. *Economic Journal* 7, 550-571.
- Eissa, N. og J. Liebman. 1996. Labor Supply Response to the Earned Income Tax Credit. *Quarterly Journal of Economics* 61, 605-37.
- Frederiksen, A, E. K. Graversen og N. Smith. 2001. Overtime work, dual job hoarding and taxation. *IZA Discussion Paper* 323, Bonn.
- Goolsbee, A. 2000. What Happens When You Tax the Rich? Evidence from Executive Compensation. *Journal of Political Economy* 108, 352-378.
- Graversen, E. K. 1996. Measuring Labor supply responses to tax changes by use of exogenous reforms. *CLS Working Paper* 97-09, Århus Universitet og Århus Handelshøjskole.
- Gruber, J. og E. og Saez. 2002. The Elasticity of Taxable Income: Evidence and Implications. *Journal of Public Economics* 84, 1-3.
- Kleven, H. J. og C. T. Kreiner. 2004. A note on Optimal Income Taxation.
- Kleven, H. J. og C. T. Kreiner, i T. Tranæs, red. 2006a. *Skat, arbejde og lighed – en undersøgelse af det danske skatte- og velfærdssystem*. Gyldendal.
- Kleven, H. J. og C. T. Kreiner. 2006b. Evaluating the Trade-Offs in the Tax and Transfer Policy: Theory and Danish Evidence. *Economic Policy Research Unit (EPRU)*.
- Kleven, H. J. og C. T. Kreiner. 2006c. The Marginal Cost of public Funds: Hours of Work Versus Labor force Participation. *Journal of Public Economics* 90, 2006.
- Ljunge, M. og K. Ragan. 2005. Labor supply and the tax reform of the century. *Arbejdsrapport, University of Chicago*, August 2005.
- Meyer, B. og D. Rosebaum. 2001. Welfare, the Earned income Tax Credit, and the Labor Supply of Single Mothers. *Quarterly Journal of Economics* 66, 1063-1114.
- Mirrlees, J. A. 1971. An Exploration in the theory of Optimum Income Taxation. *Review of Economic Studies*, 38: 175-208.
- Petersen, L. B. 2007. En analyse af omkostningerne ved det danske skattesystem- og de dertilhørende velfærdsvægte for forskellige indkomstgrupper. Kandidatafhandling ved Økonomisk Institut, Københavns Universitet.
- Triest, R. K. 1990. The effect of income taxation on labor supply in the United States. *Journal of Human Resources* 25, 491-516.
- Varian, Hal R. 1999. *Intermediate Microeconomics: A modern approach*, fifth edition. Norton.

A wage premium or penalty: An analysis of endogamous marriage effects among the children of immigrants?

Aycan Çelikaksoy

CIM and SOFI, Stockholm University, E-mail: aycan.celikaksoy@sofi.su.se

SUMMARY: This paper addresses the debate in the literature regarding marriage type and employment/wage premiums for immigrants, analyzing a cross-section of the children of guest worker immigrants in Denmark. For females, marrying someone from the same ethnic background, either a »migrant« spouse, who migrates for the purpose of marriage or an »immigrant« spouse who has already been living in Denmark, have no effect on their employment probability or wages. However, in the case of males, the results indicate that it is important to decompose endogamous marriages and investigate the effect of different marriage types. Marriage to an »immigrant« spouse has a positive impact on employment probability of males, while no effect is found on wages when spouse types are treated as exogenous variables.

1. Introduction

The two separate topics of »economic assimilation of immigrants« and »wage differentials due to marital status«¹, are of great interest to economists. However, studies at the junction of the two arrays of literature are very scarce. Most of the literature analyzing labour market integration of immigrants disregards the family's role in the assimilation process, whereas the literature on marital wage differentials mainly concentrates on white native males.

In most countries, Non-Western immigrants have lower employment and wage rates than natives and their economic assimilation is a crucial issue. However, guest worker immigrants in Europe are relatively new when compared to the immigrants in the U.S.

I am grateful to Helena Skyt Nielsen, Nina Smith, Eskil Wadensjö, Nabanita Datta Gupta, Shoshana Grossbard-Shechtman and Torben Tranæs for their comments. Furthermore, I thank the participants at the EALE conference for their comments. Financial support from The Strategic Program for Welfare Research, Project: »Integration of Ethnic Minorities« and the Graduate School for Integration, Production and Welfare, is gratefully acknowledged.

1. See Ribar (2004) for a review of the literature.

and their children are still quite young. Thus, the economic fate of the second and third generation immigrants in Western European countries is yet to be decided. Several studies in social sciences emphasize the importance of marital sorting for later social and economic outcomes. Thus, how these young immigrants sort themselves into household units through marriage may have very important implications for the integration of the current generation and their descendants.²

This paper analyzes the relationship between the spouse type of the children of guest worker immigrants in Denmark and their labour market outcomes measured by their employment probabilities and wages. Rather than looking at marital status only, I investigate whether marrying a certain type of spouse has an effect on immigrant wages.

All of the very few studies done on this issue divide marriage type into two categories, namely endogamous (marriages within the same ethnicity) and exogamous marriages (cross ethnicity marriages). What I do in this paper is to divide the endogamous marriages further into two categories and analyze them separately. Thus, immigrants have the following spouse type options: (1) imported spouse from country of origin (»migrant« spouse/marriage migration) and (2) immigrant spouse who has already been living in the host country, Denmark (»immigrant« spouse), where in both cases the spouses are from the same country of origin.

Since most societies face an increase of immigrants whose legal permits to enter the respective country are through marriage, the importance of the issue of marriage migration and its effect on the young immigrants in the destination countries is increasing. In the case of exogamous marriages, most of the sociological literature denotes interethnic marriage (marriage between immigrants and natives) as the final stage of immigrant assimilation, Gordon (1964); Qian (1999) and Qian et al. (2001). However, the direction of causality is an empirical question. So far, there have been very few attempts to answer this question empirically and an agreement has not been reached regarding the direction of the causality, Meng and Gregory (2005); Kantarevic (2004); Meng and Meurs (2006) and Duncan and Trejo (2005).

This article mainly investigates, whether the choice of bringing a spouse from country of origin has an effect on the individual's labour market outcomes, which is an issue that has never been studied before. The current analysis fails to find any association between spouse type and labour market outcomes for women, which is a finding that confirms with both the more general international literature on the relationship between marriage and labour supply decisions and wages of women and the more specific literature on the effect of endogamous marriage on immigrant women's wages. However, the results for males are not straightforward. There is a marriage penalty for

2. See Duncan and Trejo (2006) for differences in outcomes for children who are the product of either endogamous or exogamous marriages.

males for marrying a »migrant« spouse as opposed to an »immigrant« spouse for their employment probability, whereas the effect of spouse type on male wages requires further investigation due to the endogeneity of spouse type choice for males.

The paper is organized as follows: Section 2 presents the previous research. Section 3 describes the dataset and variables used. Section 4 presents the empirical analyses to test the main hypotheses. Section 5 concludes.

2. Previous research

There are a large number of studies regarding the relationship between marital status and wages, analyzing wage differentials between married men and never married men. While most of these studies investigate the issue for white men, few of them study women or African American men. Although the relationship between female wages and marriage is not very clear, all of the literature finds a positive relationship between male wages and marriage, and some of these studies establish a causal relationship in the latter case. Thus, a large proportion of this casual relationship is attributed to productivity differences between married and single men rather than discrimination or different demands for job attributes, Kermit (1991); Grossbard-Shechtman, (1993). There is also a growing literature trying to identify the reasons for these productivity differences, such as intra-household specialization or other productivity increases for men due to marriage, utilizing various datasets such as time use data or employer-employee records in addition to panel register data, Rodgers and Stratton (2005); Hersch and Stratton (2000).

The existence of a male marital wage premium has primarily been based on two theories: The first one is the selection effect, where it is argued that more productive men are more likely to be married and that marriage itself does not have an effect on the productivity of the individual, Loh (1996). The second is the productivity effect, where it is argued that marriage makes men more productive in the labour market due to various reasons; the most important one being the intrahousehold specialization, Korenman and Neumark (1991).³

In the case of immigrants, the wage differentials are mainly analyzed by comparing immigrant and native wages or the wages of different cohorts of immigrants to investigate the degree and the rate of economic assimilation for different groups of immigrants.⁴ The literature on economic assimilation mainly concentrates on the effects of number of years lived in the destination country and finds a positive correlation between the two. One well known reason for this relationship is the language proficiency improvements of immigrants. However, we still know very little about the factors that affect

3. See Becker (1973) for the theoretical underpinnings.

4. See Borjas (1999) for the definitions of economic assimilation and a review of the literature.

the assimilation process of immigrants and, surprisingly enough, the role of the family has been mostly ignored in this literature.

However, recent studies have started to turn their attention to the role of the family in relation to assimilation, where labour market decisions take place within the family context. Incorporating the family into the economics of integration, Duleep and Sanders (1993) argue that family investment strategy is an important determinant of immigrant women's labour force participation in addition to language proficiency and years since migration. They argue that whether a husband invests in human capital specific to the US labour market and also the extent of that investment, are factors affecting a woman's decision to work. Baker and Benjamin (1997) also argue that (for Canada) the family investment model explains the gender assimilation differences as immigrant wives take on »dead-end« jobs to finance their husbands' investments in human capital. They conclude that family composition is an important correlate of immigrants' assimilation, and the family investment model can account for many of the patterns in the data.

Blau et al. (2003) estimate annual hours and log wages according to family type: immigrant family and mixed family, where »immigrant family« refers to families with two immigrant spouses and »mixed family« refers to families, where one spouse is an immigrant and the other a native. In contrast to Baker and Benjamin, Blau et al. (2003) find that husbands and wives in both types of families are investing equally in their human capital. However, for both husbands and wives, the shortfall in hours and the wages shortfall at arrival are less for spouses in mixed families than for immigrant families. This reflects a lower initial level of human capital and labour market knowledge among spouses in immigrant families than in mixed families. However, most of this result disappears when they change the age at migration from 16 to 18. They mainly find similar patterns for both types of families and they find no support for the family investment model or an effect of marriage type in their study.

The sociological literature accepts mixed marriages (an immigrant marrying a native) as the final stage of assimilation, Klein (2001); Gordon (1964); Hutchinson (1957); Qian et al. (2001); Kitano & Yeung (1982). The research by Card et al. (2000) that deals with this issue as a small part of the paper also takes mixed marriages as an indicator of a high degree of assimilation. There are very few studies that investigate the relationship between marriage/spouse type and economic assimilation.

Kantarevic (2004) finds, for the US case, that immigrant men who are married to native-born spouses assimilate faster than comparable male immigrants married to foreign-born spouses. However, when he accounts for selection, this effect vanishes and he concludes that immigrants who are married to native-born spouses are a favourably selected sub-sample of the population of all married immigrants, which mainly

supports the sociological literature arguing that more assimilated immigrants marry natives⁵ rather than the spouse type having an effect on the assimilation outcomes of the immigrants. In contrast, Meng and Gregory (2005) and Meng and Meurs (2006) argue that intermarriage is an important input into the economic and social assimilation of immigrants in the host country. They find that there is an interethnic marriage premium after accounting for human capital differences among individuals for both male and female immigrants.

In summary the more general literature on marriage and wages agrees that there is a marital wage premium for men, whereas there is no agreement in the case of women. However, mostly it is found that after taking account of children and labour force attachment, there is little or no association between marital status and wages for females. The findings of the more specific literature on immigrants are more diverse. The first group of studies analyses the family investment hypothesis, that is, whether there is a difference in the assimilation profiles of husbands and wives in mixed versus immigrant families. Support for the family investment model is found for Canada but not for the U.S. However, these studies do not tackle the issues of selection into different types of marriages or into employment. The second group of studies focus on the causal relationship between spouse type and wages. Kantarevic (2004), in his study of male immigrants in the U.S., argues that there is no causal relationship between spouse type and wages and that the differences are due to selection into different spouse types. Using instrumental variable techniques, Meng and Gregory (2005) find that there is an intermarriage wage premium for both male and female immigrants in Australia, where Meng and Meurs (2006) find a similar result for immigrants in France. However, the only finding regarding the effect of endogamous marriage on female wages is that there is no effect.

2.1 Types of endogamous marriage and wages

This article combines the two bodies of literature mentioned in the previous section and extends those studies by using a register-based dataset and by specifically analysing a narrow group of immigrants, the children of guest workers, who have lived all or most of their lives in Denmark and followed the Danish educational system from the beginning. Thus, the sample is clear of such effects as years since migration or language proficiency.

Furthermore, the focus of this paper is on marriages within the same ethnicity (endogamous marriages) rather than a comparison of endogamous and exogamous

5. The definition of a »mixed marriage« to a »native« varies in the literature. For instance, Kantarevic (2004) defines it as marriage to native-born individuals although they may be immigrants born in the destination country (the US) and Meng and Gregory (2005) define it as marriage to English-speaking individuals although they may be immigrants born inside or outside the destination country (Australia).

marriages, as discussed in the international literature. There are two reasons for this. First of all, the sample size for intermarried immigrants in the current data set (3%) is not sufficient to carry out an analysis. Secondly, migration from Non-Western countries due to marriage has been a heavily debated issue in Danish politics, although there has been no empirical evidence regarding the issue. The two types of endogamous marriages analysed are: (1) individuals married to »migrant« spouses and individuals married to »immigrant« spouses. The first refers to spouses who come to Denmark for the purpose of marriage to an immigrant who is from the same country of origin (immigrant to migrant marriage), whereas the latter refers to marriages between two immigrants from the same country of origin, but where both have been living in Denmark (immigrant to immigrant marriage). These two types of marriages constitute the whole group of endogamously married individuals, which constitute over 95% of all marriages for the group of interest. Thus, I intend to analyze whether a further decomposition of the group of endogamously married individuals is necessary in the relevant analyses and investigate the importance of this issue in integration/migration studies. Hence, the first part of the paper tests whether the relationship between marriage and labour market outcomes varies by spouse type and sheds light on the direction of the observed relationship.

According to the discussion in the previous section, the observed relationship between spouse type and labour market outcomes may be either due to the different productivity enhancing effects of the different spouse/marriage types or due to the unobserved characteristics of the individuals that lead them to be sorted into different marriage types. On the other hand, there might simply be no differences in the integration process of individuals within different marriages, as Blau et al. (2003) argue.

But why might one expect different results with regard to spouse type? Following the productivity hypothesis, the specialization explanation predicts that the marriage premium for having a »migrant« spouse is higher than the premium for having an »immigrant« spouse for men and vice versa for women. That is to say that if we expect the »migrant« spouses arriving from the country of origin to have relatively more traditional preferences with regard to intra-household and labour market specialization, such that women expecting to become housewives and men expecting to be the traditional breadwinners of their family after arriving to Denmark, then the above argument should hold. However, I find no evidence for this explanation in the data. On the other hand, following the networks explanation of the cross-productivity effects within marriage, one would expect the premium difference between spouse importers and non-importers to be negative for both men and women. This is simply due to the fact that marrying someone from outside the country will not lead to the same benefits as marrying a spouse who already has his/her networks in Denmark. Marrying an immigrant spouse

already living in Denmark may lead to the formation of a new network (especially strong ties) for the individual leading to special contacts and connections due to the increase of new countrymen in one's social circle improving his/her job prospects. In addition, marrying a »migrant« may lead the family to socialize only within the specific immigrant group due to the spouse's lack of language and cultural barriers leading one to lose the existing networks.

The family investment model predicts that individuals who bring their spouses from country of origin take on the »borrower« role to finance and allow the migrating spouse to invest in host country specific human capital such as language and knowledge about culture and institutions. Individuals importing their spouse from another country have the legal responsibility to financially support their spouses, who have restricted rights to welfare benefits. In Denmark, after 1999 it is even required by law that newly arrived immigrants participate in an integration course of 36 months' duration, thus automatically making the »migrant« spouse the »investor« and her/his spouse the »borrower«. Thus, this may lead the young immigrant who has grown up in the host country to stop investing in his/her own human capital and seek employment. Hence this might reflect as taking on »dead-end« jobs and lower wages when compared to their counterparts who marry »immigrant« spouses instead and thus, are less credit constrained.

On the other hand, the selection effect hypothesis postulates that spouse type does not have an effect on the labour market outcomes even though the two may be related. This relationship may be due to some unobservable characteristics that affect both the labour market outcomes and spouse choice outcomes of the individual.

Although, neither the current analysis nor the data at hand is sufficient to clearly identify the exact mechanisms behind the observed relationships between spouse type and labour market outcomes, the above mentioned frameworks will be utilised to shed light on the issue for the children of guest worker immigrants in Denmark.

3. Data

The data used originates from the administrative registers of Statistics Denmark. It includes the total population of immigrants and descendents in Denmark in 2001. The focus of this paper is on the children of »guest workers«. Denmark imported its »guest workers« mainly from Turkey, Pakistan and Ex-Yugoslavia in the 1960s and early 1970s. The immigrant flows from Ex-Yugoslavia in the 1990s have mainly consisted of refugees. As refugees are quite a different group than that of guest workers, they have been excluded from the analysis.

The sample used in this study consists of children of guest workers, namely »middle generation« immigrants (children who are born outside Denmark but immigrated be-

fore the age of eight, i.e. they have attended primary school in Denmark) and »second generation« immigrants (children who are born in Denmark to immigrant parents), where 74% of this group are still younger than 18 in 2001. For the analysis of marriage decisions and labour market outcomes, I have selected a sample aged 18-35 years in 2001, where 37% are married. Marriage includes first and legal marriages, as cohabitation is culturally very unacceptable for the groups studied. Individuals who are divorced, widowed, interethnically married and who have spouses with missing country of origin information have been deleted from the sample. This constitutes only 6% of the sample. Furthermore, students are excluded from the sample, which constituted 40% of the sample. After deleting these individuals and individuals with missing experience information, there are 1,179 males and 914 females in the sample.

The dataset contains information on a large number of demographic and labour market characteristics of the individuals. The individuals in the sample have the following spouse type options: (1) imported spouse from country of origin (»migrant« spouse)⁶ and (2) »immigrant« spouse who has already been living in the host country (Denmark). The reason for this distinction is to allow an investigation of whether those two spouse types have different effects on the labour market outcomes of the children of immigrants.

3.1 Spouse type among the children of guest worker immigrants

Before deleting exogamous marriages from the data, one can see that 95% of males and 97% of females are married endogamously, among the married ones. A further decomposition of marriage type shows the following figures for males: (1) 82% are married to »migrant« spouses (2) 13% are married to »immigrant« spouses« (where these two types constitute endogamous marriages analysed separately in this paper) (3) 3% are married to spouses from different ethnicities (4) 2% are married to native Danes (where these two latter categories constitute exogamous marriages and are deleted from the sample for reasons stated above); the figures for females are: (1) 79% (2) 17% (3) 2% (4) 1%.

The initial guest workers who migrated to Western European countries in the 1960s and 1970s were mainly men; as a result it is not surprising that the first generation brought their existing wives or fiancées to the destination countries. Thus, I am mainly interested in the marriage behaviour of the children of guest workers. These groups of immigrants have lived all or almost all their lives in Denmark, have followed the Danish educational system from the beginning and are expected to be closer to the Danish culture than their parents. As a result, it could be reasonable to expect relative-

6. A person is defined as a »migrant« spouse or (imported spouse) if he or she has migrated in the period of two years before or after his/her marriage date.

ly more children of immigrants to marry Danes and/or each other. However, this is not quite the case. Most of the marriages that take place in most Western European countries are to spouses who arrive for marriage purposes,⁷ which has made this type of marriage an important issue in the media and the political debate in these countries, making the various consequences of such marriages an important topic to analyze. The rates of exogamous marriages above are very different from the figures for France (14% in 1999), Australia (32% in 1996) and the US (38% in 1980), where exogamous marriage rates for immigrants are higher. However, endogamous marriages are by far the most common type of marriage for immigrant groups in the world. In spite of this fact, there is very little known about this phenomenon.

Around 86% of endogamously married males are married to »migrant« spouses, where the figure is 82% for females. Thus, it is interesting that these children choose to bring their spouses from their country of origin rather than marrying someone from the same ethnic origin but who has been living in Denmark and who shares the same experiences and culture as themselves.

There are various possible reasons for this phenomenon, where one of the most important reasons is expected to be the sex ratio imbalances in the marriage market. It is generally defined as the number of men divided by the number of women and it is one of the most important determinants of how people sort themselves into marriages in that marriage market, as has been argued⁸ by most sociologists, demographers and economists. It becomes a crucial factor especially when studying ethnic marriages.

In the literature of interethnic marriages it is argued that there are two main factors that determine the rate of intermarriages; relative group size and sex ratio. The rationale behind this is quite simple, as the group size gets smaller, that is, the size of an ethnic group in the destination country gets smaller, the higher is the probability that they will marry natives (or other immigrants). The rationale behind sex ratio imbalances is similar as well. It is expected that as the sex ratio for males gets low, that is, the number of marriageable females to males within the group gets relatively low, the more likely are the males in that group to marry outside that ethnic group. In this paper, as described before, intermarriage across ethnic groups is excluded from the analysis as there are very few observations. However, I use the same rationale with regard to sex ratio and relative group size within the context of marrying »migrants«, as these are marriages that take place outside the local marriage market, ie, Denmark. Thus, in this case immigrants have the option of bringing their spouses from their country of origin rather than marrying native Danes, when faced with small relative group sizes or sex ratio imbalances.

7. See Çelikaksoy, 2003.

8. See Groves and Ogburn (1928) and Cox (1940) for the earliest empirical studies.

Table 1. Sample means for the children of guest workers by gender and spouse type. (Age group 18-35, students excluded) – 2001.

variables	spouse type					
	for males			for females		
	single	»migrant«	»immigrant«	single	»migrant«	»immigrant«
country of origin (Turkey) (0/1)	0.63	0.72	0.75	0.63	0.73	0.77
middle generation (0/1)	0.27	0.64	0.53	0.25	0.54	0.48
employed (0/1)	0.54	0.51	0.59	0.56	0.52	0.47
log hourly wage	4.86	5.00	5.07	4.75	4.85	4.81
experience (nb. of years)	1.40	6.15	6.32	1.18	4.49	3.35
experience squared/100	0.69	7.13	7.88	0.51	4.23	2.58
education (nb. of years)	9.98	10.11	10.58	10.47	10.31	9.89
age	22.98	27.11	27.15	22.50	26.36	26.00
age squared/100	5.39	7.49	7.48	5.18	7.08	6.90
nb. of children (age ≤ 6)	0.03	1.07	1.10	0.07	1.08	1.16
nb. of children (6 < age ≤ 17)	0.00	0.42	0.25	0.02	0.44	0.42
sex ratio	1.02	1.01	1.02	0.98	0.98	0.97
relative group size*100	1.04	0.89	0.91	0.99	0.86	0.90
Number of obs.	704	419	56	414	403	97

Table 1 shows sample means with regard to sex and spouse type for the children of guest worker immigrants who are aged 18-35 years. The variables showing whether the individual is »employed« or not, »country of origin« and type of immigrant are indicator variables. The variable »experience« shows the number of years of labour market experience the individual has accumulated, where »education« shows the total number of years of completed schooling and post schooling. The variables related to children show the individual’s number of children at the specified ages. The variables »sex ratio« (i.e. sex ratio for men is defined as the number of women to the number of men within the relevant group and vice versa for women) and »relative group size« (i.e. relative group size for men is defined as the number of women within the relevant group to the total number of women in Denmark) are explained in detail in the next section.

Looking at the male sample, it can be seen that, on average, married men have higher levels of education and experience as well as higher rates of log hourly wages than single males, which might partly be due to their older age on average. Whereas in the case of women, single women have a higher employment rate than married women on average and they have an average of 0.37 years longer completed level of education when compared to married women although they are younger on average. Looking at the raw figures in Table 1 it can be seen that the marital wage premium for males is

only 4%. This might seem contradictory to the figures of marital wage premium of around 10%-30% for white Western men however; marital wage premium is expected to be lower for immigrant men who are endogamously married. The comparable figures for Australia and the US are 7% and 14% respectively.

Looking at males who are married to »migrant« and »immigrant« spouses separately, it can be seen that males who are married to »immigrant« spouses have, on average, higher employment rates, wage, education and experience levels. There is only a 1% marital wage premium for having an »immigrant« spouse compared to having a »migrant« spouse. This difference is low when compared to, for instance, an exogamous marriage premium of 5% as opposed to endogamous marriages in the US. However, this paper focuses on the differences within endogamous marriages. Furthermore, it is an empirical question whether this premium is a return for higher human capital and how it changes after controlling for other factors. In the case of females, on the other hand, the ones who are married to »migrant« spouses have higher levels of education, experience, employment and wage rates on average when compared to the ones who are married to »immigrant« spouses and there is a 1% wage premium for having a »migrant« spouse compared to having an »immigrant« spouse.

In summary the premia difference for average log hourly wages by spouse type is in different directions for males and females and the observed differences in the raw data are very small. These issues will be investigated further in the next section.

4. Empirical strategy

4.1 Employment

Variable log hourly wages can only be observed when the individual is employed as a wage or salary earner. Hence, the wage equations that will be estimated may be biased due to a sample selection problem. In addition, due to the highly compressed wage structure and a generous welfare benefit system in Denmark, being employed is a very crucial indicator of the assimilation process of immigrants. Thus, it is important to look at the determinants of being employed first and to analyze the relationship between spouse type and being employed.

In this section, the hypothesis of whether marriage type has an effect on employment probabilities for the children of immigrants is tested. To do this, I start by estimating a simple probit model for the probability of being employed and I test whether the spouse types for both men and women are statistically significant. In the probit model, I also control for other variables, either exogenous variables or variables that may be endogenous in the employment equation. As discussed before, I decompose endogamous marriages further into two categories, namely, marriages to »migrants«

Table 2. Marginal effects from estimation of the probability of being employed - 2001. Probit specification. (Std. err. in parantheses).

	Whole sample		Sample 1	Sample 2	Sample 3
	Males	Females	spouse type: »migrant« vs. single	spouse type: »immigrant« vs. single	spouse type: »migrant« vs. »immigrant«
experience	0.164*** (0.012)	0.189*** (0.015)	0.168*** (0.012)	0.274*** (0.021)	0.099*** (0.014)
experience squared	-0.055*** (0.005)	-0.066*** (0.006)	-0.057*** (0.005)	-0.099*** (0.009)	-0.031*** (0.005)
education	0.027*** (0.006)	0.023*** (0.007)	0.030*** (0.006)	0.025*** (0.008)	0.024*** (0.007)
<i>Spouse type:</i>					
marriage migrant	-0.014 (0.053)	-0.064 (0.059)	-0.026 (0.054)		-0.137** (0.067)
immigrant	0.156* (0.082)	-0.021 (0.076)		0.141 (0.118)	
age	0.010 (0.043)	0.079 (0.051)	0.003 (0.044)	0.050 (0.060)	-0.008 (0.080)
age squared	-0.084 (0.083)	-0.213** (0.098)	-0.071 (0.086)	-0.212* (0.121)	-0.025 (0.146)
number of children (age <= 6)	-0.039 (0.030)	-0.159*** (0.034)	-0.031 (0.031)	-0.015 (0.060)	-0.055* (0.031)
number of children (6 < age <= 17)	-0.066** (0.030)	-0.021 (0.035)	-0.067** (0.032)	0.081 (0.080)	-0.054* (0.030)
Wald F-test for migrant = immigrant prob.	4.17 0.04	0.45 0.50			
# observations	1179	914	1123	760	475
Hausman test for endogeneity of spouse type (P-value)	0.520	0.147			
LR-test for rho = 0 (prob.)	0.100	0.913			

Note: (1) *** indicates significance at the one percent level, and ** and * at the 5% and 10% levels respectively. (2) Sample 1 includes only singles and individuals who married migrants; sample 2 includes only singles and individuals who married immigrants; sample 3 excludes singles and compares individuals who have migrant spouses versus immigrant spouses. (3) An indicator variable is added for missing education information.

and »immigrants«, to analyze the effect of spouse type in more detail and to see whether studying spouse type in broad categories is misleading.

Table 2 presents the marginal effects (evaluated at sample means) from probit estimations of the probability of being employed. The first two columns are the results for

the total sample, while columns 3, 4 and 5 presents the estimation results for different samples, that have been constructed to be able to conduct pair-wise comparisons, as will be discussed further later in this section. Sample 1 excludes individuals married to »immigrant« spouses and compares individuals married to »migrant« spouses with singles; Sample 2 excludes individuals married to »migrant« spouses, where Sample 3 excludes singles and compares individuals married to »migrants« as opposed to the ones married to »immigrants«.

Looking at the first two columns, it can be seen that the size of the marginal effects of labour market experience is large and significant at the one percent level for both males and females. An additional year of experience increases the probability of being employed by 16 percentage points for males and by 19 percentage points for females. The larger size of the effect for females even after controlling for education and number of children might be due to the relative importance of labour market attachment for women's employment decisions. Education clearly exerts a positive impact on employment behaviour of both males and females, where the effect size is larger for males. This might be partly due to lower rates of return for education for females compared to males, hence effecting their participation in employment. The descriptives of the raw data also show that log hourly wages for males are 4% higher on average than that of females although both have on average 10 years of education.

For females, an additional child in the age category 0-6 reduces the employment probability by 16 percentage points on average, while it has no effect on males. This is a common finding in the international literature and thus, career interruptions for females are also a common partial explanation for lower rates of return to education for females relative to that for males'. However, the negative impact of having older children for males is unexpected. It can be seen in the descriptive statistics that males who are married to »migrant« spouses, compared to the ones married to »immigrant« spouses, have more children on average in the age category 7-17. For an average male to have a child at the age of 12 for instance the father has to be 15 at birth. Thus this might be a negative effect of early child bearing for males who are married to »migrant« spouses. Looking at the next three columns also show that this is a phenomenon for males who are married to »migrant« women.

Finally, turning attention to the main focus of this paper, it can be seen that for males having an »immigrant« spouse increases their employment probability by 16 percentage points, which is an effect that is significant at the 10 percent level. While, the sign of the effect of having a »migrant« spouse is negative, it is not statistically different from zero. This is also the case for both of the spouse type effects for females. Furthermore, for females I cannot reject the hypothesis that the effects of the two marriage types are not statistically different from that of staying single. Hence, there is no

statistical justification for a further decomposition of endogamous marriage for females or of an effect of endogamous marriage on women's propensity to be employed as opposed to staying single. Also, Meng and Meurs (2006) find no evidence of an effect of marriage type on labour force participation of immigrant women in France. However, for males the Wald F-test indicates that the difference between the two marriage types is significant, which justifies the decomposition. Thus, the pairwise comparisons have been carried out only for males. It is clear in sample 3, which compares the effect of having a »migrant« spouse as opposed to having an »immigrant« spouse, that the former has a negative effect of 14 percentage points on males' employment probability.

However, it is important to consider the possibility that unobserved characteristics or omitted variables in the employment and spouse choice equations are correlated. Certain unobservable characteristics might appeal to both employers and certain types of potential spouses (such as cultural assimilation or good looks). However, it can be seen in Table 2 that I cannot reject the hypothesis that spouse type is exogenous in the employment equation. Both the Hausman test and the likelihood ratio test for the correlation of error terms of the bivariate probit model that simultaneously estimate employment decisions and spouse type choices endogenizing the spouse types, indicates that marriage types can be treated as exogenous variables. Thus, the simple probit model is preferred over instrumental variables methods. Two instruments namely, sex-ratio and relative group size are used as instruments in these tests. The instruments used for the tests are explained below in this section in detail.

This section concludes that it is important to analyze the effect of two types of marriages separately for males and that marrying an »immigrant« spouse increases men's employment probability. However the distinction of the marriage types in the case of females does not add any input to our understanding of the issue and I fail to find any association between marriage type and employment probabilities of females. The finding for males might simply be due to the well documented importance of network effects on employment for immigrants. Thus, family and kin networks can be seen as the closest network circle for information and screening employment opportunities. In this case marrying an insider (»immigrant« spouse), who knows the language, culture and has her own established networks in the country should impose a premium for her husband's employment opportunities. However, we do not observe such an effect in the case of females.

Instruments:

Regarding the instruments, they must fulfil two criteria: First of all, they must be strong, i.e. they should exert a strong influence on the endogenous variable, and they

must be valid, i.e. they should only influence the outcome variable through the endogenous variable. Validity is not testable, however it is difficult to imagine that an individual's employment or wages would depend on the number of individuals of the opposite sex in the same age-ethnic group in that year as a ratio of the total stock of the opposite sex of marriageable age. The same rationale goes for the sex ratios.

The variables that are used in creating the instruments are available for the period 1984-2001: These are country of origin, sex, and age for the whole population of Turkish and Pakistani immigrants. The source of this dataset is the same as the one explained in Section 3. The other variables, ie, the total number of males and females for the relevant age group in Denmark in each year have been taken from the official page of Denmark Statistics. The instruments are as follows:

1. Country of origin, age and gender-specific sex ratios:

$$SR_{it}^m = \frac{n_t^f}{n_t^m} \text{ for males and } SR_{it}^f = \frac{n_t^m}{n_t^f} \text{ for females}$$

2. Country of origin, age and gender-specific relative group size:

$$RF_{it}^m(M) = \frac{n_t^f}{N_t^f} \text{ for males and } RF_{it}^f(M) = \frac{n_t^m}{N_t^m} \text{ for females}$$

The instruments are computed for the year when the individual is 18 years old to have the most relevant figures for each individual. To be more precise, in this paper sex ratios are defined, such that sex ratio for an individual who is male and from Turkey is:

$$SR_t^m = \frac{n_t^f}{n_t^m}, \text{ where } n_t^f \text{ is the number of females in Denmark, who are originally}$$

from Turkey and aged 18-25 years at year t , and n_t^m is the number of males in Denmark, who are originally from Turkey and aged 18-25 years at year t , and vice versa for a female where year t is the year when the individual is 18. Thus, sex ratio measures the competition in a certain ethnicity and age specific marriage market for males (or females). The larger this ratio the less severe the competition is within the group and

the less likely the individual will search for partners outside the local marriage market, that is, in this case, »migrant« spouses.⁹

Relative group size (relative frequency) is created the same way, where it is defined with regard to country of origin, sex and age, such that the relative group size variable for a male from Turkey is:

$$RF_{it}^m (M) = \frac{n_t^f}{N_t^f} , \text{ where } n_t^f \text{ is the number of females in Denmark who are}$$

originally from Turkey and within the age group 18-25 at year t , and N_t^f is the total number of females in Denmark, within the age group 18-25 at year t , that is, when the individual is 18 years old. The larger the value of this variable the more likely the individual to meet a partner from the same ethnic-age group within the local marriage market, that is an »immigrant« spouse from Denmark.

4.2 Wages

This section presents results regarding the relationship between spouse type and wages. Table 3 reports the estimates from OLS cross-sectional log wage regressions for the whole sample and for the three different samples. Looking at the first column, it can be seen that there is no association between male wages and spouse type. Furthermore, the Wald F -test also indicates that being married to either type of spouse is not statistically different from staying single in terms of wages. This finding contradicts the general literature for white Western males, who have around 10% to 30% wage premium due to marriage. However, there is only one study that finds a marital wage premium (9%) for male immigrants for being endogamously married, which is for Australia, Meng and Gregory (2005). Thus, given the highly compressed wage structure in Denmark, the different findings for Australia and Denmark might be expected.

Also, in the case of females, none of the spouse type variables are significant and the formal tests indicate that the variables are neither significantly different from each other nor from the state of being single. The only finding in the literature regarding the relationship between wages of immigrant females and endogamous marriages is that there is a marriage penalty of around 7% as opposed to the state of being single, while this effect disappears after endogenizing marriage type, Meng and Gregory (2005).

The main finding of Table 3 is that, in spite of the expected differences in wage outcomes by spouse type due to the discussions in Section 2, the current OLS estimations

9. See Angrist (2002) and Meng and Gregory (2005).

Table 3. Estimates of Ordinary Least Squares Regression for log hourly wages. 2001. (Std. err. in parantheses).

	Whole sample		Sample 1	Sample 2	Sample 3
	Males	Females	spouse type: »migrant« vs. single	spouse type: »immigrant« vs. single	spouse type: »migrant« vs. »immigrant«
experience	0.079*** (0.010)	0.057*** (0.012)	0.079*** (0.010)	0.097*** (0.015)	0.066*** (0.012)
experience squared	-0.026*** (0.004)	-0.019*** (0.005)	-0.026*** (0.004)	-0.031*** (0.007)	-0.020*** (0.005)
education	0.028*** (0.006)	0.010 (0.007)	0.028*** (0.006)	0.021** (0.009)	0.033*** (0.007)
<i>Spouse type:</i>					
migrant	0.017 (0.041)	0.072 (0.046)	0.021 (0.041)		-0.063 (0.068)
immigrant	0.074 (0.070)	0.039 (0.069)		0.025 (0.078)	
constant	4.378*** (0.061)	4.469*** (0.072)	4.378*** (0.062)	4.414*** (0.088)	4.470*** (0.103)
Wald <i>F</i> -test for migrant = immigrant prob.	0.68	0.26		0.41	0.61
# observations	641	473	602	393	287
Hausman test for endogeneity of spouse type (P-value)	0.006	0.737			

Note: (1) *** indicates significance at the one percent level, and ** and * at the 5% and 10% levels respectively. (2) Sample 1 includes only singles and individuals who married migrants; sample 2 includes only singles and individuals who married immigrants; sample 3 excludes singles and compares individuals who have migrant spouses versus immigrant spouses. (3) An indicator variable is added for missing education information.

fail to find any association between the different marriage types and wage outcomes for male and female immigrants. Considering the very small differences by spouse type (1%) in the raw data for average log hourly wages for both males and females, this might be an expected result. However, there are two main potential problems with the current OLS estimates. The first one is related to selection into employment. As was mentioned earlier wages are observed only for the individuals who are employed, which is 54% of males and 52% of females in the sample. Hence, as being employed or not is potentially non-random, there might be a serious sample selection bias in the OLS estimates of the wage regressions. This potential problem can be accommodated by using the two-step procedure developed by Heckman

(1979). Thus, a probit model for the probability of being employed is estimated, and the estimates from this model are used to construct a sample selection correction term (IMR) Inverse Mill's ratio,¹⁰ where in principle, the inclusion of this term ensures consistent estimates. The probit model estimated is the same as the one in Table 2 and the exclusion restrictions assumed to identify the employment model are age, age squared and the number of children in different age groups.

The second potential problem regarding the OLS estimates is related to the endogeneity of the spouse types if the decision to marry and the choice between marrying a »migrant« or an »immigrant« depends, in part, on the potential earnings within each state/marriage type or on the other hand if there is either a positive or a negative correlation between some unobservable characteristics that jointly have an effect on wages and spouse type choices, such as tastes, cultural assimilation, looks, etc. Hausman specification test statistics presented in Table 3 indicate that exogeneity of spouse types for males is rejected while it is not rejected for females. Thus, simple OLS is preferred over IV estimations for the female sample.

Endogenizing a set of mutually exclusive indicator variables is not straightforward. Thus, I have chosen to focus on pair-wise comparisons, which reduces the problem to one of endogenizing one binary indicator variable, which is straightforward.¹¹ Thus, three samples are constructed for this purpose as discussed earlier in the paper. To deal with the above mentioned problems a Heckman two-step estimation is used, where the coefficients for the spouse type variables, capture the average effect of each spouse type on wages. The instruments, »sex ratio« and »relative group size« are used to identify the reduced form equations. In addition, the term IMR is added to both stages to take account of selection into employment, as discussed above. Table 1A in the Appendix reports the results for males.

Looking at the determinants of the two types of marriages, that is, the second column in each sample, it can be seen that education and experience have a large positive and significant effect on the likelihood of having married a »migrant« spouse compared to both staying single (sample 1) and having married an »immigrant« spouse (sample 3). On the other hand, these coefficients are not significant for the choice of an »immigrant« spouse. Furthermore, the coefficients for the selection correction term (IMR) for both stages of the wage equations are positive and significant in samples 1 and 3 unlike sample 2, which indicates that men with an above average probability of being employed, given their observed characteristics, have a higher than average pro-

10. Thus, lambda is computed as follows: First, I estimate a standard probit model for being employed and I compute the Inverse Mills ratio: using standard probit for $S_{it} = 1$ such that $P(S_{it} = 1 \mid W_i) = \Phi(W_i \delta)$, compute $\hat{\lambda} = \phi(\hat{\delta}' W_i) / \Phi(\hat{\delta}' W_i)$.

11. The approach of pair-wise comparisons has been applied in the matching literature by Gerfin and Lechner (2002) and Lechner (2002).

bability of marrying »migrant« women and have higher than average wages, again given their observed characteristics. In addition, a positive and significant »rho« term for samples 1 and 3 indicates that the error terms for the wage and »migrant« marriage equations are positively correlated.

Thus, in summary, men who marry »migrant« women as opposed to singles and men who are married to »immigrants« are a positively selected group, on average given their observed characteristics. However, marrying a »migrant« spouse decreases wages by 45% when compared to singles and by 53% when compared to men married to »immigrants«. The positive selection can simply be due to the difficulty and the financial burden of importing a spouse, where the individual has to prove that he is not on welfare benefits, has a certain level of income and is able to support his »migrating« wife to be able to gain entry for his wife. In addition, it is reasonable to think that there are extra costs involved in finding a spouse from another country such as travelling costs. Thus, individuals with less financial constraints, higher ability or ambition are more able to search for a spouse abroad and are more able to bring them to Denmark. However, there is a negative effect of marrying a »migrant« spouse after accounting for observable and unobservable characteristics, which might be explained by the family investment model. Thus, within this framework, the individual takes on the »borrower« role due to the credit constraints imposed upon him by his marriage type, as discussed in the previous section, which is associated with relatively lower investments in education and a flat earnings profile after labour market entrance. However, the current analysis is not sufficient to conclude either in support or rejection of the proposed mechanism behind the results.

Turning attention to the IVs as determinants of marriage types, the joint F-statistic in Table 1A clearly rejects the null of omission of the instruments from the models for both types of marriages. However, the instruments do not have the expected signs. For instance, as mentioned earlier in the paper, sex ratio is expected to measure the competition in a certain ethnic-age specific marriage market for males (or females). Hence, the larger this ratio, the less severe the competition is within the group and the less likely the individual will search for partners outside the local marriage market, that is, a negative coefficient of sex ratio is expected for the likelihood of marrying a »migrant« spouse. In theory both instruments should work well in explaining the choice between the different spouse type options and these instruments have been successfully utilised in the international literature as determinants of different marriage types. However, in the case of Denmark, where there is clearly a strong preference for »migrant« spouses over »immigrant« spouses, they clearly do not work well in the current case at hand. Thus, these results should be read by caution and are only reported in the Appendix.

Table 4. Estimation of log hourly wages with sample selection correction. 2001. (Std. err. in parantheses)

	Whole sample		Sample 1	Sample 2	Sample 3
	Males	Females	spouse type: »migrant« vs. single	spouse type: »immigrant« vs. single	spouse type: »migrant« vs. »immigrant«
experience	0.093*** (0.015)	0.059*** (0.017)	0.096*** (0.016)	0.114*** (0.022)	0.072*** (0.017)
experience squared	-0.031*** (0.006)	-0.020*** (0.007)	-0.033*** (0.007)	-0.038*** (0.010)	-0.023*** (0.006)
education	0.030*** (0.006)	0.011 (0.007)	0.031*** (0.006)	0.021** (0.009)	0.035*** (0.008)
Spouse type:					
migrant	-0.005 (0.045)	0.068 (0.053)	-0.005 (0.046)		-0.080 (0.072)
immigrant	0.070 (0.071)	0.035 (0.073)		0.016 (0.078)	
constant	4.266*** (0.115)	4.456*** (0.114)	4.244*** (0.122)	4.414*** (0.088)	4.368*** (0.180)
(λ) lambda	0.099 (0.086)	0.011 (0.078)	0.114 (0.088)	0.088 (0.086)	0.099 (0.143)
Wald F-test for migrant = immigrant prob.	1.13 0.29	0.25 0.61			
# observations	1179	914	1123	760	475

Note: (1) *** indicates significance at the one percent level, and ** and * at the 5% and 10% levels respectively. (2) Sample 1 includes only singles and individuals who married migrants; sample 2 includes only singles and individuals who married immigrants; sample 3 excludes singles and compares individuals who have migrant spouses versus immigrant spouses. (3) An indicator variable is added for missing education information.

As a next step, the instruments are dropped and spouse types are treated as exogenous variables and a two-step Heckman model is fitted to account for selection into employment. Table 4 reports the results. The coefficient estimates, compared to their OLS estimates, are mainly unaffected by the selectivity correction, and the estimated coefficients of the lambda terms do not indicate the presence of sample selectivity. Thus, the conclusion of this section is that, although there is clearly an endogeneity problem for male spouse choice decisions and their wages, the commonly used instruments in the literature were not useful to tackle this problem for the current analysis. However, in the case of females all results indicate that neither the two spouse types separately nor endogamous marriage as a whole has an association with labour market outcomes of

females. As discussed earlier, this is actually a common finding both for the general literature looking at white Western females and also in the case of the more specific literature focusing on immigrant females. However, few studies show that these findings might be misleading due to the endogeneity of education, experience and labour market attachment of females both in their labour supply and marital decisions and earnings. However, these issues are left for future research.

5. Conclusion

This paper addresses the debate in the literature regarding marriage type and employment/wage premiums for immigrants. However, rather than dealing with inter-marriages, a very narrow definition of spouse types has been utilised in this paper. In particular, the relationship between endogamous marriages and labour market outcomes for the children of guest worker immigrants is studied. The two types of endogamous marriages are analysed separately in two categories, namely marriages to »migrant« spouses and »immigrant« spouses, to test whether bringing one's spouse from the country of origin rather than marrying someone from the same ethnic background who already lives in Denmark has an effect on the individual's labour market outcomes. The results for females confirm the findings of the international literature, where there is no evidence of an effect of endogamous marriage on female employment probabilities or wages. Furthermore, there is no statistical evidence that the two marriage types should be analysed separately for females. However, in the case of males, marrying an »immigrant« spouse increases the probability of being employed and the formal tests justify the decomposition of endogamous marriages further into two categories. This result is interpreted in relation to the importance of network effects on the propensity of finding employment for immigrants, where marriage to a spouse from within the country is associated with an employment premium.

In the case of male wages, the endogeneity problem regarding spouse choice does not mimic the international phenomenon of the decision of immigrants of whether to marry within the relevant local marriage market or to search for outside options. Thus, it has not been possible to utilise competition within the group and the relevant relative group sizes to fully explain the spouse type choices as expected. The results, where spouse types are treated as exogenous variables, show that marriage type does not have an effect on male wages. However, the group of immigrants who are the focus of this study are still quite young, thus more recent data will be utilised in future analyses, where labour market outcomes and marital choices are revealed for a larger sample.

Literature

- Baker, M. and Benjamin, D. 1997. The Role of the Family in Immigrants« Labor-Market Activity: An Evaluation of Alternative Explanation, *American Economic Review*, 87 (4): 705-27.
- Becker, G. S. 1973. A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4): 813-46.
- Blau, F. D., Kahn, L. M., Moriarty, J. Y., and Souza, A. P. 2003. The Role of the Family in Immigrants Labour Market Activity: An Evaluation of Alternative Explanations: Comment, *American Economic Review*, 93(1): 429-47.
- Borjas, G. J. 1985. Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics*, 3(4): 63-489.
- Borjas, G. J. 1995. Assimilation and changes in cohort quality revisited: What happened to immigrants earnings in the 1980s?, *Journal of Labor Economics*, 13(2): 201-45.
- Borjas, G. J. 1999. Economic analysis of immigration, in Orley Ashenfelter and David Card (eds.) *Handbook of labor economics* 3A:1697-1760, Elsevier Science, Amsterdam.
- Card, D., DiNardo, J, and Estes, E. 2000. The More Things Change: Immigrants and the Children of Immigrants in the 1940s, 1970s and 1990s, in G. J. Borjas (ed.) *Issues in the Economic of Immigration*, University of Chicago Press, Chicago.
- Chiswick, B. R. 1978. The Effect of Americanization on the Earnings of Foreign-born Men., *Journal of Political Economy*, 86(5): 897-921.
- Chiswick, B. R. 1979. The Economic Progress of Immigrants: Some Apparently Universal Patterns., in Fellner, W. (ed.) *Contemporary Economic Problems*, American Enterprise Institute, Washington, 357-99.
- Cox, O. C. 1940. Sex Ratio and Marital Status among Negroes., *American Sociological Review*, 6:37-72.
- Çelikaksoy, A. 2003. Marriage Migration: The case of Denmark, Unpublished Paper.
- Duleep, H. O. and Sanders, S. 1993. The decision to work by married immigrant women, *Industrial and Labor Relations Review*, 46 (4): 677-90.
- Duncan, B. and Trejo, S. J. 2005. Ethnic identification, intermarriage, and unmeasured progress by Mexican Americans, *IZA discussion paper series* No. 1629.
- Gerfin, M. and M. Lechner. 2002. A Microeconomic Evaluation of the Active Labour Market Policy in Switzerland, *Economic Journal* 112: 854-93.
- Gordon, M. M. 1964. *Assimilation in American Life*, Oxford University Press, New York.
- Grossbard-Shechtman, S. 1993. *On the Economics of Marriage: A Theory of Marriage Labor and Divorce*, Boulder, Westview Press.
- Groves, E. R. and Ogburn, W. F. 1928. *American Marriage and Family Relationships*, Henry Holt and Company, New York.
- Heckman, J. J. 1979. Sample Selection Bias as a Specification Error, *Econometrica*, 47 (1): 153-161
- Hersch, J. and Stratton, L. S. 2000. Household specialization and the male marriage wage premium, *Industrial and Labor Relations Review*, 54(1): 78-94.
- Hutchinson, B. 1957. Some evidence related to matrimonial selection and immigrant assimilation in Brazil, *Population Studies*, 11(2): 149-56.
- Kantarevic, J. 2004. Interethnic marriages and economic assimilation of immigrants, *IZA discussion paper series*, No. 1142.
- Kermit, D. 1991. Does marriage make men more productive?, University of Chicago, *Economics Research Center*, No. 92-2.
- Kitano, H. H. L. and Yeung, W. T. 1982. Chinese interracial intermarriage, *Marriage and Family Review*, 5(1): 35-47.
- Klein, T. 2001. Intermarriage between Germans and Foreigners in Germany, *Journal of Comparative Family Studies*, Calgary, 32(3): 325-47.
- Korenman, S. and Neumark, D. 1991. Does Marriage Really Make Men More Productive?, *The Journal of Human Resources*,

- 26(2): 282-307.
- Lechner, M. 2002. Some Practical Issues in the Evaluation of Heterogenous Labour Market Programmes by Matching Methods, *Journal of the Royal Statistical Society A* 165(1): 59-82.
- Loh, E. S. 1996. Productivity differences and the marriage wage premium for white males, *Journal of Human Resources*, 31(3): 566-89.
- Meng, X. and Gregory, R.G. 2005. Intermarriage and the Economic Assimilation of Immigrants, *Journal of Labor Economics* 23(1): 135-75.
- Meng, X. and Meurs, D. 2006. Intermarriage, Language and the Economic Assimilation Process: A Case Study of France, *IZA discussion paper series*, No. 2461.
- Qian, Z. 1999. Who Intermarries? Education, Nativity, Region, and Interracial Marriage, 1980 and 1990. *Journal of Comparative Family Studies* 30(4): 579-97.
- Qian, Z., Blair, S. L. and Ruf, S. D. 2001. Asian American interracial and interethnic marriages: Differences by education and nativity, *The International Migration Review*, 35(2): 557-86.
- Ribar, D. C. 2004. What do social scientists know about the benefits of marriage? A review of quantitative methodologies, *IZA Discussion Paper Series* No. 998, Bonn.
- Rodgers, W. M. and Stratton, L. S. 2005. The male marital wage differential: Race, training, and fixed effects, *IZA Discussion Paper Series*, No. 174.

Appendix

Table 1A. Instrumental variable estimates for log hourly wages. 2001. (Std. err. in parentheses).

	Sample 1 spouse type: »migrant« vs. single		Sample 2 spouse type: »immigrant« vs. single		Sample 3 spouse type: »migrant« vs. »immigrant«	
	Wage	Migrant	Wage	Immigrant	Wage	Migrant
sex ratio		6.301*** (1.547)		0.070 (2.659)		2.645 (3.089)
relative group size		-0.303 (0.223)		-0.870** (0.364)		0.642 (0.534)
experience	0.199*** (0.042)	0.887*** (0.074)	0.124*** (0.029)	0.083 (0.127)	0.106*** (0.024)	0.564*** (0.113)
experience squared	-0.068*** (0.015)	-0.298*** (0.031)	-0.039*** (0.010)	-0.010 (0.050)	-0.036*** (0.009)	-0.209*** (0.042)
education	0.042*** (0.008)	0.128*** (0.028)	0.025*** (0.010)	-0.037 (0.044)	0.040*** (0.008)	0.122*** (0.041)
(IMR) for selection into employment	0.533*** (0.176)	3.682*** (0.428)	0.149 (0.139)	-1.412* (0.844)	0.283** (0.141)	3.931*** (0.798)
<i>Spouse type:</i>						
migrant		-0.451*** (0.182)				-0.525** (0.219)
immigrant			0.099 (0.313)			
constant	3.765*** (0.212)	-12.324*** (1.763)	4.232*** (0.183)	-0.171 (3.049)	4.484*** (0.179)	-7.224** (3.275)
# observations		599		389		282
rho		0.637		0.180		0.680
		»migrant«		»immigrant«		
joint <i>F</i> -tests of the <i>IV</i> 's		spouse		spouse		
<i>prob.</i>		30.70		12.04		
		0.000		0.000		

Note: (1) *** indicates significance at the one percent level, and ** and * at the 5% and 10% levels, respectively (2) An indicator variable is added for missing education information.

Debat og kommentarer

Estimating the Effects of Delayed Entry into Higher Education: A Discussion

Maria Knoth Humlum

School of Economics and Management, University of Aarhus, E-mail: mhumlum@econ.au.dk

SUMMARY: In Denmark many high school graduates choose to delay their entry into higher education. A number of studies have investigated the effects of the delay on educational and labour market careers. The existing studies are likely to suffer from selection bias, and this problem is unlikely to be remedied by use of control variables. More advanced techniques that can eliminate the selection bias are needed in order to identify causal effects. For policy purposes it would thus be preferable to focus on the economic losses associated with the years lost in the skilled labour market which are potentially substantial.

1. Introduction

A new entrant into higher education in Denmark has a median age of almost 23. International comparisons place Denmark at the absolute low end with respect to making quick transitions from education to work. E.g. while about 40% of the Danish population aged 25-29 are enrolled in an education, this is only the case for about 23% of the Swedish and 15% of the Norwegian populations.¹ At a time when the population is ageing, the dependency ratio is rising, and the demand for skilled labour is increasing, delayed entry into higher education poses a considerable problem for the Danish welfare state. There are several ways to expand the labour force, but one obvious way would be to reduce the relatively unproductive years spent between high school graduation and entry into higher education. These years have been dubbed »sabbatical years« in Denmark. However, the advocates of sabbatical years claim that the sabbatical years are productivity-increasing in that they help to »mature« students before their entry into higher education, by letting them work, travel, etc. The somewhat loose argument is that individuals acquire valuable experience during these years spent outside the formal educational system. Of course, the effects of delayed entry are likely to differ

I am indebted to Helena Skyt Nielsen and Rune Majlund Vejlin for many helpful comments and suggestions. I appreciate the useful comments given by two anonymous referees. The usual disclaimer applies.

1. OECD (2005), tables C2.2 and C4.2a.

depending on the nature of the delay, e.g. a year of travelling is likely to have different effects on future outcomes than a year of working. Studies that investigate the effects of the delayed entry on the educational and labour market careers do not find conclusive evidence, and these studies generally do not account for the potential endogeneity of the delay in the outcome equations. I argue that accounting for the potential endogeneity of the delay is crucial for giving the estimated effect a causal interpretation and, therefore, also for the policy conclusions that can be drawn from these studies. It will be relevant for future research in this area to extend the theoretical arguments for the choice of delayed entry. This will be essential for the choice of suitable instrumental variables and for producing relevant policy analyses.

Apart from increasing the size of the labour force, the government can choose to intervene in young people's decision process for a number of reasons. First, a free-rider problem arises because individuals who delay their entry into higher education end up contributing less to the welfare state. Second, if individuals postpone their entry due to »market failures« such as informational constraints, the government could intervene to alleviate such constraints if possible. Third, delayed entry could be associated with peer effects, i.e. an individual's choice of delay is affected by the choices of his peers. In this case, one can imagine the existence of multiple equilibria such that intervention on the part of the government might be Pareto-improving.

The problem of delayed entry into higher education has been present for many years. In fact, the average years of delay per high school cohort was increasing during most of the period 1978-1986.² In recent years the problem has gained a lot of media and political attention, and the number of individuals who enrol immediately after high school graduation has been increasing. The delayed entry into higher education is also a cause of growing concern in other countries, e.g. Sweden, Holmlund et al., (2006) and the US, Monks (1997).

In both DORS (2003) and Velfærdskommissionen (2005) it was recommended that the duration of the delay be reduced. The main motivation is to expand the labour force and waste fewer productive years on non-productive activities. On the other side of the debate are the student bodies and their sympathizers that argue that the delay is spent performing productive activities, such as information gathering, increasing motivation, and simply accumulating valuable skills that are different from those obtained in the educational system. It is important to keep in mind that even if delayed entry increases productivity at later stages these productivity effects would still have to outweigh the loss incurred due to the fact that the individual spends less years on the skilled labour market, cf. the following simple example.

Consider two individuals who differ only in terms of time of enrolment in higher

2. See figure 1 in section 3.

education. Suppose that the annual income before enrolment is 75,000 DKK reflecting that many individuals do not work full-time during this period, the annual income during education is 75,000 DKK, and the annual income after graduation is 275,000. Furthermore, the age at the time of high school graduation is 19, an individual spends 5 years in higher education, and retires from the labour market at age 65. To make things even more simple consider the case where the wage profile is flat, and assume a discount factor of 0.05. Using these assumptions one can calculate the percentage increase in annual income after graduation that would compensate for delaying entry for a given number of years. E.g. in this very simple setting an individual who delays his education for two years will have to earn 9% more annually for the rest of his life than an individual who does not delay, in order to get the same present value of lifetime income at age 19. Of course, such calculations are very sensitive to the assumptions made about incomes, discount factors, wage profiles etc., but examples of this sort are very illustrative of the importance of the extra years spent in the skilled labour market in terms of lifetime income. In order to perform more realistic calculations, one needs detailed knowledge of the effects of delayed entry on the wage, the retirement age, the time spent in higher education etc. In the above example all dynamic effects of delayed entry are assumed to be non-existent. The main point to deduce from this example is that the drop in lifetime income caused by a shorter labour market career is potentially substantial.

The plan of the paper is as follows. In section 2, I briefly go over some theoretical considerations on this issue. In section 3, I present some descriptive statistics on the delayed entry into higher education in Denmark. In section 4, I summarize the findings of (primarily Danish) studies of the effects of these delays on the educational and labour market careers and discuss the problems with these studies and the interpretation of their results. Section 5 discusses potential instruments that could be used to identify a causal effect. Section 6 concludes and discusses policy implications.

2. Theoretical considerations

From the perspective of conventional human capital theory, the delayed entry into higher education is somewhat of a conundrum. According to the Ben-Porath (1967) model it is optimal for individuals to accumulate human capital in the early stages of the lifecycle and decumulate human capital in later stages of the lifecycle. If one imagines that the delay is spent e.g. working or consuming pure leisure, the prevalence of delays is not in strict accordance with the human capital theory. It would appear that individuals who delay their entry into higher education are decumulating human capital at a very early stage of their lifecycle.

Taking a broader view of human capital, investments in human capital can encompass as diverse elements as education, health, information search and on-the-job trai-

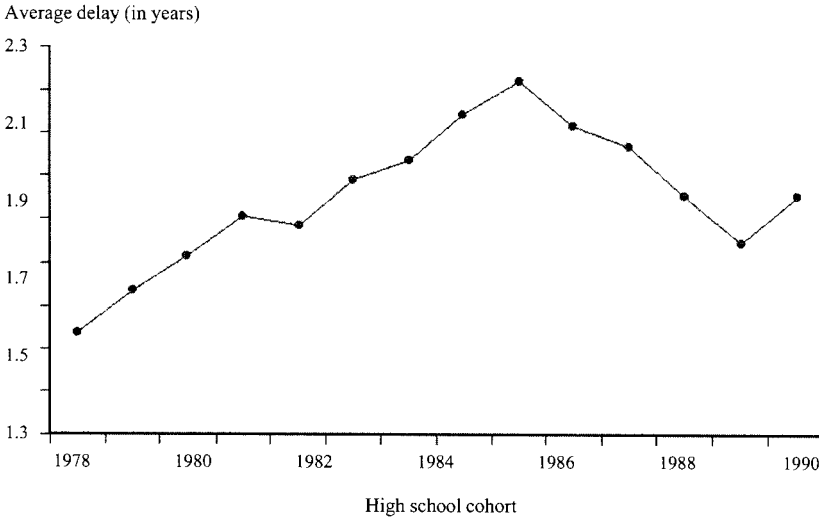


Figure 1. The average delay for each high school cohort, 1978-1991.

ning. Thus, the delay can be interpreted as just another period of human capital accumulation. After all, individuals are usually working, travelling, doing military service or acquiring other (informal) types of schooling. Also, if acquiring an education is a risky investment there is an incentive to gather more information about the risk and return of the investment prior to making the investment, Becker (1993).³ Consequently, it could be optimal to delay entry into higher education in order to make a more informed educational decision. Another potential explanation is that for some individuals, a longer delay might extend their educational choice set.⁴ Finally, the leisurely years spent between high school graduation and entry into higher education could be associated with a sizeable consumption value.

3. Descriptive statistics

The descriptive statistics presented in this section are the author's own calculations based on the AKF 10% panel data sample of the Danish population. Each individual is observed once a year in the period 1984-2001. In addition to detailed information on earnings, parental background etc. there is also information on the educational histories of the individuals in the sample. All statistics in this section are based on individuals who graduated from an ordinary high school in the years 1978-1991, and whose age at high school graduation was less than 22.⁵ The delay is defined as the number of years

3. In the Ben-Porath (1967) model there is no uncertainty.

4. This is a direct consequence of the structure of the admission system in Denmark.

5. The choice of cohorts reflects that more recent cohorts have not yet settled into the labour market.

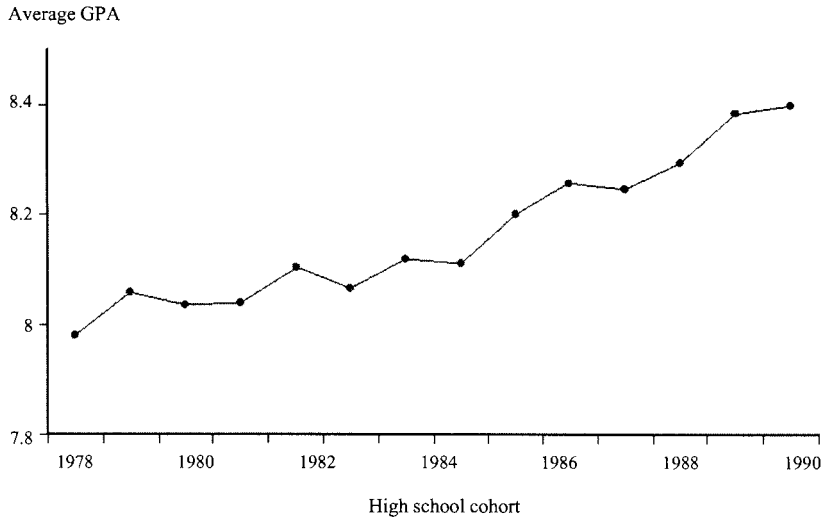


Figure 2. Average GPA for each of the high school cohorts, 1978-1991.

passed between (the first) high school graduation and (the first) entry into higher education.⁶ Whether or not the education is eventually completed does not influence the delay. A more detailed description of the data and definitions of variables are given in Humlum (2006).

Figure 1 shows the historical development in sabbatical years by high school cohort. In only 8 years the average number of sabbatical years increased by about 0.6 years. The size of the delay was generally increasing for the cohorts 1978-1986 and decreasing for later cohorts. This could be a reflection of the big backlog of students built up in the aftermath of the reform of the admission requirements in 1977. The number of study places at the University of Copenhagen was reduced every year during the years 1977-1981. Moreover, the size of the high school cohorts increased dramatically during the period 1978-1985 and reached its peak in 1985 when a high school cohort counted about 20,000 individuals compared to only about 12,000 individuals in 1978.⁷ The increasing cohort sizes are likely to be associated with an increasing demand for higher education. Another interesting development is that the average GPA for a high school cohort actually increased during this time. Figure 2 shows that

6. The sample selection criteria were chosen in order to obtain a relatively homogenous sample of high school graduates that are likely to have followed the same overall educational path. If individuals differ in their educational paths prior to high school graduation then these differences would, in principle, need to be incorporated in an analysis.

7. Ministry of education databases: number of pupils, completed language-based or mathematically based high school.

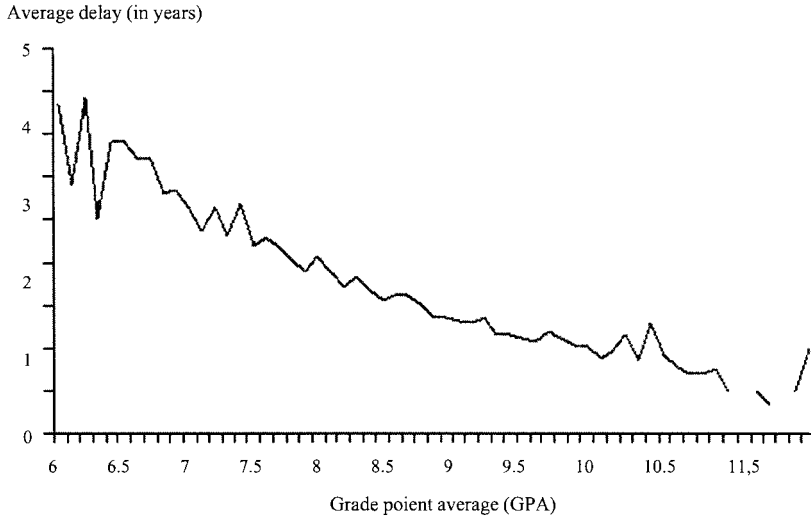


Figure 3. Average delay by high school grade point average.

from 1978 to 1991 high school GPAs steadily increased from around 8.0 to 8.4. In the presence of such grade inflation, the distribution of GPAs is clearly not constant across high school cohorts. Consequently, including GPA as an explanatory variable when estimating the effects of delay on, e.g., earnings, can lead to biased estimates of the effects of the delay.⁸

Figure 3 shows the average years of delay for each high school GPA. On average, individuals with a higher GPA have shorter delays. In the data the correlation between years of delay and GPA is statistically significant and equals -0.25 . There are two straightforward explanations of this phenomenon. First, it is plausible that high-ability individuals have higher returns to education and/or stronger preferences for education. Secondly, the admission requirements for higher education effectively restrict the access to higher education of some individuals with low GPAs, and the setup of the admission system induces them to delay their entry.

The duration of the delay also varies considerably depending on the education in which the individual chooses to enrol. Figure 4 shows the average years of delay for different educational groups. Again, differences in admission requirements could be part of the explanation for this pattern. However, since the long-cycle educations are generally associated with a smaller delay and relatively strict admission requirements, this does not appear to be the main explanation. Another explanation could be that the benefits associated with delaying entry into education are higher in some occupations

8. I will discuss this problem in more detail in section 4.

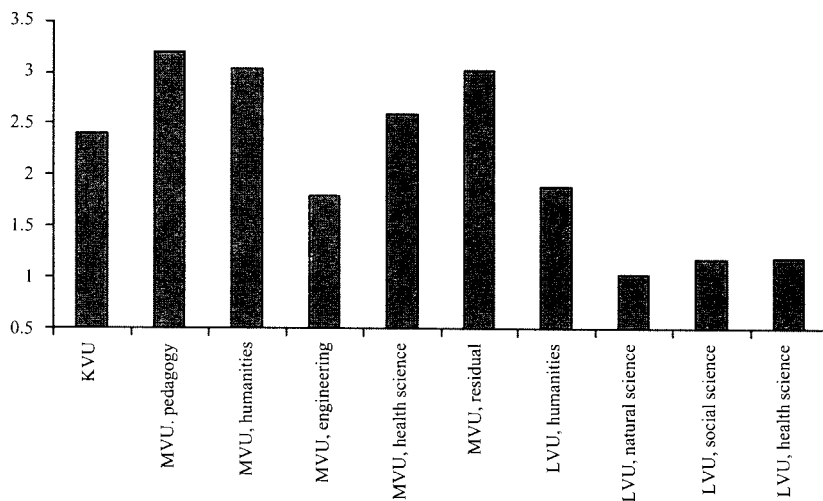


Figure 4. Average delay (in years) by length and field of education. *KVVU*: short-cycle higher education, *MVU*: medium-cycle higher education, and *LVU*: long-cycle higher education.

than in others, e.g. social educators, teachers and nurses might experience age premiums. In general, it is troublesome to separate the effects of age from those of delayed entry.

Table 1 shows average earnings measured in 2001 DKK and the probability of dropping out of an education by years of delay. Earnings are measured 1 and 5 years after graduation from higher education. For both earnings measures the general tendency is that as the delay increases the earnings decrease. This pattern is in accordance with a negative causal impact due to e.g. decumulation of human capital or a negative spurious relation due to self-selection if individuals with low ability and low motivation self-select into longer delays.

For longer delays the probability of dropping out is lower. Also, the average delay for individuals who do not drop out is 2.0 years, while the average delay for those who drop out is only 1.53 years. This supports the view of those who consider the delay to be productive, either in the sense of being able to choose the »right« education, or in the sense of having acquired skills that increase educational performance. If high-ability individuals self-select into short delays, as suggested above, then this is a somewhat puzzling result. One would expect high-ability individuals to have smaller probabilities of dropping out. A thing that could be playing a role here is that the gains from dropping out are likely to be lower for individuals who have already delayed. They

Table 1. Earnings and the probability of dropping out by number of years of delay.

Delay (in years)	Earnings 1 year after graduation ^(a)	Earnings 5 years after graduation ^(a)	Drop out probability
0	238.561	327.646	0,35
1	226.016	294.677	0,30
2	217.017	269.873	0,25
3	213.718	255.900	0,24
4	205.486	251.607	0,20
5	201.831	243.613	0,20

Note: (a) Observations with zero earnings are excluded.

have fewer years left before retirement and making a new educational investment will be associated with a relatively low return.⁹

The descriptive statistics in this section suggest that the delay is correlated with observables such as earnings, probability of dropping out, choice of education, high school GPA, etc. Other schooling decisions are widely believed to be heavily correlated with unobservable characteristics as well as observable characteristics. For instance, non-cognitive skills which are generally unobserved are likely to influence educational decisions. In summary, there is no reason to think that the delay is uncorrelated with the unobservables, and simple comparisons of observationally equivalent individuals with different durations of delays will not be informative about causal interpretations.

4. Discussion of existing studies and selection bias

Economic theory does not provide us with an exact answer to whether or not the delayed entry is productivity-enhancing, and establishing this empirically poses some fundamental problems. As is the case with education, to delay or not is a decision made by the individual, implying that individuals with specific characteristics (both observable and unobservable) will self-select into shorter and longer delays. As a result, individuals who delay their entry into higher education are likely to be different from those who do not, and they are also likely to experience higher gains from delaying their entry. In this case simple comparisons of outcomes for these groups will not be informative about the direction and size of the causal effect of the delay on the outcome of interest. To estimate a *causal* effect of the delay on an outcome variable, one must employ more advanced techniques such as IV estimation.

9. The figures in table 1 are aggregated across all educational groups implying that they are also a reflection of the heterogeneity of earnings and drop out rates across educational groups. A similar table for each educational group displays a less clear pattern, but the number of observations in each cell is heavily reduced in this case.

Examples of unobservable characteristics that could be imagined to affect both an individual's tendency to delay his enrolment and his educational and labour market outcomes could be intrinsic ability and non-cognitive skills such as motivation and determination. Taking a conventional view this would probably lead to a downward bias in estimates of the effect of delayed entry, since these unobservables are likely to be negatively correlated with the delay, but positively correlated with earnings. On the other hand, one might also imagine personality traits that are positively correlated with delayed entry, and yet, which are also positively correlated with wages, e.g., individuals who are adventurous, independent, risk-taking etc. might be more likely to delay their entry into higher education. However, these are also traits that might be rewarded in the labour market implying that one could also argue that the estimated effect of the delay on e.g. earnings is biased upwards. In conclusion, the direction of bias cannot be straightforwardly deduced, but a traditional perspective would suggest that estimates of the effects of delayed entry will be downward biased if the selection bias is not accounted for.

As mentioned in the introduction, a potential productivity effect of delay on labour market outcomes would not only need to be positive, it would also have to be of a substantial size in order to justify delaying enrolment. Existing studies do not suggest that this is the case. These studies have focused on the effects of delayed entry into education on two outcomes: earnings and the probability of dropping out of an education. Vejrup-Hansen and Agersnap (2002) find a small, but positive, effect of smaller delays on the starting wage, while Velfærdskommissionen (2005) calculates the gain of entering the labour market one year earlier as being about 2% of lifetime disposable income. The estimation technique used in the latter study is propensity score matching, but this does not eliminate the problems of selection bias, since it is only possible to match individuals on observable characteristics. Klausen (2006) focuses on the question: Do younger graduates have higher earnings? He concludes that younger graduates do have higher earnings and that the results hold both for earnings in the beginning of and later in the career. However, in the presence of selection bias it is impossible to say whether younger graduates have higher earnings because they are younger, or because *younger* graduates have higher intrinsic ability, or are more motivated etc.

Another source of bias in these kinds of estimates occurs when you compare younger and older graduates. When comparing graduates who graduate at the same time, but are 7 years apart in age, you will often also be comparing individuals who are 7 years apart in terms of year of high school graduation. In the 1980s there was moderate grade inflation in Denmark in terms of high school GPAs, see figure 2, and this phenomenon will lead to a bias in this type of estimate, unless one believes that the ability of the high school cohorts has actually increased. The point is that the GPA will be ne-

gatively correlated with the graduation age and positively correlated with the time of high school graduation. A direct implication of this is that the estimated effect of graduation age on earnings will be downward biased.

One common way to investigate how the delayed entry into education influences the educational career is to look at the effect on the probability of dropping out of an education. DORS (2003) and Velfærdskommissionen (2005) find similar results, namely that the delayed entry into education tends to increase the probability of dropping out. However, this is contrasted by the findings of DST (2006) and Humlum (2006) who find that the delayed entry tends to decrease the probability of dropping out. One reason for this disparity in results (using relatively similar methods and control variables) could be differences in the definitions of the sample populations and the delay. Furthermore, the first two studies mentioned use indicator variables for years of delay, whereas the two latter studies use a count variable. Using indicator variables allows for possible non-linearities in the effect of the delay on the probability of dropping out.

A few international studies have also investigated the effects of delayed entry on earnings. A study using Swedish data, Holmlund et al. (2006), finds that delaying entry into higher education by two years is associated with lower lifetime earnings. In present value terms the lower lifetime earnings correspond to about 40-50% of annual earnings at age 40. In a study on US data, Monks (1997) investigates the effect of delayed college enrolment on earnings using a fixed effect framework to try to account for self-selection, and he finds that older graduates have lower earnings than younger graduates.

The Danish studies do not account for the potential self-selection into shorter and longer delays. The central issue is that unless one firmly believes that an individual who has delayed his entry into higher education by e.g. 3 years is in every respect (that is not explicitly controlled for) equal to an individual who has not delayed entry, then the existing Danish studies suffer from selection bias, and policy recommendations based on these studies are thus invalid.

5. IV estimation and potential instruments

One way to solve the problem of selection bias is to apply instrumental variables (IV) strategies. Different IV approaches have been used extensively in the literature on returns to schooling, and various, more or less, ingenious instruments have been used for instrumenting years of schooling: distance to college, Card (1995), school reforms, Black et al. (2005) etc. The main problem with IV estimation is that finding a valid instrument can be a real challenge. And, without a valid instrument IV estimates can be even more biased than OLS estimates.

The IV procedure rests on the assumption that some component of the variation in the endogenous variable is exogenous. The appropriate instrument should be able to explain some or all of this exogenous variation in the endogenous variable. A feasible instrument must be both *valid* and *strong*. For an instrument to be valid it must be uncorrelated with the unobservables in the second stage regression. For an instrument to be strong it must be »sufficiently« partially correlated with the endogenous variable. Using an instrument that satisfies these two criteria implies that the causal effect of the delay on the outcome of interest can be identified.

5.1 Potential instruments for the number of years of delay

In the following I will discuss some potential instruments for the years of delay to illustrate what kind of changes might induce exogenous variation in the duration of delay. I consider 4 potential instruments: changes in the admission requirements, changes in the student grant system, compulsory military service and distance to education.

5.1.1 Instruments based on the admission requirements

Admission to higher education in Denmark is restricted. The admission requirements are primarily based on high school GPA and they change each year with the number of applicants and the number of available places. This implies that two completely identical individuals, differing only in their year of high school graduation, will face different opportunity sets when they graduate. They may not be able to enter the same set of educational courses since the admission requirements they face in the year of high school graduation will not be the same. It is not uncommon that individuals, who have decided on the education they want and who actually want to enrol, have to wait several years before they have collected enough points to gain admission, or alternatively, before the admission requirements loosen. Intuitively, if admission requirements were tight in the year of high school graduation, the high school graduates will, on average, have longer delays. This idea is supported by the negative relationship between GPA and the delay, see figure 3. Figure 1 suggested that there is considerable variation in the average number of sabbatical years over high school cohorts. This could also be interpreted as a result of changing opportunity sets for the high school cohorts. In the 1990s as many as 20-40% of the applicants to higher education were rejected each year, Ministry of Education (1998), suggesting that the effect of admission requirements on the extent of the delay is of a considerable size. Furthermore, if one assumes away general equilibrium effects that may occur in the long run, admission requirements in the year of high school graduation are likely to be exogenous in the outcome equation: an individual cannot influence the admission requirements, and the year of high school graduation is more or less fixed.

5.1.2 Instruments based on the student grant system

Students at the institutions of higher education in Denmark qualify for student grants. A student receives a student grant each month which should cover living expenses and other expenses incurred during the education. Introduced in 1970, the student grant system has been subject to various reforms over time, including minor changes from year to year. An increase in the student grant would be an additional incentive for an individual to enrol in higher education.

As changes in the student grant system are likely to affect the incentives of individuals choosing the number of years to delay their entry into higher education, an instrument that takes advantage of such variations in the student grant system might be able to explain some of the variation in the durations of the delay. Changes that would be of interest in this context are changes that induce some individuals to enrol who would otherwise not have enrolled at that point in time, or vice versa. For example, changes in the student grant system that correspond to an increase (decrease) in the student grants from year to year, would reduce the costs associated with taking an education. One historical example where this was indeed the case was a reform of the student grant system effectuated in the academic year 1988/89. From the year 1987/88 and to the following year the monthly student grant was significantly increased. On the basis of this change in the student grant system, one might suspect that the high school cohorts of 1987 and 1988 systematically differed in their choice of sabbatical years.¹⁰ In figure 1 it is clear that the average delay is in fact slightly shorter for the 1988 cohort which was provided with increased incentives to enter higher education immediately. However, the difference in average delays is small, especially when compared with the differences between other cohorts.

Two adjacent high school cohorts are not likely to exhibit large differences in observed characteristics. But, if other important variables such as the admission requirements differ across cohorts, an indicator for high school cohort may well pick up other effects than the one originally intended. With regard to the strength of this instrument it is not obvious that increasing the size of the student grants would induce a lot of individuals to enrol sooner.

In the current debate it has been suggested that the student grant should be smaller for individuals who postpone enrolment for some specific number of years. If this reform was introduced, it would probably be useful in providing some exogenous variation in the timing of enrolment.

10. Of course, one would have to assume that the future increase in the student grant was known at the time the 1988 cohort applied for enrolment. Furthermore, if the change had been anticipated for a number of years, some individuals might have delayed their entry until the new system was in place. In this case the reform would have the additional implications that some individuals delayed entry more than they would have done otherwise.

5.1.3 Instruments based on the compulsory military service

In Denmark young men are required to come up before a draft board, typically in the year in which they turn 18. The purpose is to determine their eligibility for military service. At the draft board the young men are required to take a written test, a health examination, and subsequently, the draft board establishes the eligibility of the draftee. Those who are declared eligible for military service draw lots, and the ones who draw low lottery numbers are required to do their compulsory military service. Typically, the young men who are drafted do their military service in the year following graduation from high school. An important implication of this is that individuals who draw lower lottery numbers are more likely to delay their entry into higher education. A binary variable indicating whether or not the individual drew a low lottery number would therefore be a potential instrument for the average years of delay. The randomness of the instrument ensures that it will be exogenous in the earnings equation. A similar instrument using the draft lottery in the US was employed in Angrist (1990) to estimate the effect of veteran status on earnings.

One issue that arises concerning this instrument is whether or not doing your compulsory military service has an effect on your future earnings. However, since the duration of the compulsory military service is rather short, and draftees are not sent to war, an assumption of no effect of lottery number on earnings except through the effect on the delay seems reasonable. The major limitation is, of course, the instrument's limited scope in that it is only applicable for males.

5.1.4 Instruments based on distance to education

Distance to education has been used to instrument years of schooling in earnings regressions, Card (1995). The idea is that the costs associated with the acquisition of more schooling are lower, when the distance to education is lower. This might be due to transportation costs, boarding costs, psychological costs associated with moving etc. Consider the decision problem of an individual who has to time his entry into higher education. Direct pecuniary costs such as relocation costs etc. are relatively constant over time, and should therefore only have a very small effect on the choice of delay, if there are no borrowing constraints. However, it is easy to imagine non-pecuniary costs of moving that decrease with age. First, individuals who have just completed high school may not wish to leave the security and stability of their homes. Secondly, social networks may tie them to their home town. Such non-pecuniary costs are likely to decrease with age, as the individual matures and wants to leave home, and as the years pass, the social network is often dissolved as individuals begin to move away from their home town. For individuals with low distance to educational institutions these costs are not necessarily present. There might be problems with the validity of this instrument if individuals self-select into locations.

5.1.5 General considerations

Whether or not any of the 4 types of instruments mentioned here are applicable in the case of estimating the causal effect of delaying entry into higher education would have to be further investigated. One can argue that each of the four is a valid instrument, but only direct investigations in the form of first-stage regressions will give a definitive answer.^{11, 12} Even if an instrument is valid it may not be sufficiently strong to provide feasible estimates of the causal effect, cf. Bound et al. (1995) who describes some of the problems related to the use of weak instruments. Specifically, IV estimates can be even more biased than OLS estimates. Many of the instruments discussed here rely only on time series variation which can lead to problems with weak identification. Generally, it is preferable to use an instrument that relies on both time series variation and cross section variation for identification.

When interpreting IV estimates one should keep in mind that the estimated effect can only be interpreted as an average treatment effect (ATE) if it is assumed that the effect is constant across the population. In the presence of heterogeneous effects the IV estimate has a local average treatment effect (LATE) interpretation.

LATE is the average treatment effect of those individuals who are affected by the instrument. Consequently, whether or not LATE is an interesting parameter depends on the instrument and the research question. For the compulsory military service instrument LATE would be the average treatment effect for those who were induced to have longer delays because they drew a low lottery number.

6. Discussion and policy implications

The main objective of this paper was to demonstrate that existing Danish studies of the effects of the delayed entry into higher education on the educational and labour market careers cannot be given *causal* interpretations. Hence, it is hard to deduce policy conclusions from these studies. Using IV estimation to identify the causal effects of the delayed entry on educational and labour market careers is one possible direction for future research in this area. I have suggested four types of potential instruments for the number of years of delay: instruments based on the admission requirements for higher education, instruments based on the student grant system, instruments based on compulsory military service, and finally instruments based on the distance to education.

To a large extent the focus of the debate and the existing studies have been on the

11. Holmlund et al. (2006) explore the following three instruments: the timing of military service enrolment, the local labour market conditions at the time of high school graduation, and an admissions reform. They report that the first-stage estimates are too imprecise to be informative.

12. In Humlum (2006), I use an instrument based on the admission requirements for higher education as an instrument for years of delay, but the coefficients of the instruments in the first-stage regressions were not always of the expected sign, and therefore no conclusions could be made about the causal effect.

productivity effects of the delayed entry into higher education. However, the economic losses associated with the years lost in the skilled labour market are likely to be substantial – for society and for the individual. An individual who spends fewer years on the skilled labour market will have a lower lifetime income, *ceteris paribus*. To counter this effect, the productivity effects of delaying entry would have to be substantial. At present, the empirical evidence does not support this conclusion.

Literature

- Angrist, J. D. 1990. Lifetime Earnings and the Vietnam Era Draft Lottery, *American Economic Review*, Vol. 80, No. 3, pp. 313-36.
- Becker, G. S. 1993. *Human Capital. A Theoretical and Empirical Analysis, with Special Reference to Education*, The University of Chicago Press, 3rd edition, Chicago.
- Ben-Porath, Y. 1967. The Production of Human Capital and the Life Cycle of Earnings, *Journal of Political Economy*, Vol. 75, No. 4, Part 1, pp. 352-65.
- Black, S. E., P. J. Devereux og K. G. Salvanes. 2005. Why the Apple Doesn't Fall Far: Understanding the Intergenerational Transmission of Human Capital, *American Economic Review*, Vol. 95, No. 1, pp. 437-49.
- Bound, J., Jaeger, D. A. and R. M. Baker. 1995. Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak, *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 443-50.
- Card, D. 1995. Using Geographic Variation in College Proximity to Estimate the Return to Schooling, in: Louis N. Christofides, E. Kenneth Grant and Robert Swidinsky, eds., *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, University of Toronto Press, Toronto, Canada, pp. 201-22.
- Card, D. 1999. *The Causal Effect of Education on Earnings*, in *Handbook of Labour Economics*, Vol. 3, Elsevier, Amsterdam.
- DORS. 2003. *Dansk økonomi, efterår 2003*, Det Økonomiske Råd, København.
- DST. 2006. Frafald og fuldførelse på videregående uddannelser, in: *Statistisk Tiårsoversigt*, Statistics Denmark, Copenhagen.
- Holmlund, B., Q. Liu og O. N. Skans. 2006. Mind the Gap? Estimating the Effects of Postponing Higher Education, *Working Paper 2006:17*, Department of Economics, Uppsala University.
- Humlum, M. K. 2006. *The Causal Effect on Wages of the Delayed Entry into Higher Education in Denmark*, Speciale, Institut for Økonomi, Aarhus Universitet.
- Imbens, G. og J. Angrist. 1994. Identification and Estimation of Local Average Treatment Effects, *Econometrica*, Vol. 62, pp. 467-76.
- Klausen, T. B. 2006. *Humankapital. Hvad betyder dimittendalder for indkomsten?*, Report no. 1 from the project: *Catching up qualifications or fooling around?*, Rockwool Foundation and Centre for Comparative Welfare Studies, Aalborg University, *CCWS Working paper*, 2006, 37.
- Ministry of Education. 1998. Værd at vide om de videregående uddannelser – *Statistik 1997/98*, Undervisningsministeriet, København.
- Monks, J. 1997. The Impact of College Timing on Earnings, *Economics of Education Review*, Vol. 16, No. 4, pp. 419-23.
- OECD. 2005. *Education at a Glance: OECD Indicators – 2005 Edition*, OECD, Paris.
- Vejrup-Hansen, P. og F. Agersnap. 2002. Studieførløb og karriere, *Nationaløkonomisk Tidsskrift*, Vol. 140, pp. 185-220.
- Velfærdskommissionen. 2005. *Fremtidens velfærd – vores valg*, Analyserapport.

Jesper Jespersens disputats

Jesper Jespersen: Makroøkonomisk metodologi^{1, 2}

Lars Pålsson Syll

Malmö Högskola, E-mail: Lars.Palsson-Syll@mah.se

Inledande kommentarer

Avhandlingen börjar med en resumé och ett påföljande kapitel 0 där författaren positionerar sitt arbete utifrån den pågående forskningshorisonten. Vad han väsentligen gör här är att klargöra de avgörande metodologiska dragen hos främst den neoklassiska teorin och olika varianter av Keynes-inspirerad ekonomi. Det hela mynnar ut i en figur (s 40) som beskriver ledande makroekonomiska skolor och deras intellektuella ursprung. En av huvudpoängerna – som jag förstår det – är att visa hur den skolbildning som kallas nykeynesianism egentligen inte har så mycket med Keynes att göra utan snarare ska räknas in i traditionen neoklassisk syntes/monetarism/nyklassicism. De verkliga arvtagarna till Keynes ekonomiska idéer är olika varianter av post-keynesianism, till vilka väl också författaren själv måste räknas. Kapitlet har en tydlig doktrinhistorisk ådra och den taxonomi den mynnar ut i är i mina ögon väsentligen invändningsfri.

I kapitel 1 fördjupar författaren sin diskussion av de Keynes-inspirerade ekonomernas teori utifrån ett metodologiskt perspektiv. Eftersom Keynes själv inte lämnade ett metodologiskt manifest efter sig försöker författaren också att ge en bild – utifrån spridda nedslag i böcker, artiklar och brev – av de bärande delarna i Keynes metodologiska förhållningssätt.

I kapitel 2 – om makroekonomisk metodologi – argumenterar Jespersen utifrån ett kritisk realistiskt perspektiv för att makroekonomer i större utsträckning måste göra ontologiska överväganden och låta dessa spela roll för hur analysen av verkligheten ska designas. Studieobjektet måste i hög grad styra de valda metoderna, för att i möjligaste mån minimera gapet mellan metod och objekt.

1. Afhandlingen blev forsvaret for dr.scient.adm.-graden den 9. november 2007 på Roskilde Universitetscenter. Bedømmelsesudvalget bestod af: Professor Lars Pålsson Syll, Malmö Högskola, Professor Gudmundur Magnusson, Islands Universitet, Professor Klaus Nielsen, Roskilde Universitetscenter og Birkbeck College, University of London (udvalgets formand).

2. Redaktøren for Nationaløkonomisk tidsskrift har valgt at bringe to opponent bidrag fra hhv. anden opponenter, Lars Pålsson Syll's og opponent ex auditorium, Niels Kærgård. Formand for bedømmelsesudvalget Klaus Nielsens opposition, der indledningsvist gav en sammenfatning af afhandlingen, vil blive trykt i *Samfundøkonomen*.

Många makroekonomer – både ortodoxa och heterodoxa – är överens om detta (även om man sällan gör explicita dylika ställningstaganden). Skillnaderna mellan olika skolbildningar är i praktiken därför mer en gradskillnad än en fråga om antingen/eller. Alla forskare måste göra »avsnörningar« (som Jespersen vidare diskuterar i ett avsnitt i kapitel 6 om »semi-closures«) för att kunna göra relevanta analyser och avtäckta underliggande mekanismer i verklighetens djupa strukturer. Frågan är alltså mer vilka slutningar som är relevanta eller adekvata för analysen, än om slutningar måste göras eller inte. Detta gäller även om man på det *ontologiska* planet konstaterat att det »system« man undersöker är öppet. För att kunna genomföra en *analys* måste systemet delvis slutas (som Joan Robinson brukade säga har vi ingen glädje av en karta i skala 1:1). Detta *behöver* inte innebära att man *a fortiori* också uppfattar systemet som realiter slutet. Jag tror att även kritiska realister blivit mer mottagliga för denna argumentation (diskussionerna om betydelsen av »demi-regs« tyder på det). Lite tillspetsat skulle man kunna formulera frågan som om det går att analysera öppna system med metoder som tar sin utgångspunkt i (delvis) slutna system. En rimlig position här borde väl *prima facie* vara att få graden av slutenhet att vara så kongruent som möjligt mellan system och analys. Eller som Jespersen (s 20) uttrycker det: »hvis sigtet med analysen er at forstå væsentlige træk af den makroøkonomiske virkelighed, så er det nødvendigt at tilstræbe betydelig kongruens mellem det reale og det analytiske niveau.«

Den berättigade kritik Jespersen riktar mot ortodox makroekonomisk teori är att den ofta *oargumenterat* utgått från att verkligheten är homogen och oföränderlig. Men om den inte är det är det inte troligt att de metoder och modeller som används är adekvata/relevanta. Argumentationen är analog med den som Keynes riktade mot ekonometrin under 1930- och 40-talen och som (tyvärr) fortfarande får anses vara relevant för mycket av det som görs i modern tillämpad mikro- och makroekonometri.

Självklart måste man när man analyserar i någon mån konstanthålla saker. Frågan är bara vilka ting och för hur lång tid. Här tror jag det av Jespersen använda begreppet »path-dependence« kan hjälpa till att finna de relevanta avgränsningarna och ge en stark grund för den argumentation som alltid måste ges.

Jespersen tangerar också i kapitel 2 frågan om de teoretiska abstraktionernas status. Keynes och andra vetenskapsteoretiska realister inser naturligtvis nödvändigheten av abstraktioner, men kvalificerar detta ställningstagande med att de måste vara realistiska och inte som den deduktivistiska makroekonomin som företrädesvis bygga på isolering/idealisering där faktorer helt sonika bortses från eller där man arbetar med fiktioner bara för att få analysen att hålla ihop.

Detta sagt måste man nog ändå medge att både abstraktioner och isolering innebär något slags slutning av modellen. Keynes betonande av homogenitet/invarians (både i *A Theory of Probability* och den senare kritiken av ekonometrin) var ju implicit en

plädoyer för sökande efter denna typ av stabilitet i ekonomin (att han sen menade att den kanske inte så ofta förelåg är en annan sak). Keynes betonande av osäkerhet och hans »weight of argument«-aspekt på sannolikhetsbegreppet gjorde honom dock immun mot ett alltför långtgående användande av slutna modeller i sin analys. I en värld genomsyrad av osäkerhet kan man möjligen urskilja tendenser, men knappast händelseregelbundenheter. Därav också en större öppenhet för behovet av triangulering, fallstudier och allmän metodpluralism i den ekonomiska vetenskapen.

Popper och kritisk realism

Jespersen kommer också i kapitel 2 in på frågan om man kan förena Popper och ett kritisk-realistiskt perspektiv på makroekonomi. Jespersens diskussion är här intressant och så vitt jag kan bedöma uttryck för ett bland kritiska realister nytt positivt förhållningssätt till Popper.

Vad Jespersen tar fasta på hos Popper (se t ex *A world of propensities* (1990)) är dennes tredelning av relationen mellan teori och verklighet i World 1 (som är den verklighet vi försöker förstå och förklara), World 2 (den analytiska) och World 3 (som representerar våra teorier och föreställningar om verkligheten).

Noteras bör att beskrivningarna ovan av de tre världarna är Jespersens. Poppers egna är – se exempelvis hans Tannerföreläsningar »Three Worlds« (1978) – väsentligen annorlunda. Jespersen har uppenbarligen låtit sig inspireras av Poppers trikotomi, utan att känna sig bunden av upphovsmannens egna explikationer.

Här tror jag dock också att det är viktigt att klarare än Jespersen betona att svaret på frågan om Poppers vetenskapsteoretiska hållning kan vara något att bygga vidare på utifrån kritisk-realistiska synpunkter i stor utsträckning hänger samman med *vilken* Popper vi talar om. Om vi talar om den sene Popper som skrev exempelvis *A world of propensities* är det tänkbart. Om vi talar om den yngre Popper (-1970) är jag tveksam.

I flera av sina tidiga arbeten – t ex *Conjectures and Refutation* (1962) och *The Poverty of Historicism* (1957) – hävdar Popper med emfas att analyser av historiska och samhällsliga skeenden måste utgå från vad han betecknar som en »situationell logik«. Samhället och historien är ett så komplext och mångfacetterat system att vi omöjligt kan förutse eller förklara singulära händelser (»tokens«) utan måste inskränka oss till att med utgångspunkt från »rationalitetsprincipen« försöka konstruera modeller som möjliggör prediktion av klasser av händelser (»types«) genom att beskriva typiska »sociala situationer«.

Detta följer helt i linje med Poppers deduktivistiska inställning till vetenskapen som sådan och enkannerligen hans avståndstagande från all typ av induktion (och alla former av abduktion/retroduktion).

När det gäller förklaringar inom samhällsvetenskaperna menar Popper som sagt att vi måste skilja på om vi försöker förklara enskilda händelser eller typer av händelser.

Samhällsvetenskaperna sysslar sällan med att besvara frågor av den första typen. När de använder modeller gör de det för att besvara frågor av den andra typen och enligt Popper gör man det bäst med hjälp av *situationsanalys* som ger oss modeller av typiska sociala situationer. I dessa modeller antas aktörerna handla på modellens villkor i enlighet med *rationalitetsprincipen* att göra det som är »passande för situationen«.

Popper är metodologisk individualist, men menar att detta tvärtemot vad som oftast anförs när det gäller ändamålsförklaringar, inte medför att man inom samhällsvetenskaperna är hänvisad till en psykologisk metod. Tvärtom. Psykologismens reduktion av sociala fenomen och regelbundenheter till psykologiska dito håller inte.

Popper anför i stället sin idé om situationslogik för att förklara våra handlingar. Tanken bakom »situationens logik« är rimligen att det givet en beskrivning av den situation aktören befinner sig i ska gå en rak linje till hans handlande via antagandet om rationalitet. Det är alltså ett slags »single-exit«-lösning som söks och därför blir en beskrivning som går in på aktörens motiv, önskningar och historia inte så mycket fel som överflödigt och onödigt. Här finns en slags effektivitets- och sparsamhetstanke. Kan vi förklara handlingar med en variabel, varför ska vi då söka flera och svårtillgängliga variabler? Problemet är väl dock att det är svårt att övertygas av Poppers berättigande av ett sådant antagande.

En minst sagt knivig fråga är också hur man ska kunna förena Poppers falsifikationism och hans »metafysiska realism« och »kritiska rationalism«.

När det gäller den senare tror jag nog ingen har problem att skriva under på vikten av ett kritiskt förhållningssätt inom vetenskapen. Det förenar t o m kritiska realister och de mest utrerade socialkonstruktivisterna. Föga förvånande är också detta något Popper höll fast vid både i sitt tidigare och senare liv. Men detta ger oss tyvärr inte så många ledtrådar till hur vi ska bygga våra modeller eller genomföra våra analyser. Vad det däremot lär oss – och det är kanske nog så viktigt – är att vi måste ta kritik av våra modeller och deras tillkortakommanden på stort allvar.

Popper är känd som en av Humes skarpaste kritiker. Ett av falsifikationismens syften var att ersätta den senares (och hans efterföljares) användande av induktivism inom vetenskapen med »bold conjectures and refutations« sprungna ur den deduktiva logiken. Men om nu en teori/hypotes inte falsifierats trots upprepade »severe tests«, vad säger oss detta om teorin? Det är här problemet med realismen dyker upp, för enligt Popper kan det inte säga oss något om teorins sanningshalt. Teorin har korroberats (överlevt falsifikationsförsök) men säger oss inget om huruvida den är i överensstämmelse med verkligheten. Självklart är detta problematiskt för någon som vill både vara

falsifikationist *och* realist. Det är svårt (omöjligt) att på ett konsistent sätt vara både falsifikationist och realist. Oftast (alltid) får man välja det ena eller det andra.

För Popper är den vetenskapliga metoden i princip deduktion + experimentellt testande, där vetenskapliga teoretiska satser konfronteras med empiriska evidens och via *modus tollens* falsifieras eller inte. Kruxet ur realismens synpunkt är bara att de som inte falsifieras inte kan antas vara sanna trots att Popper menar att både teorier (universella satser) och evidens (singulära påståenden) tolkas realistiskt. Visserligen laborerar Popper också med sitt begrepp *verisimilitude*, men inte heller detta kan fungera som broslagning mellan hans falsifikationistiska epistemologi och hans realistiska ontologi (något som också den sene Popper motvilligt medgav). En realist måste kunna argumentera för att hans teorier och modeller inte bara är instrumentella fiktioner utan faktiskt också ger oss kunskap om världen sådan den är. Just detta kan Popper inte göra eftersom han bara vill medge att falsifikationsförsöken visat att några teorier »överlevt« (även om Popper också var en varm anhängare av Tarskis sanningsteori, ska man komma ihåg att han uppfattade sanningskriteriet i denna enkom som en regulativ idé).

Den sene Popper har på många sätt en annan ontologi, som ligger närmre den kritiska realismen. Han betonar nu att »vi lever i en värld av benägenheter« och att det existerar »tendenser eller benägenheter« som är »något i stil med krafter«. Världen kan inte längre uppfattas som en kausal maskin. Precis som hos kritiska realister uppfattas den som öppen och strukturerad.

Popper diskuterar också frågan hur man ska kunna *testa* teorier och hypoteser som försöker beskriva dessa strukturella egenskaper i vår värld. Han menar att testet måste ske via *experiment*.

Tyvärr ger Popper inte utöver citatet ovan uttryck för någon uppfattning om hur denna nya ontologiska uppfattning återverkar på våra förklaringsmodeller. Vad man skulle vilja veta är ju t ex om Popper menar att experimentmetoden också är applicerbar inom den ekonomiska vetenskapen.

Den sene Popper är kanske förenlig med kritisk realism, men det beror i lika hög grad på vad Popper säger som vad han *inte* säger. Ger man inte uttryck för någon deciderad uppfattning i viktiga metodologiska frågor är man så klart förenlig med det mesta i metodologisk väg. Tyvärr gör det också ens uppfattning tom.

Experiment och tänkta ekonomier

Just frågan om experimentens vara eller icke-vara inom samhällsvetenskapen i allmänhet och den ekonomiska vetenskapen i synnerhet är sedan några år tillbaka föremål för intensiva diskussioner.

Även om den experimentella ekonomin fått en allt högre status – under inflytande av arbeten av förgrundsgestalter som Charles Plott och Vernon Smith – är många fort-

farande skeptiska till den experimentella ekonomin. Framför allt grundar sig detta i ett ifrågasättande av om de resultat man får fram i experimenten är överförbara till världen utanför laboratorierna. Gapet mellan experimentsituationen och den ekonomiska verkligheten sägs vara för stort för att experimenten ska vara verkligt intressanta.

I experimentell ekonomi har laboratorieexperiment varit helt dominerande eftersom förekomsten av naturliga experiment är begränsad och fältexperiment ofta är överväldigande dyra och ibland etiskt tveksamma.

Både Plott och Smith har försvarat den experimentella ekonomin utifrån en empirisk-realistisk hållning. Om ekonomisk teori verkligen ska vara så generell som den ger sig ut för att vara, måste den också vara applicerbar i den mer speciella experimentella situationen. Att denna inte återspeglar ekonomin i stort är därför inget problem. Tvärtom är själva idén med experimenten att kunna kontrollera och konstanthålla vissa variabler för att lättare kunna se hur vissa andra variabler interagerar med varandra och kunna dra slutsatser om kausala relationer. Experimentvärlden är enklare, men lika verklig, som världen där ute.

För att experimenten ska vara intressanta måste de kunna replikeras och varieras för att se om teorier går att bekräfta eller falsifiera, om de är robusta och så vidare. Teorierna kan på grundval av experimenten förfinas och fås att bli så sanna som möjligt. Detta är en av grundvalarna för Smiths *parallism*, enligt vilken individers beteenden och institutioners prestationer som testats i laboriemässiga mikroekonomier också kan appliceras på icke-laboriemässiga mikroekonomier där »motsvarande« villkor gäller.

Experiment är ju i en uppenbar mening både artificiella och orealistiska, men parallellismen säger att situationerna ändå kvalitativt kan vara likartade. Till syvende og sidst måste man alltid testa om resultaten från de mer eller mindre kontextfria miljöer som laborier utgör också är applicerbara på »fältet«.

En del filosofer – främst Nancy Cartwright – har mot detta resonemang invänt att regelbundenheter bara går att producera under artificiella förhållanden, som i laborier, eller vad Cartwright kallar »nomologiska maskiner«. Och kritiska realister som Tony Lawson har också kritiserat experiment som kunskapsform. Experiment är ett sätt att skapa slutna modellvillkor. Att dessa skulle säga något om en essentiellt öppen verklighet är tveksamt. Händelseregelbundenheter äger bara rum när vi har »laddat tärningen«. De resultat som den experimentella ekonomin levererar kan inte utan berättigande överföras till den verkliga världen. Ju mer kontroll vi skapar i laboriet över de variabler vi är intresserade av, ju mindre applicerbara är resultaten.

Överlag kan man nog konstatera att den experimentella ekonomin har gett stöd för den kritik som framförts av Jespersen och många andra mot den ortodoxa ekonomiska teorin vad gäller antaganden om att ekonomins aktörer är rationella. I experiment visar

sig individer ofta värdera, välja och handla på ett annat sätt än teorin förutser. I olika experiment kring t ex auktionsförfarande har man kunnat visa att vad som är ekonomiskt effektivt marknadsbeteende i hög grad avhänger kontextuella förhållanden. Hur människor agerar på marknaden avhänger hur denna är organiserad.

Slutsatsen av detta resonemang måste bli att det inte längre lika lätt går att avfärda möjligheten till användande av experiment inom den ekonomiska vetenskapen som på Poppers tid. Självklart måste användandet av experiment grundas på en föregående ontologisk och metodologisk reflektion som också klargör experimentens omfångslogiska status.

Här kanske man dock också måste hissa varningsflagg när det gäller Jespersens återkommande kritik av laboratorie-metaforen, som han främst använder i kritiken av Torben Andersen när denne säger att syftet med teoretiska analyser är att »opstille et laboratorium for afprøvning af forskellige hypoteser.« Jespersen återkommer vid upprepade tillfällen till denna metafor, men den typ av laboratorium som här avses har så vitt jag kan förstå inget alls att göra med de experiment och laboratoriestudier som experimentella ekonomer sysslar med. För mig blir Andersen och andra ortodoxa ekonomers laboratorier inget annat än tänkta ekonomier.

Jespersen diskuterar också i kapitel 5 utförligt (s 174) »den laboratorieskabte virkelighed« och hur denna »idealmødel« inom generell jämviktsteori »er blevet gjort til genstandsfeltet« och där »de analytiske resultater fra 'laboratoriet' vil være udsagn om ideal-mødelens egenskaber og altså ikke med den intention at beskrive (og forstå) den 'virkelige' virkelighed.« Här kan man inte annat än instämma i en berättigad kritik. Att oargumenterat likställa modell och verklighet är helt otillfredsställande (även om det kanske kan vara av intresse att använda modeller av detta slag som »bench-marking«, men värderingen måste då rimligen bero på vad bench-markingen syftar till).

Abduktion

Jespersen kommer också i det för avhandlingen så centrala kapitel 2 in på frågan om vad för slags inferensmetod som är den mest ändamålsenliga för att förstå och förklara en komplex och stratifierad verklighet och kunna »afdaekke« dess djupa kausala mekanismer. Han beskriver denna »retroductive arbejtsform« – som byggande på (s 109) »en vekselvirkning mellem »common sense«, deduktion, observation og induktion.«

På senare år har det förekommit en omfattande diskussion bland filosofer och vetenskapsteoretiker om abduktion och dess förtjänster och eventuella brister. Traditionellt har man i huvudsak försökt karakterisera den vetenskapliga metodens basala struktur i termer av induktion och/eller deduktion. Men den vetenskapliga metoden ska helst både vara *ampliativ* – utöka våra kunskaper – och öka vår epistemiska *beviskraft* för de resultat den producerar.

Att den vetenskapliga metoden ska utvidga våra kunskaper är egentligen en självklar utgångspunkt för det kritisk-realistiskt förhållningssätt som Jespersen beskriver i kapitel 2 under rubriken »Hvad er Kritisk Realisme?« Det är dock inte alltid så lätt att förena ampliatio och epistemisk beviskraft. Vad är det som ger beviskraft åt en teori/förklaring snarare än en annan när vi försöker gå utöver våra sinnesintryck?

En rent deduktiv metod skulle försäkra oss om att slutsatsen är lika beviskraftig som de premisser den bygger på. Problemet är bara att deduktionen också är helt icke-ampliativ. Ska vi använda oss av en ampliativ metod måste vi därför acceptera att beviskraften inte kan vara av deduktiv kaliber. Våra data garanterar aldrig att endast *en* teori/förklaring är giltig. Innebär detta att *alla* teorier/förklaringar har samma grad av giltighet eller beviskraft? Är alla katter grå? Inte nödvändigtvis. Om detta är en riktig hållning kan inte enbart avgöras på formallogiska grunder utan måste föras tillbaka på hur vår värld ser ut och är strukturerad.

Induktionens problem är att den bygger ampliatio från »några« instanser till »alla« på generalisering. Denna »mer av samma« metod utökar våra kunskaper på ett enbart »horisontellt« sätt. I det avseendet innebär induktionen en minimal ampliatio av våra kunskaper. Den ger oss inga teorier/förklaringar till orsakerna bakom de observerade fenomenen. Detta är av avgörande vikt för samhällsforskare som delar med Jespersens grundsyn att (81) »den del af landskabet, som vi evner at observere, billedligt talt kun er toppen af isbjerget, idet alle de væsentlige forhold, der er gemt i undergrunden, ikke kan gængives i en landskabsmodel.«

Abduktion bygger på att vi gör en slutledning utifrån det faktum att om en viss hypotes skulle *förklara* hypotesen så är hypotesen *sann*. Generellt gäller givetvis att flera olika hypoteser skulle kunna förklara evidensen, så man måste innan man kan göra slutledningen kunna utesluta de andra hypoteserna. Olika hypoteser kontrasteras mot varandra och den som givet vår bakgrundskunskap bäst förklarar »varför x?« väljs som den troliga förklaringen. Här kan det naturligtvis inte handla om annat än sannolikhet/trolighet. I deduktionen är det uppenbart att konklusionen är nödvändig, medan i induktionen och abduktionen konklusionerna endast är rimliga/troliga/sannolika. Detta är priset vi får betala för att våga oss utanför den rena logiska formens säkerhet.

Vad jag med denna långa utläggning vill ha sagt är att om vi ska vi ta Jespersens resonemang om ekonomin som ett essentiellt öppet system på fullt allvar måste vi också acceptera att våra inferenser inte kan ha den beviskraft som deduktion kan ge (och som ortodoxa makroekonomiska modeller ofta tar sikte på), utan vi får nöja oss med möjliga och sannolika inferenser. Jag tror därför också att Jespersens diskussion om öppna/slutna modeller och »semi-closures« hade kunnat vinna på en tydligare koppling till den moderna diskussionen om abduktion (eller »inference to the best explanation« som Peter Lipton och en del andra vetenskapsteoretiker föredrar att kalla inferens-typen).

Jespersen *har* en diskussion i kapitel 2 under rubriken »Retroduction« om abduktion. Liksom en del andra kritiska realister föredrar han denna term (som liksom abduktion kommer från Peirce). Men som det ofta är när detta begrepp används skjuter man då in sig på inferensens bejakande av icke-observerbara entiteter (strukturer, krafter etc). Som kritisk realist kan man inte annat än bejaka förekomsten av sådana kausala faktorer, men jag tycker ändå att abduktionsmetoden framställs onödigt »smalt« på det viset och inte lika tydligt framstår som en bred och bärkraftig inferensmetod på nivå med induktion och deduktion.

Det makroekonomiska landskapet och osäkerhet

I kapitel 3 försöker författaren att utveckla en metodologi som ska kunna grunda vårt sökande efter de avhörande makroekonomiska drivkrafterna. Han visar övertygande att detta måste ske utifrån en ontologisk reflektion.

I kapitel 4 tydliggör författaren på ett förtjänstfullt sätt (främst de makroekonomiska) implikationerna av Keynes och postkeynesianers åtskillnad mellan osäkerhet och risk. Målet är bl a att visa på inkompatibiliteten mellan (neoklassiska) rationella förväntningsmodeller och ett erkännande av att världen genomsyras av genuin osäkerhet. I appendix diskuterar han hur man ska se på relationen mellan den till övervägande delen filosofiska *A Theory of Probability* och de senare ekonomiska arbetena. Också frågan om hur man ska uppfatta Keynes sannolikhetsbegrepp behandlas. Ska man rikta någon liten kritik mot framställningen här är det väl att Frank Ramseys (för den efterföljande diskussionen så viktiga) kritik av Keynes inte relateras.

Atomistiska felslut och representativa aktörsmodeller

I ett mycket intressant kapitel 7 behandlar Jespersen begreppet »atomistisk felslutning«. Han har i flera av sina tidigare böcker också extensivt diskuterat detta begrepp och är väl den moderne makroekonom som (efter Johan Åkerman) grundligast problematiserat det. Han argumenterar för att felslutet främst är en konsekvens av att ekonomer har en metodologisk-individualistisk utgångspunkt och använder sig av antagandet om generell jämvikt, och att en övergång från slutna till öppna modeller skulle innebära att man undviker att begå detta fel. För Jespersen bygger mycket av dagens makroekonomiska teori på otillåtna och okritiska slutningar från individ/mikro till kollektiv/makro. Om den värld vi lever i väsentligen måste betraktas som öppen – vilket Jespersen övertygande argumenterar för – kan vi inte bygga våra analyser på modeller som bortser från emergens, supersiliens och andra företeelser som direkt motstrider möjligheten av en makroekonomisk metodologisk individualism som lever upp till rimliga standarder vad gäller adekvans och relevans.

Jespersen ger i samma kapitel också ett bra exempel på vikten av en genomtänkt ontologisk positionering för makroekonomiska forskare i sin diskussion av de inom modern ortodox makroekonomi så vanliga representativa aktörsmodellerna.

Typiskt för den nyklassiska skolbildningen är – förutom att man i likhet med nykeynesianer understryker vikten av att ge mikroekonomiska fundament till makroekonomin – att man arbetar med enkla generella jämviktsmodeller där representativa aktörer har fullständig kunskap, transaktionskostnader är frånvarande och marknader är kompletta.

I modeller med representativa aktörer förutsätts alla individer vara identiska, vilket får långtgående konsekvenser för analysen. Så kan till exempel inte situationer karakteriserade av asymmetrisk information – som nykeynesianer menar är legio – uppkomma i sådana modeller. Självklart kan man som Jespersen ifrågasätta relevansen av sådana mikrofundament om tanken är att bygga en makroanalys på dem

Den nykeynesianska informationsansatsen är kritisk mot användandet av representativa aktörsmodeller eftersom dessa inte inkorporera fördelningseffekter. Ekonomins överväldigande koordinationsproblem är också svåra, för att inte säga omöjliga, att analysera med den reala konjunkturteorins »Robinson Kruse-modeller« där just de skillnader mellan olika aktörer som utgör en av drivkrafterna i de nykeynesianska analyserna är frånvarande.

Olika individers val måste kunna visas vara koordinerade och konsistenta. Detta är svårt att göra om de ekonomiska modellerna ska ge utrymme för heterogena individer. Representativa aktörsmodeller blir visserligen mer hanterliga, men inverkar också menligt på deras förklaringspotential och relevans.

Jespersen hävdar – med rätta menar jag – att mikrofundamentstanken och representativa aktörsmodeller i grunden är förfelade och att man också skulle kunna valt att acceptera att makro inte går att reducera till mikro. I verkligheten måste mikroekonomiska resonemang bygga på makroekonomiska förutsättningar.

Också ledande mikroekonomer (som Alan Kirman) har hävdad att användandet av representativa aktörsmodeller är oberättigat och kan leda till slutsatser som är både vilseledande och felaktiga. Den representativa aktören är en fiktion som makroekonomen använder för att berättiga ett användande av jämviktsanalys och ge ett slags pseudo-mikrofundament. I takt med att ekonomiska modeller blir alltmer komplexa, blir också antagandet mindre trovärdigt.

Keynes var som Jespersen riktigt påpekar anhängare av Marshalls »en sak i sänder«-metodologi. Men när makroekonomin utvecklades under efterkrigstiden kom de modeller som åberopade Keynes att mer och mer knytas till Walras och hans generella jämviktsmetodologi. Överhuvud kan man säga att formella modeller kommit att utgöra det nav runt vilket modern ortodox makroekonomi snurrar. De som hävdad att

ekonomin är för komplicerad för att säga något formellt om har i stort förblivit ohörda. De modeller som byggts och bortsett från verklighetens komplikationer har som Jespersen konstaterar inte kunnat ge några större insikter i verkliga makroekonomiers funktionssätt.

Den »nya syntesen« inom makroekonomin – med sina dynamiska stokastiska generella jämviktsmodeller – har visat att mycket av det som tidigare makroekonomiska teorier trott sig kunnat visa i själva verket bara stämmer under ännu mer långtgående och strikta antaganden. Men även om de är eleganta och konsistenta, är det mer tveksamt hur relevanta de är för förklarings-, prediktions- och policysyften. Likaså kan man starkt ifrågasätta om de har så mycket med Keynes tänkande och visioner att göra. Keynes idéer om osäkerhet går inte att reduceras till stokastiska riskfluktuationer. Den nya syntesens modeller bygger fortfarande på vad man utifrån Jespersens diskussion skulle betrakta som enormt orealistiska rationalitets- och informationsantaganden. Individer känner lika lite som ekonomer till den »korrekta« modellen (om den ens existerar). Precis som hos sina walrasianska föregångare får man intrycket att elegans och konsistens köps på bekostnad av realism.

Slutomdöme

Jag har i dessa kommentarer inte kunnat göra full rättvisa åt alla de tankedigra, spännande och i många stycken nyskapande teoretisk-metodologiska synpunkter som förs fram i Jespersens avhandling. Jag har fått nöja mig med att göra några djupdykningar på avsnitt som jag funnit vara speciellt intressanta eller tankeväckande. Mitt intryck av avhandlingen i stort är att den vittnar om stor beläsenhet och en god förmåga att tillgodogöra sig och ta fasta på de väsentliga frågor som modern makroekonomisk metodologi står inför. En avhandling av denna typ skulle lätt kunna urarta i en tråkig exponering av »Lese-früchte«. Den som behärskar sitt forskningsområde så väl som Jespersen hemfaller dock inte åt denna horrör.

Jaspersens avhandling måste ses som ett viktigt bidrag till den makroekonomiska metodologiska förnyelsen inom heterodox ekonomisk teori. Den bygger vidare på och utvecklar den givande korsbefruktning som – med arbeten av bl a Paul Davidson, Sheila Dow, Victoria Chick och Tony Lawson – nu under en tid ägt rum mellan kritisk realism och postkeynesiansk ekonomisk teori. Speciellt med sin betoning på makroekonomins metodologi har den tillfört den ekonomiska metodologin nya bärande synpunkter som varje makroekonom borde (måste) förhålla sig till i sin forskargärning.

På sina ställen har jag anfört några kritiska synpunkter. Detta ska dock mest ses som ett positivt uttryck för att varje arbete av dignitet och med en genomarbetad argumentation väcker forskarkollegors lust att klargöra sina egna försanthållanden och ställningstaganden.

Post-keynesiansk disputats gør op med »mainstream« økonomien

Niels Kærgård

Fødevareøkonomisk Institut, Københavns Universitet, E-mail: nik@life.ku.dk

Indledning

Højtærede rektor, ærede professorer og doktorer, mine damer og herrer, kære præses.

Det var den klassiske måde at indlede en opposition ved et dansk disputatsforsvar. Og Jesper Jespersen disputats er på flere måder en meget klassisk – for ikke at sige gammeldags – disputats. Det at skrive en monografi på dansk og indlevere den som disputats er ikke set i økonomi i de sidste 15-20 år.

Ved Københavns Universitet er min egen disputats fra 1991 den sidste økonomiske disputats, og hvis man ser på økonomernes sprog, så er alle de 63 ph.d.-afhandlinger, der er forsvarede i økonomi i København i det her årtusinde, skrevet på engelsk. Jesper Jespersens disputats er også en gammeldags disputats i den forstand, at den har et ekstremt klart synspunkt, som kan disputeres. Hele afhandlingen er en lang argumentation for, at den post-keynesianske skole er den mest frugtbare tilgang til økonomisk forskning.¹

Den danske pragmatisme

Med dette klare udgangspunkt i en økonomisk skole er disputatsen en utypisk repræsentant for den danske økonomiverden. De danske økonomers forhold til de internationale skoler har været lidt som sommerfuglens til blomsterne. De er normalt flakket fra skole til skole og har taget det, de mente at kunne bruge. Man finder meget få danske ortodokse tilhængere af de kendte skoler. Og sådan har det været i århundreder, se Kærgård (1997).

Det kan der gives mange eksempler på. Toldpolitikken fra toldloven af 1797 og frem til EU-tilslutningen var f.eks. et klart kompromis mellem Adam Smiths frihandelsesteorier og de tyske økonomers beskyttelsestold. Man kendte utvivlsomt begge synspunkter, men gik aldrig renlivet ind i nogen af lejrene.

Dette indlæg er en let tilrettet udgave af Niels Kærgårds ex auditorie opposition ved Jesper Jespersens disputatsforsvar på Roskilde Universitet den 9. november 2007. Peder Andersen, Anders Møller Christensen, Henrik Hansen, Claus Thustrup Kreiner og Peter Birch Sørensen takkes for nyttige kommentarer til tidligere udgaver af dette papir.

1. Selve forsvarshandlingen på RUC den 9. november indikerede også, at disputatsforsvar ikke længere er det naturlige forum for fagøkonomiske debatter. Auditoriet var fyldt med prominente folk, navnlig af den ældre generation, fra den danske samfundsdebat (Asger Baunbak Jensen, Niels Blomgren Hansen, Steen Folke, Niels I. Meyer, Bent Rold Andersen, Jørgen Rosted, Per Schulz Jørgensen, Anders Ølgaard med m.fl.), men landets to førende økonomiske universitetsinstitutter (KU og AU) sås overhovedet ikke i salen –

fortsættes næste side...

Harald Westergaard er et andet typisk eksempel. Medens den tyske metodestrid mellem den neoklassiske teoretiske skole og den historiske skole rasede stærkest i slutningen af det 19. århundrede, var Westergaard i tæt kontakt med begge skoler. Han korresponderede med Jevons om matematisk-økonomiske problemer og udgav en indflydelsesrig dansk lærebog bygget på neoklassiske principper. Men samtidig var han katedersocialismens ledende danske repræsentant, der kæmpede energisk for sociale reformer.

Et tredje eksempel er diskussionen mellem neoklassikere, monetarister og keynesianere i 1960'erne og 1970'erne. I denne periode spillede Anders Ølgaard en central rolle som formand for Det Økonomiske Råd og arbejdede der med udgangspunkt i keynesianske synspunkter, hvilket kom til udtryk i en lang række politikanbefalinger og i strukturen i Rådets økonometriske modeller, men samtidig skrev han en tyk afhandling om neoklassisk vækstteori.

Det er ligeledes typisk, at de anvendte økonometriske modeller (Det Økonomiske Råds SMEC og regeringens ADAM) er startet som rene keynesianske modeller i begyndelsen af 1970'erne, men nu uden noget stærkt brud er udbygget med bl.a. lønrelationer, så de har fået næsten fuld crowding out på langt sigt ved finanspolitisk efterspørgselsstimulering. Her har ikke været nogen metodestrid, men en gradvis pragmatisk tilpasning.

Men sådan er Jesper Jespersens disputats ikke. Han er – jeg havde nær sagt troende – post-keynesianer. Det gælder også i hele hans øvrige produktion. Man skal faktisk – måske med undtagelse af nogle enkelte CEPOS-liberalister – tilbage til 1970'ernes marxister for at finde lige så skoletro økonomer.

Det sande keyneske og »mainstream økonomien«

Man kan diskutere Jesper Jespersens disputats på to måder. *Enten* givet de post-keynesianske præmisser diskuterer, hvor meget videre disputatsen bringer denne skole, *eller* udefra diskutere den post-keynesianske skoles stærke og svage sider. Det er det sidste, jeg vil gøre. Jeg er en typisk pragmatisk dansk økonom, der somme tider er blevet omtalt som en, »der ikke har glemt, hvem Keynes var«. Men jeg føler mig ikke

fortsat...

de optrådte kun som skydeskive in absentia. Det var helt modsat f. eks. debatten i 1940'erne og 1950'erne; ved Brems' disputatsforsvar i 1950 kom Århus-økonomerne næsten bogstaveligt talt i busser for at nedgøre københavnerne håb. På RUC i 2007 var det præsentation af et debatindlæg på hjemmebane. Jesper Jespersen præsenterede først sine synspunkter i en fuld time og blev takket med et langt bifald fra auditoriets side (glemt var Hector Estrups bidende bemærkning, når han styrede disputatsforsvar: »Man klapper af professor Tribini på Bakken, men ikke af universitetets professorer, når de diskuterer videnskabelige spørgsmål«). Til gengæld blev ex auditorie opponenterne skåret ned til en halv time, hvoraf 10 minutter var reserveret præses til svar. Det var ikke den kritiske debat, der var i højsædet den dag.

rigtig hjemme i nogen af Jesper Jespersens kasser, hverken hos »gammel-keynesianerne«, »ny-keynesianerne« eller »post-keynesianerne«.

Det er et af problemerne ved disputatsen, at alle andre skoler i så høj grad optræder som karikaturer og skydeskiver. Hele bogen igennem bliver moderne mainstream makroøkonomi repræsenteret stort set alene ved Hans Jørgen Whitta-Jacobsen og Peter Birch Sørensens lærebog beregnet til 2.-3. års studerende og så Torben M. Andersens oversigtsartikel beregnet til efteruddannelse af ældre økonomer. Det er ikke noget bredt og nuanceret billede af mainstreamøkonomi, man derved får.

Desuden er Jesper Jespersen næsten lige så sikker på, hvad Keynes ville mene, som Søren Krarup er på, hvad Luthers synspunkt ville være. Jesper Jespersen er meget sikker på, hvordan Keynes skal fortolkes, og det skal jeg klogeligt holde mig fra at diskutere – det er for meget udebane for mig. Jeg vil dog lige pege på det lidt specielle i, at Hicks' »neoklassiske syntese« med IS-LM-diagrammet kom i artiklen »Mr. Keynes and the »Classics« i *Econometrica* i 1937 (altså ni år før Keynes' død) uden at Keynes selv protesterede, men ikke desto mindre bliver den af Jesper Jespersen betragtet som en klar fejltolkning. Jesper Jespersen har selv den meget ærlige formulering:

John Hicks' IS-LM-diagram, der blev præsenteret i efteråret 1936, Hicks (1937) kom i de første ti år til at stå som eksponent for den fællesmængde som keynesianerne – i skarp opposition til datidens mainstreamøkonomer – kunne enes om var hovedbudskabet i Keynes' nye makroteori. ... Det bør også nævnes, at Keynes ikke tog afstand fra Hicks' fremstilling, hvilket han måske burde have gjort. [Jespersen, 2007, s. 50-51].

Jesper Jespersen har også – efter en beskrivelse af hvorledes der er en slugt mellem keynesk og neoklassisk teori – en sigende formulering om, at hvad der er sand »keynesianisme«, og hvad Keynes egentlig mente, først blev klarlagt adskillige årtier efter hans død:

For uanset Keynes' advarsel om mangler på realisme, så var hans nye metodiske ståsted endnu ikke blevet uddestilleret. Det blev det først som konsekvens af den post-keynesianske teoriudvikling fra 1970'erne og fremefter. [Jespersen, 2007, s. 273].

Jesper Jespersens forhold til det keyneske minder om kirkearkitekten H. B. Storcks forhold til romansk kirkearkitektur. H. B. Storck stod i sidste halvdel af 1900-tallet for restaureringen af en lang række danske kirker, bl.a. Viborg Domkirke. Disse kirker bestod ofte af en autentisk middelalderlig romansk kerne med gotiske og senere tilbygninger og modifikationer; Storck rev ofte det meste af det autentiske ned for at opføre, hvad han anså for en stilren romansk kirke. Det er spørgsmålet om ikke Jesper Jesper-

sens keyneske økonomiske teoribygning i lige så høj grad som Storcks kirkebygninger savner historisk autenticitet.

Det kunne måske også være værd at nævne, at den kendte teoriehistoriker Robert Dimand, der bl.a. har skrevet en bog om »The origins of the Keynesian Revolution«, har et ganske interessant citat af Harry D. Johnson, for Harry D. Johnson er ingen hr.-hvem-som-helst. Robert Dimand skildrer ham som »a Chicago open-economy macroeconomist who had been a Keynesian in Chambridge, who had been a Chambridge Apostle, who had heard Keynes present his last paper and had taken tea with Keynes' parents, and who was married to and writing with the editor of four volumes of The Collected Writings of John Maynard Keynes«, der derfor naturligvis »was moved to reflect on parallels and contrasts between the historical phenomena of Keynes and his followers at Cambridge and of monetarism in Chicago« (Dimand, 2001, s. 668-669). Resultatet af Harry D. Johnsons overvejelser er følgende:

There are really two theories in the General Theory. One is the theory that, given the money wage rate, the equilibrium levels of aggregate income and the rate of interest are determined by the propensity to consume, the investment demand schedule, liquidity preference, and the quantity of money, the first three being stable functions in the short run. The other, which finds its fullest expression in the »Notes on the Trade cycle«, is that all three of the fundamental propensities are highly unstable, under the influence of changing expectations about the future. It is the first theory, of course, which has become the Keynesian theory, inevitably so because it alone offers a system of relationships amenable to theoretical manipulation and empirical application. The second theory, carried to its logical extreme, would amount to a negation of the first theory. This was clearly not Keynes' intention; but he offered no formal analysis of the formation of expectations, and it was left to later writers to develop his short-period analysis into theoretical and econometric models of cycles and growth. [Johnson, 1962, se Dimand 2001 s. 676].

Men jeg vil som sagt ikke blande mig i debatten om, hvad Keynes egentlig mente, blot påpege, at Jesper Jespersens tolkning langtfra er enerådende. Hvor det kan diskuteres, hvad Keynes egentlig mente, så er der derimod ingen tvivl om, at den keyneske revolution, der slog igennem i 1940erne og 1950erne, var noget helt andet end Jesper Jespersens »post-keynesianisme«. Det var den del af Keynes, som Harry D. Johnson kunne lide. Det var Hicks' IS-LM-model, Tinbergen og Kleins makroøkonometriske Keynes-modeller og de klare politik-konklusioner af disse modeller. Klein, der fik en Nobelpris for sine økonometriske modeller, publicerede bogen »The Keynesian Revolution« i 1947 samtidigt med, at han arbejdede med de økonometriske modeller, som »kan opfattes som et forsøg på verifikation af de keyneske teorier« (Andersen, 1975, s. 45). Var Keynes først og fremmest kommet til at stå for den post-keynesianske me-

todologi, der beskrives af Jesper Jespersen, så er det et spørgsmål, om der nogensinde var kommet en keynesiansk revolution.

Jesper Jespersens keynesianisme er altså en speciel »post-keynesiansk« form for keynesianisme, og den lægges op som en klar kontrast til al mainstream-økonomi. Gammel-keynesianere, monetarister, nyklassikere og nykeynesianere slås sammen i en kasse over for alternativet postkeynesianerne I-III (s. 40). Side 49 anføres f.eks:

Gammel- og nykeynesianerne på den ene side og post-keynesianerne på den anden side benytter sig af en fundamental forskellig makroøkonomisk metode, der udspringer af to distinkte videnskabsteoretiske ståsteder, hhv. *idealisme* og *realisme*. [Jespersen, 2007 s. 49].

Og side 50:

Gammel- og nykeynesianerne har primært fokuseret på, hvorledes teorien om den effektive efterspørgsel, som de anser for den væsentligste arv fra Keynes, kan inkorporeres i en generel ligevægtsmodel [Jespersen, 2007 s. 50].

Også politikmæssigt ser Jesper Jespersen ingen store forskelle på ny-klassikere og ny-keynesianere; de har alle i større eller mindre grad interesseret sig for struktur-reformer:

...vedrørende strukturreformer på arbejdsmarkedet (større fleksibilitet og mindre velfærdsydelser). Her er det for alle praktiske formål umuligt at skelne mellem anbefalinger fra monetarister og fra ny-keynesianere. [Jespersen, 2007, s. 57-58].

I Jesper Jespersens optik bliver alle andre end post-keynesianerne til en mere eller mindre sammenhængende teoribygning af »mainstreamøkonomer«, der bygger på individuel optimering, repræsentative agenter, generel ligevægt og rationelle forventninger. Det er meget vanskeligt ikke at betragte dette billede som en karikeret stråmand, der ikke har andet formål end at blive skudt ned. Med nogle få undtagelser er alle de centrale økonomiske debatter og kontroverser fra 1945 til i dag foregået inden for det, Jesper Jespersen slår sammen til »mainstreamøkonomi«.

Det post-keynesianske alternativ

Over for denne skydeskive opstiller Jesper Jespersen så den post-keynesianske teori, der består af en række centrale byggesten:

- En genuin usikkerhed, der ikke kan reduceres til sandsynlighedsregning.
- En fundamental makrobetragtning, »the economy as a whole«, som er forskellig fra »the sum of the parts« – økonomien som et landskab.

- Åbne systemer, der er afhængige af den historiske udvikling (path-dependency), og ikke ligevægtsmodeller.

Dette giver en »realisme«, der tager udgangspunkt i »den eksisterende, sociale virkelighed«. Jesper Jespersen bruger (som flere af skolens tilhængere) betegnelsen »kritisk realisme« om denne metode. Han er ganske kritisk over for modeller og udvikling af en modelramme, der kan fungere som et »laboratorium«. Laboratorieforsøg og generelle ligevægtsmodeller, »den laboratorieskabte virkelighed« og lignende er flere gange brugte kritiske karakteristika af f.eks. ny-keynesianerne. Han citerer med klar afstand-tagen Torben M. Andersens formulering:

Formålet med teoretiske analyser er at opstille et laboratorium for afprøvning af forskellige hypoteser, [Andersen, 2000, se Jespersen, 2007, s. 55].

Men dette er jo langt fra noget neoklassisk synspunkt. Ragnar Frisch – der var en indædt modstander af at overlade ressourceallokering og indkomstfordeling til det frie uregulerede marked og langt fra var nogen neoklassiker – har en næsten parallel formulering:

I vor tanke skaber vi en liten modellverden for oss selv. En modellverden som ikke er mer komplisert enn at vi kan overskue den. Og som er konstruert slik at den gir faste holdepunkter for tanken. Og så tager vi for oss å studere den modellverden i stedet for den virkelige verden. Dette lille trick er det som konstituerer den rasjonelle forskningsmetode [Frisch, 1928/29, citeret efter Kærgård, 2000, s. 333].

Usikkerhed, makroøkonomi og path-dependency kan være meget fornuftigt. Det er ikke det, der er problemet. Problemet er, at fremstillingen bliver så snæver. Diskussionen bliver holdt inden for, hvad post-keynesianerne og Keynes har sagt om disse ting, og på alle disse områder er der mange værdifulde »ikke-keynesianske« bidrag. Det kan næppe være frugtbart på alle disse områder at tage udgangspunkt i et økonomisk forfatterskab, der er afsluttet for 60 år siden.

Vedrørende *usikkerhed* så bliver det fremstillet som et keynesk fænomen, f.eks. s. 63:

Den erhvervsmæssige kerne, der således adskilte Keynes fra hans forgængere (eller rettere samtidige) var navnlig fænomenet »usikkerhed«. Begrebet »Usikkerhed« har en helt anden epistemologisk betydning, når en generel ligevægtsmodel anvendes som en (relevant) analytisk ramme. I så fald kan usikkerheden reduceres til risiko, der bl.a. med henvisning til »de store tals lov« yderligere kan reduceres til sikker (statistisk) viden på makroplan. [Jespersen, 2007, s. 63].

Men hele analysen af forskellen mellem egentlig usikkerhed og statistisk risiko stammer først og fremmest fra Frank Knight's *Risk, uncertainty and profit* fra 1921, og Frank Knight var ikke keynesainer – langt fra – han var Chicago-skolens grundlægger og dermed en slags bedstefar til monetarismen. Han nævnes da heller ikke af Jesper Jespersen. Men han er en indikation af, at der er en helt stor teoribygning om »egentlig« usikkerhed uden for den post-keynesianske skole. Også hele teorien om forsigtighedsprincippet i miljøøkonomien er et forsøg på at kæmpe med en optimerende agents håndtering af egentlig usikkerhed, se f.eks. Andersen (1998 og 2000) og Bichel-udvalget (1999). Men heller ikke denne litteratur nævnes.

Lige sådan med *ligevægt og åbne systemer*. Under overskriften »To adskilte positioner« s. 197 sondres mellem »den neoklassiske laboratorium-model (hvor) ligevægt (er) et veldefineret matematisk begreb« og post-keynesianerne og Keynes, der bruger equilibrium om en empirisk eller teoretisk begrundet »stilstand«, der ikke nødvendigvis er forårsaget af markedsclearing, der er blevet afløst af »path-dependent« analyseformer (Jespersen, 2007, s. 198).

Men det, at der er mange forskellige ligevægtsbegreber, er jo langt fra noget nyt eller post-keynesiansk. Mainstream-økonomien (i Jespersens brede forstand) har talt om markedsligevægte med individuelt maksimerende agenter og markedsclearing, om makroligevægte i den simple Keynes-models 45-graders-diagram uden maksimerende agenter og med arbejdsløshed og om arbejdsmarkedsligevægte i fagforeningsmodeller med maksimerende agenter og arbejdsløshed. Allerede i 1960erne illustrerede Hector Estrup ligevægtsbegrebets mangetydighed med historien om Mars-manden, der blev ansat i Konkurrencestyrelsen og sendt ned på Nørregade for at finde ud af, om der var ligevægt, og det rapporterede han, at der var. Da så chefen, formodentlig en jordisk neoklassiker, kom ned på gaden, så han lange køer overalt og bebrejdede Mars-manden, at han kaldte det ligevægt. Estrups Mars-mand svarede: »Der er da ligevægt; de står jo roligt i køerne«.

Man kan altså definere alle mulige former for ligevægt, og det er næppe frugtbart at indskrænke ligevægtsbegrebsdiskussionen til for eller imod markedsclearing og så bruge det til at sondre mellem neoklassikere og keynesianere. Der eksisterer en langt alsidigere debat: F.eks. forsøgte Poul Buch Hansen på Marstrand-mødet i 1971 at argumentere for helt at undlade at bruge ligevægtsbetingelser og i stedet lade beholdningsuligevægte indgå i adfærdsrelationerne, f.eks. lade lagerbeholdningerne indgå i udbudsfunktionen. Markedsclearing har så en relevant mening, hvis udbuddet justeres hurtigt med henblik på nedbringelse af lagrene.

Diskussioner af ligevægtsbegrebet er altså ikke nogen specielt post-keynesiansk position, og det er path-dependency heller ikke. Hele den omfattende økonometriske litteratur om integration og kointegration fra de sidste 20 år handler om at give begre-

bet »path-dependency« en præcis mening. Tilsvarende er den omfattende arbejdsmarkeds litteratur om hysteresis et andet eksempel på path-dependency i teoretiske og empiriske modeller, der intet har med Keynes at gøre. Men heller ikke denne litteratur er diskuteret i disputatsen.

Hele ideen om, at de økonomiske systemer flakker uforudsigeligt rundt og er systematisk ustabile, passer dårligt med en lang række empiriske undersøgelser. Der er i økonomien nogle forbløffende stabile sammenhænge. I Kærgård (1991) undersøges en række økonomiske sammenhænge, og der findes en del adfældsrelationer, der synes uændrede hen over to verdenskrige, og tilsvarende findes relationer estimeret på perioden 1878-1970, som er i stand til at forudsige udviklingen for 1971-81. Analyser af de økonomiske prognoser, der laves af Det Økonomiske Råd, indikerer heller ikke, at det økonomiske system er fuldstændigt uforudsigeligt. Der kan komme overraskelser, men stort set er prognoserne 1-2 år frem i tiden rimeligt sigende og i hvert fald væsentligt bedre end naive random-walk-forudsigelser, se Kærgård (1999). Det er ikke tilfældigt, at man omkring 30 år efter de økonometriske makromodellers storhedstid stadig bruger ADAM og SMEC som arbejdsheste hos de centrale aktører ved tilrettelæggelse af den danske økonomiske politik.

Kan katten fange musen?

Økonomisk teori skal først og fremmest kunne bruges til at besvare relevante spørgsmål om »det økonomiske landskab«. Her er Jesper Jespersens overordnede metodologiske standpunkt i mange situationer ikke til megen hjælp. Hvilken metode, det er hensigtsmæssigt at bruge, kan i høj grad afhænge af, hvilket problem der skal løses. Modellerne er netop modeller, og det vil sige, at de ikke er hele sandheden, for så var de jo virkeligheden. Selv den mest nøjagtige skibsmodel falder for det simple test, at modellens størrelse er meget forskellig fra det virkelige skib. Bare det at veje skibmodellen viser, at den er signifikant forskellig fra virkelighedens skib.

Det gør imidlertid ikke modellerne ubrugelige. Men det indikerer, at man skal bruge meget forskellige modeller og metoder efter hvilke problemstillinger, der skal analyseres. Her stiller »mainstream«-økonomien med en stor kasse med mange redskaber. Det kan være en neoklassisk baseret cost-benefit-analyse, der forudsætter, at priserne afspejler nytten og omkostningerne. Det kan være en valutakursteori, der siger, at valutakurserne på lang sigt afspejler købekraftspariteter, men ikke siger meget om kurserne på kort sigt. Det kan være en neoklassisk vækstteori, der ser bort fra de kortsigtede konjunkturbetingede udsving i kapacitetsudnyttelsen. Det kan være en »gammelkeynesiansk« konjunkturteori, der forklarer cykler i den økonomiske vækst og i arbejdsløsheden, men ikke siger noget om kapacitetsudviklingen på langt sigt. Der er et væld af metoder fra spilteorien, økonometrien og den eksperimentelle økonomi.

Forholdet mellem disse metoder og modeller har givet og giver anledning til debatter, konflikter og kompromisser. Men det meste af alt dette slås af Jesper Jespersen sammen til en »neoklassisk mainstream«-metode, som han forkaster. Eller det vil sige, det er lidt uklart, hvor meget af det, han vil erstatte med en »post-keynesiansk metodologi«. Betyder det »åbne« ikke bare, at han som næsten alle traditionelle modeller vil have nogle eksogene variable? Og nok tales der om »the economy as a whole«, men »konkrete afsnøringer« er acceptable, og hvordan afviger det så fra, hvad vi plejer at gøre?

Enhver økonomisk metode skal vurderes på, om den kan give ny indsigt i økonomiens funktion og nye resultater, der kan bruges til en hensigtsmæssig tilrettelæggelse af de økonomiske institutioner og den økonomisk politik. Som den pragmatiske kinesiske leder Deng Xiaoping sagde for mange år siden, så er det afgørende ikke, om katte er hvid eller sort, men om den kan fange mus.

Men i disputatsen er der intet spor af konkrete økonomiske analyser eller henvisninger til sådanne post-keynesianske økonomiske analyser. I disputatsen skrives der meget om realisme, men bogen er helt rensset for eksempler på brugen af metoden på virkeligheden. Hvordan vil den konkrete post-keynesianske analyse se ud? Kan den post-keynesianske kat fange mus? Det er det, der må afgøre, om denne disputats er et nyttigt bidrag til den økonomiske metodes udvikling.

Litteratur

- Andersen, E. 1975. *Træk af makroøkonomiske modellers historie og udvikling*, Akademisk Forlag, København.
- Andersen, P. 1998. Kan man bruge økonomisk teori i forbindelse med forsigtighedsprincippet? i *Miljønyt* nr. 31 s. 36-39, Miljøstyrelsen, København.
- Andersen, P. 2000. Forsigtighedsprincippet økonomi, *Samfundsøkonomen*, Nr. 1 s. 27-30.
- Andersen, T. M. 2000. Makroteori, i Chr. Hjorth-Andersen, red. *Udviklingslinier i økonomiske teorier*. DJØFs Forlag. København.
- Bichel-udvalget. 1999. *Rapport fra underudvalget om Produktion, Økonomi og Beskæftigelse*, København.
- Dimand, R. W. 2001. Harry G. Johnson as a Chronicler of the Keynesian Revolution, *American Journal of Economics and Sociology*, Vol. 60 (3), s. 667-91.
- Jespersen, J. 2007. *Makroøkonomisk metodologi – i et samfundsvidenskabeligt perspektiv*, Jurist- og økonomforbundets Forlag, København.
- Knight, F. 1921. *Risk, uncertainty and profit*, London school of Economics, London (senere udgaver Kelley 1957 og Harper 1965).
- Kærgård, N. 1991. *Økonomisk Vækst – En økonomisk analyse af Danmark 1870-1981*, Jurist- og Økonomforbundets Forlag, København.
- Kærgård, N. 1997. Karakteristiske træk ved dansk nationaløkonomi 1750-1997, *Nationaløkonomisk Tidsskrift*, bd 135 s. 32-43.
- Kærgård, N. 1999. Økonomiens elendighed: Myte eller realitet, *Nationaløkonomisk Tidsskrift*, bd 137 s. 20-32.
- Kærgård, N. 2000. Osloskolen – sønorsk eller tidstypisk skandinavisk? *Nationaløkonomisk Tidsskrift*, bd 138 s. 329-48.
- Sørensen, P. Birch og H. J. Whitta-Jacobsen. 2005. *Introduction Advanced Macroeconomics: Growth & Business Cycles*, MacGraw-Hill Education, Maidenhead.

Replik

Makroøkonomisk metodologi – i et samfundsvidenskabeligt perspektiv. Et svar til Lars Pålsson Syll og Niels Kærgård

Jesper Jespersen

Roskilde Universitetscenter, E-mail: jesperj@ruc.dk

Makroøkonomisk metodologi

The composition of this book has been for the author a long struggle of escape...from habitual modes of thought and expression. The ideas which are expressed so laboriously are extremely simple and should be obvious. [Fra Keynes' forord til *The General Theory*, s. viii].

Afhandlingens hovedtema er videnskabsteori. Sigtet er at diskutere sammenhængen mellem metodevalg og samfundsvidenskabelige resultater – dvs. metodologiens betydning. Dette gøres inden for rammen af den makroøkonomiske disciplin. Det indebærer, at de eksempler, jeg benytter til at illustrere den overordnede problemstilling, helt overvejende er af makroøkonomisk karakter. Metodediskussionen vil dog kunne overføres til andre samfundsvidenskabelige discipliner og har derfor også et tværvideenskabeligt perspektiv.

En af afhandlingens hovedkonklusioner er påvisningen af, at der inden for makroøkonomisk forskning opnås markant forskellige resultater afhængig af, hvilken metode, der anvendes, hvilket nødvendigvis må lede til forskellige politikanbefalinger. Af helt afgørende betydning for vurderingen af metodens relevans er graden af kongruens mellem genstandsfeltet (den makroøkonomiske virkelighed) og analysemetoden. Derfor er det vigtigt, at den anvendte metode underkastes en grundlæggende vurdering, inden analysen gennemføres navnlig med henblik på at afklare metodens relevans. Dette er ikke noget simpelt spørgsmål at afklare, idet valget af metode i det mindste delvist bør være bestemt af genstandsfeltets fremtrædelsesformer – i dette tilfælde makroøkonomiens sociale ontologi. Til bestemmelse heraf argumenterer jeg for, at der med fordel kan gennemføres en ontologisk refleksion på baggrund af makroøkonomiens topologi. Hvordan fremtræder den makroøkonomiske virkelighed? Er det muligt at opridsse et ma-

Jeg har haft stort udbytte af Bjørn-Ivar Davidsens kommentarer til forholdet mellem den »tidlige«Karl Popper og kritisk realisme fremført uden for den officielle forsvarshandling.

kroøkonomisk landskab, der kan danne udgangspunkt for den efterfølgende metodiske refleksion? Ifølge sagens natur kan der kun være tale om en kontekstafhængig skitse, idet det billedligt talt gælder for de makroøkonomiske sammenhænge, som for beskrivelsen af et isbjerg af kun en beskedent del er umiddelbart synlig. Denne manglende synlighed må indebære, at en væsentlig del af den makroøkonomiske analyse består i at afdække de underliggende strukturer, magtpositioner, konventioner og institutionelle forhold, der alle er medbestemmende for landskabets konkrete form og – ikke mindst – for dets forandring.

Forståelsen af det makroøkonomiske system vil må denne baggrund altid være karakteriseret ved genuin usikkerhed. Et forhold, der er uundgåeligt i og med, at fremtiden er ukendt, hvilket indebærer, at de økonomiske aktører i ordets egentligste forstand famler (delvist) i blinde. Aktørerne kender ikke fremtiden, ligesom de kun i varierende grad kender konsekvenserne af deres handlinger. Der kan være ganske stor forskel på, hvad der forventes, og hvad der bliver realiseret. De ved aktørerne godt; men det fritager dem ikke fra at træffe dispositioner, der rækker ud i den usikre fremtid. Hertil kommer forskernes begrænsede mulighed for overhovedet at erkende aktørernes beslutningsgrundlag og konsekvenserne heraf i form af aggregeret økonomisk adfærd. En væsentlig forskningsmæssig opgave ligger derfor i at udarbejde en analysemetode, der på en adækvat måde kan sikre, at analyseresultaterne har tilstrækkelig affinitet til den makroøkonomiske virkelighed, som udgør genstandsfeltet. Det bliver således argumenteret i afhandlingen, at enhver makroøkonomisk analyse bør gøres kontekstafhængig. Der vil aldrig være to makroøkonomiske landskaber, der er ens, hverken på overfladen eller i de underliggende strukturer, der ydermere kun er delvist erkendbare. Hertil kommer, at ethvert makroøkonomisk udviklingsforløb vil være path-dependent – det vil sige sti-specifikt, hvilket indebærer at der heller ikke vil være to udviklingsforløb som er ens – selv om initialbetingelserne helt undtagelsesvist måtte være identiske.

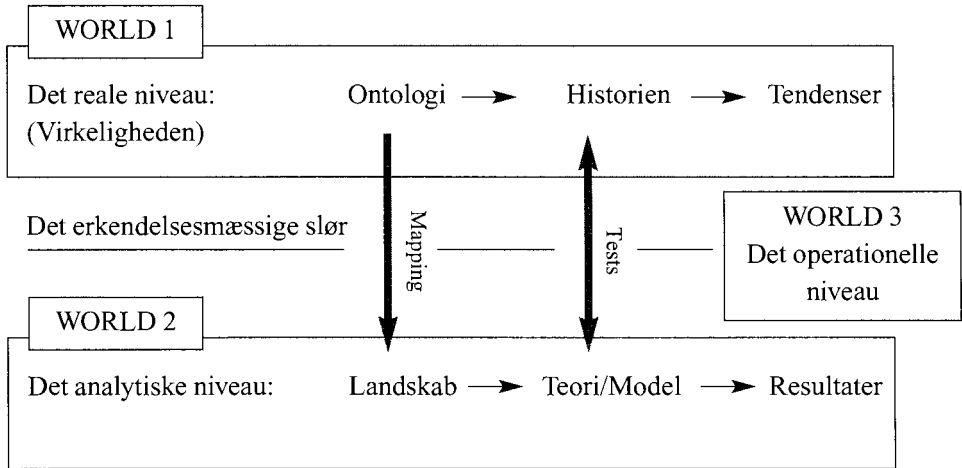
Det analytiske felt benævner jeg i afhandlingen World 2 til forskel fra den virkelighed, som vi ønsker at forstå, men som vi aldrig vil få den fulde og hele forståelse af og som er benævnt World 1. For at komme fra World 1 til World 2 må der foretages en række begrænsende forudsætninger vedrørende karakteren af det makroøkonomiske landskab og af de kræfter, der er (med)bestemmende for dets udvikling i et konkret tidsmæssigt forløb. Disse forudsætninger øver ligesom valget af metode naturligvis indflydelse på de analyseresultater, der opnås. De vil være model- og metodeafhængige og derfor ikke direkte udsagn om World 1. Det er vigtigt, at der sikres den størst mulige kongruens mellem disse to verdener, hvilket kan ske gennem empiriske tests af de hypoteser, der foreslås anvendt i World 2. Empirisk bekræftede teorier er således sikret at have en vis forankring i det konkrete makroøkonomiske landskab, hvilket er med til at sikre relevans for analysen af et faktisk udviklingsforløb. Den afgørende for-

skel mellem realisme (bekræftede teorier) og idealisme (hypoteser) er indeholdt i det ubetingede krav om empirisk verifikation/falsifikation. Uden en sådan empirisk forankring kan »anything goes«, hvilket i så fald gør det umuligt at sondre mellem videnskab og ideologi. Det er den empiriske bekræftelse, der gør hypoteser til teorier om virkeligheden. Denne fremgangsmåde, hvor der veksles mellem induktion og deduktion benævnes ofte retroduktion (eller abduktion), hvor analyseresultaterne er kontekst-afhængige.

Men selv når der er sikret en vis empirisk forankring, kan kvaliteten af de bekræftede teorier naturligvis altid diskuteres yderligere, idet World 2 aldrig kan blive gjort identisk med World 1. Derfor kan analyseresultater ikke direkte benyttes til forståelse af World 1 sammenhænge. De må efterfølgende tolkes i lyset af de gjorte forudsætninger og den anvendte metode, hvorved forskeren overfører sine resultater fra World 2 til World 3 gennem en vurdering af forudsætningernes betydning for de opnåede analyseresultater. Forskningsresultater kan i den forstand aldrig være fuldstændigt objektive. Der vil altid være et element af personlig fortolkning ved overgang fra World 2 til World 3, hvilket det naturligvis også er vigtigt at informere om ved fremlæggelsen af de endelige forskningsmæssige resultater (og evt. politikanbefalinger).

Denne forskningsmæssige fremgangsmåde (metodologi) med at opdele processen i tre »worlds«: den reale, den analytiske og den strategiske »verden« er i høj grad inspireret af den videnskabsteoretiske retning, der benævnes *Kritisk realisme*, som blandt andet Tony Lawson (1997 og 2003) og Sheila Dow (2002) m.fl. står som eksponenter for. Nogle væsentlige pejlemærker herfra kan sammenfattes således: for det første at den sociale ontologi (virkeligheden) er stratificeret, hvoraf kun en lille del, som nævnt, direkte kan observeres. For det andet, at den sociale ontologi epistemologisk må opfattes som et åbent system, og endelig bør realismen søges sikret gennem empirisk forankring af den analytiske model (og metode).

Kritisk realisme er med dette udgangspunkt udviklet i modsætning til på den ene side logisk positivisme og på den anden side konstruktivisme med svag (eller ingen) empirisk forankring. Desuden har jeg fået en betydelig inspiration fra den mere modne Karl Poppers værker, hvor han eksplicit forholder sig til de videnskabsteoretiske problemer, der knytter an til samfundsvidenskaberne. Han er her fuldt opmærksom på, at de sociale ontologier er fundamentalt forskellige inden for naturvidenskab og samfundsvidenskab. Levende væsener reagerer delvist på baggrund af selvrefleksion, såkaldt situationslogik – i modsætning til fysiske fænomener. Inden for den samfundsvidenskabelige forskning stiller Popper derfor mindre skarpe krav om en egentlig falsifikation som det ultimative kriterium for erhvervelse af ny viden; men fastholder dog ubetinget kravet om en form for empirisk forankring. Det er således åbenbart, at Pop-



Figur 1. Makroøkonomisk Metodologi baseret på kritisk realisme.

Note: Popper (1999) benytter også en kategorisering i form tre »verdener«; dette har inspireret mig til World 1 og World 2, der er stort set sammenfaldende med Poppers brug af begreberne; men min World 3, det operative niveau, hvori analyseresultaterne fortolkes til strategisk brug, afviger fra Poppers World 3.

Kilde: afhandlingen, s. 83.

per hverken er naiv empirisk eller logisk positivist. Han formulerer selv sin position på følgende måde:

To be a »positivist« is tantamount to being an opponent of all philosophical speculation and especially an opponent of realism, Popper (1999:23).

Denne mere nuancerede udgave af Poppers kritiske rationalisme har også opblødt den tidligere afvisende holdning som toneangivende kritiske realister har indtaget over for Poppers videnskabelighedskrav. Lawson er med tiden blevet mere positiv i sin vurdering af Karl Poppers videnskabsteoretiske krav, hvilke han bl.a. har formuleret således:

I think that it entails replacing the Popperian of modern mainstream economists with the quite different Popperianism of Karl Popper, Lawson (2002:27)

Der synes således at kunne etableres en ikke ubetydelig videnskabsteoretisk fællesmængde inden for den samfundsvidenskabelige metodologi mellem kritisk realisme og kritisk rationalisme. Et forhold der er blevet sammenfattet af en af Poppers elever på følgende vis:

Those readers with a Popperian background have always taken »critical realism« for granted... Basically, the main question is: do model's assumptions truly represent reality, that is, represent the real, objective world? Boland (2003:284)

Lars Pålsson Syll stillede i sin opposition bl.a. det vigtige spørgsmål, om laboratorie-metodologien overhovedet er anvendelig inden for økonomi mere generelt? Det har jeg ikke direkte forholdt mig til i afhandlingen, idet mine eksempler alle er af makroøkonomisk karakter. Her gælder det – for øvrigt ligesom inden for astronomi – at der ikke kan gennemføres kontrollerede eksperimenter.

Makroøkonomiske skoler

[This book] is an attempt by an economist to bring to an issue the deep divergences of opinion between fellow economists which have for the time being almost destroyed the practical influence of economic theory, and will, until they are resolved, continue to do so, (The General Theory, s.vi)

Niels Kærgård (NK) er i sin opposition skeptisk over for min opdeling af makroøkonomer bestemt af deres tilknytning til forskellige skoler. Han mener, at det bryder grundlæggende med den danske tradition for et blandet metodisk og teoretisk tilhørsforhold. I hans optik ser han den danske økonomiverden som bestående helt overvejende af en broget skare af pragmatikere. Deri har han muligvis ret, men i så fald ligger nok noget af forklaringen på den manglende internationale gennemslagskraft, som kendetegner dansk økonomisk forskning måske netop i dette uklare tilhørsforhold. Ser vi på vore to skandinaviske nabolande, så er situationen unægtelig en anden. Her var hhv. Stockholm- og Osloskolen dominerende og netop karakteriseret ved en betydelig vægt på at gøre den anvendte metode eksplicit.¹

Som allerede nævnt er mine eksempler helt overvejende taget fra den makroøkonomiske tradition. Giver det her mening at opdele den makroøkonomiske litteratur i »skoler«. På baggrund af et omfattende litteraturstudium af forskellige teoretikers metodevalg er mit svar et entydigt »ja«; for det bidrager til at tydeliggøre en vigtig pointe. Også her står jeg i en betydelig intellektuel gæld til Keynes, idet han i 1934 påviste en markant skillelinje (»gulf«) inden for den makroøkonomiske forskning:

On the one side [of the gulf] are those who believe that the existing economic system is, in the long run, a self-adjusting system, though with creaks and groans and jerks

1. Med mulig undtagelse af Frederik Zeuthen, (H. Brems i *Danske økonomer*, Socialøkonomisk Forening, København, 1974) fra den ældre generation og Karl Vind og Birgit Grodal fra nyere tid. Disse tre økonomer var anerkendt for en betydelig metodisk stringens i deres teoretiske arbejder.

and interrupted by time lags, outside interference and mistakes.... On the other side of the gulf are those that reject the idea that the existing economic system is, in any significant sense, self-adjusting. The strength of the self-adjusting school depends on it having behind it almost the whole body of organised economic thinking of the last hundred years...I think I am siding with the latter group, CWK XIII, (1934:485).

Metodologisk er der en afgørende forskel mellem de makroøkonomer, der a priori antager, at et markedsøkonomisk system med fleksibel pris- og løndannelse automatisk (omend med visse forsinkelser) af sig selv konvergerer mod en generel ligevægt, hvor alle markeder clearer. Denne antagelse legitimerer, at en generel ligevægtsmodel benyttes som benchmark-model, hvilket f.eks. var tilfældet med Velfærdskommissionens brug af DREAM-modellen.

Med inspiration fra Keynes' refleksioner over betydning af valget af makroøkonomisk metode, formulerer jeg i afhandlingen det spørgsmål: hvis denne a priori antagelse om generel ligevægt ikke afspejler den *makroøkonomiske* virkelighed (World 1), hvilken metodologi bør så bringes i anvendelse, således at de makroøkonomiske resultater bliver i (bedre) overensstemmelse med den makroøkonomiske virkelighed?

At Keynes vidste, hvad han talte om, er uomtvisteligt, idet han selv helt frem til og med udarbejdelsen af *A Treatise on Money* (udgivet i 1930) havde arbejdet inden for rammen af en generel ligevægtsmodel. Så han kendte dens analytiske svagheder – ikke mindst da der skulle rådgives om den økonomiske politik i 1930'erne. Et afgørende teoretisk problem var, at fænomenet *usikkerhed* i betydningen »we simply do not know« ikke øver indflydelse på den generelle ligevægtsløsning. Dette er den metodologiske udfordring, som Keynes og siden hen post-keynesiansk forskning har taget op, og lige netop derfor giver det mening på det metodologiske plan at sondre skarpt mellem resultater baseret på neoklassisk generel ligevægtsmodel og den makroøkonomiske forskning, der inddrager usikkerhed eksplicit.

Herved kommer med inspiration fra Keynes de metodologiske spørgsmål til at stå i centrum for den skole, der kan betegnes som *realistisk* makroøkonomisk forskning i modsætning til den idebaserede forskning.

Makroøkonomiske modeller

*Economics is a science of thinking in terms of models joined to **the art** of choosing models which are **relevant** to the contemporary world. It is compelled to be this, because, unlike the natural science, the material to which it is applied is, in too many respects, not homogeneous through time, (CWK, XIV, 1937:296/9, min fremhævelse).*

Denne anbefaling fra Keynes' hånd blev som bekendt overhørt eller i hvert fald ikke taget tilstrækkelig alvorligt af den række fremtrædende fortolkere af *The General Theory* bl.a. John Hicks, Lawrence Klein og Franco Modigliani, der kom til at stå som eksponenter for den »keynesianske revolution« i de første tiår efter krigens afslutning, og som efterfølgende meget betegnende er blevet kaldt den »neoklassiske syntese«. Her blev Keynes' efterspørgselsteori implanteret i den eksisterende generelle ligevægtsmodel. NK kan dog ikke ud fra en pragmatisk synsvinkel se noget problem i denne syntese og benytter den selv i sin disputats omhandlende udviklingen i dansk økonomi igennem små hundrede år. Og han kan heller ikke se, hvorfor Keynes burde have været mere kritisk over for Hicks' ISLM-model fra 1937. Men det kunne Hicks (1980) derimod; men først adskillige år efter Keynes' død. Hicks lægger således selv afstand til den neoklassiske fremstilling af *The General Theory* i en grad, så han ligefrem afstår fra at bruge modellen i sin forelæsningsrække fra 1974 *The Crisis in Keynesian Economics*. Derfor er det interessant at diskutere, i hvilken grad Keynes tilbage i slutningen af 1930'erne i fuldt omfang havde erkendt de metodologiske implikationer af at analysere makroøkonomi inden for rammen af et åbent system præget af usikkerhed i modsætning til en generel ligevægtsmodel. På det makroøkonometriske felt lod Keynes dog ikke Tinbergen i tvivl om, at sidstnævnte var gået alt for langt i sit forsøg på at forstå udviklingen i den hollandske økonomi i mellemkrigstiden ved brug af en til ISLM svarende model.

De store makroøkonometriske modelbyggere i den Hicks/Tinbergske tradition kunne med fordel have lyttet til Keynes' metodologiske advarsler, hvis de ville have undgået, at modellerne kuldsejlede en efter en, da efterkrigstidens stabile makroøkonomiske udvikling i 1970'erne blev afløst af en »uforklarlig« stagflation. Den »keynesianske revolution« baseret på den neoklassiske syntese blev her lagt i ruiner,² hvilket var en medvirkende årsag til, at også Keynes' og den senere post-keynesianske makroteori af mange yngre økonomer (og politikere) blev betragtet som et fejlslagent forskningsprogram.

Den »keynesianske revolution« blev i stedet afløst af en makroøkonomisk teori, der i eksplicit form benyttede sig af et neoklassisk mikroøkonomisk fundament samt af rationel forventningsdannelse (fuld forudseenhed), der blev holdt sammen på makro-niveau af en walrasiansk generel ligevægtsmodel. Den neoklassiske syntese blev således afløst af noget, der mest af alt mindede om en »keynesiansk antitese«.

Jeg dokumenterer dette udviklingsforløb i afhandlingen både internationalt gennem en række citater fra bl.a. Robert Lucas (1987) og i dansk sammenhæng ved brug af ci-

2. En på mange måder parallel udvikling er gengivet i Andersen (1991), hvor den tidlige modeludvikling i Danmark beskrives.

tater fra hhv. Torben M. Andersens oversigtsartikel med titlen »Makroteori« og Birch Sørensen og Whitta-Jacobsens lærebog, »Advanced Macroeconomics«, der benyttes som grundbog på bl.a. 2. år af polit-studiet. Det er derfor overraskende, at NK kalder disse citater taget direkte fra centrale main-stream kilder for »en karikeret stråmand«. Hvis disse oversigtsværker, der i øvrigt giver et udmærket grundlag til at forstå de i Danmark ofte benyttede makromodeller (DREAM, ADAM og SMEC), er karikerede stråmænd, så er det svært at forstå, at en sådan uenighed ikke kommer eksplicit frem ved præsentationen af main-stream makroøkonomi for 2.-3. års økonomistuderende og ældre kolleger, der ønsker at blive opdateret. Her har NK efter min mening et forklaringsproblem; for det er ikke tilstrækkeligt at henvise til f.eks. Joseph Stiglitz og George Akerlofs (2002) væsentlige og ganske avancerede bidrag til at forstå en række markedsmæssige imperfektioner i samfundsøkonomien, der beklageligvis ikke altid når frem til lærebøgerne. For disse fortjenstfulde partielle analyser (af bl.a. kreditmarkeder og brugtvoغنsmarkeder) bliver nemlig af main-stream økonomerne efterfølgende sat ind i en generel ligevægtsmodel, når deres makroøkonomiske betydning skal analyseres, hvilket Andersen (2000) med forbilledlig klarhed har redegjort for.

Dette indbyggede dilemma mellem på den ene side partielle og til tider ganske realistiske analyser af specifikke markedsdannelser og på den anden side de makroøkonomiske helhedsanalyser er uløst inden for main-stream økonomien. Her finder vi måske en del af baggrunden for, at NK på trods af sine mange kritiske bemærkninger til afhandlingen alligevel slutter med følgende anerkendende konklusion, »at man skal bruge meget forskellige modeller og metoder efter hvilke problemstillinger, der skal analyseres.« (s. 8). Det kunne have været et citat fra afhandlingen. Hvorefter NK fortsætter, »her stiller »mainstream«-økonomien med en kasse af redskaber«, hvilket lige netop er problemet, at det gør mainstream-økonomien jo ikke. Den stiller med et urealistisk mikroøkonomisk fundament og en generel ligevægtsmodel, der har vist sig ikke at være relevante, når ambitionen er en *realistisk* makroøkonomisk forskning. Den af mainstream-økonomerne anbefalede metode undertrykker væsentlige aspekter af virkeligheden. Generel ligevægt og rationelle forventninger må anses for at være en analytisk antitese til forståelse af et kontekst og path-dependent makroøkonomisk udviklingsforløb.

Afhandlingens hovedkonklusion vedrørende metodologiens betydning står således uantastet, at hvis ambitionen er at gennemføre en realistisk makroøkonomisk analyse, så må det kræves, at den metode, der benyttes, er »*relevant to the contemporary world*«. Det vil sige at åbenbare og virkelighedsnære forhold inddrages, som f.eks. at fremtiden er usikker, og at det markedsøkonomiske system ikke er selvregulerende.

Afsluttende om det post-keynesianske alternativ

Those [economists], who are strongly wedded to what I call »the classical theory«, will fluctuate, I expect, between a belief, that I am quite wrong and a belief that I am saying nothing new. It is for others to determine if either of these or the third alternative is right. (The General Theory, s.v)

NK er således mildt sagt skeptisk med hensyn til, at den post-keynesianske skole skulle repræsentere en dybere forståelse af en række makroøkonomiske fænomener. Som Keynes bemærkede i sit forord, så tilkommer det ikke forfatteren af afhandlingen at besvare dette spørgsmål. Jeg har lagt argumenterne frem, hvorfor jeg mener, at et mikroøkonomisk fundament baseret på repræsentative agenter, rationel forventningsdannelse og generelle ligevægtsmodeller sjældent om overhovedet kan være en velegnet analyseramme til at forstå den makroøkonomiske virkelighed.

Men åbenbart er det, at post-keynesianere (og kritiske realister) stiller sig skeptisk an over for, det synspunkt, at lange tidsmæssige forløb meningsfuldt skulle kunne analyseres inden for rammerne af en uforanderlig generel ligevægtsmodel. Det kan naturligvis ikke afvises, at der muligvis er der nogle underliggende strukturer i det dybe stratum, der kan forme et skelet for den makroøkonomiske udvikling i et 5 til 10 års perspektiv; men forståelsen heraf må opnås gennem en *trial and error process*, hvor også udformningen af den økonomiske politik (og måske navnlig forventningerne her til) vil øve en væsentlig indflydelse. Forhold der ifølge sagens natur må være delvist eksogene, delvist endogene og derfor gøre ethvert makroøkonomisk udviklingsforløb kontekst og path-dependent. Det vil til eksempel være rimeligt på baggrund af den viden vi allerede har at antage, at der (også) i fremtiden vil være en positiv korrelation mellem f.eks. privat forbrug og disponibel indkomst og mellem effektiv efterspørgsel og beskæftigelse. Det er robuste empiriske sammenhænge, der kan gives en teoretisk forklaring; men hvor den mere præcise angivelse af de størrelsesmæssige sammenhænge, der vil gælde i fremtiden naturligvis må være præget af en betydelig usikkerhed og vil variere på en uforudsigelig måde.

NK efterlyser konkrete eksempler på post-keynesianske analyser og politikanbefalinger, der ville have gjort en forskel – kan katten fange mus? Alene det forhold, at de ovenstående metodologiske overvejelser i sig selv leder frem til en mere ydmyg holdning til, hvilke spørgsmål, makroøkonomisk forskning overhovedet kan besvare, end der er tradition for blandt main-stream økonomer, gør mig tilbageholdende. På den anden side kan der måske alligevel være basis for at henvise til nogle få tilsyneladende robuste sammenhænge, som blev formuleret i Jespersen (1996). Heraf kan nævnes mulighederne for at opnå fuld beskæftigelse gennem en dybere forståelse af begrebet

»effektiv efterspørgsel« (jfr. også afhandlingens kapitel 8), påvisning af at ØMU-konstruktionen er baseret på ligevægtsteori, der kan medføre reduceret vækst i eurozonen, en robust sammenhæng mellem økonomisk vækst og energiforbrug.

I afhandlingens afsluttende kapitel har jeg derfor valgt at give en række anvisninger på centrale og aktuelle makroøkonomiske problemstillinger, der med fordel kan analyseres ved brug af en post-keynesiansk/kritisk realistisk makroøkonomisk metodologi.

Litteratur

- Akerlof, G. A. 2002. Behavioral Macroeconomics and the Macroeconomic Behavior (a Nobel lecture), *American Economic Review*, June, pp. 411-33.
- Andersen, E. 1991. *Modeludvikling i Danmark*, kapitel 3 i J. Jespersen, red. Model og Virkelighed, København: DJØFs Forlag.
- Andersen, T. M. 2000. Makroteori, i Chr. Hjorth-Andersen, red. *Udviklingslinjer i økonomisk teori*, København: DJØFs Forlag.
- Boland, L. A. 2003. *The Foundation of Economic Method: a Popperian perspective*, 2nd ed. London: Routledge.
- Dow, S. 1996. *The Methodology of Macroeconomic Thought – a Conceptual Analysis of Schools of Thought in Economics*, Cheltenham: Edward Elgar.
- Dow, S. 2002. *Economic Methodology: An Inquiry*, Oxford: Oxford University Press.
- J. Jespersen, 1996. *Økonomi og virkelighed*, København: Fremad.
- Jespersen, J. 2004. Kritisk Realisme – med makroøkonomiske eksempler, i L. Fuglsang og P. Bitsch Olsen, red., *Videnskabs-teori for Samfundsvidenskaberne*, Frederiksberg: Samfundsfaglitteratur.
- Jespersen, J. 2007. *Makroøkonomisk metodologi i et samfundsvidenskabeligt perspektiv*, København: DJØFs Forlag.
- Hicks, J. 1980/81. IS-LM: An Explanation, *Journal of Post Keynesian Economics*, Winter.
- Keynes, J. M. *The Collected Writings of Keynes*, (CWK), vol. i-xxx, edited by Donald Moggridge and published by Macmillan & Cambridge University Press for The Royal Economic Society in the years 1972-1989.
- Keynes, J. M. 1936. *The General Theory of Employment, Interest and Money*, CWK, vol. VII.
- Lawson, T. 1997. *Economics and Reality*, London: Routledge.
- Lawson, T. 2003. *Reorienting economics*, London: Routledge.
- Lawson, T. 2002. *Social explanation and Popper*, discussion paper/lecture given in September 2002, University of Galway, kan down-loades fra Tony Lawsons hjemmeside: www.econ.cam.ac.uk/faculty/lawson/
- Lucas, R. E. 1987. *Models of Business cycles*, Oxford: Basil Blackwells.
- Popper, K. R. 1999. *All Life is Problem Solving*, London: Routledge.
- Sørensen, P. B. og H. J. Whitta-Jacobsen. 2005. *Introducing Advanced Macroeconomics*, New York: McGraw Hill Education.

Bog anmeldelser

Anders Christian Hansen: *Uden for hovedstrømmen – Alternative strømninger i økonomisk teori*. Anmeldt af Nicolai J. Foss, Copenhagen Business School.

Krisenak før og nu

For tre til fire årtier siden var der i fin overensstemmelse med tidsånden megen tale om økonomisk teoris krise. En række alternative, »heterodoxe«, »non-mainstream« retninger blev etableret eller taget frem af gemmerne og støvet af som alternativer til den formodet monolitiske, neoklassiske hovedstrøm. Det drejer sig primært om evolutionær økonomi, østrigsk økonomi, post-Keynesianisme, neo-ricardianisme og den »gamle« institutionalisme (samt kombinationer af disse). Alternative »societies« og ganske mange nye tidsskrifter, hvoraf de fleste stadig eksisterer (f.eks. *Journal of Post-Keynesian Economics* og *Review of Austrian Economics*), blev startet. Flere af de etablerede, men måske lidt falmende, stjerner, som Joan Robison (1971), Nicholas Kaldor og Oskar Morgenstern (1972), meldte sig som markante kritikere af den neoklassiske hovedstrøm. Samtidig profilerede en række retninger, der reelt var tæt på hovedstrømmen, sig som kritikere af denne. Det drejer sig for eksempel om »new institutional economics«, herunder teori om ejendomsrettigheder, Coase (1988), samt for eksempel public choice-teori, Buchanan (1978).

Med neo-ricardianerne som undtagelse var kritikerne forenede i en argumentation om, at den dominerende generel ligevægts-teori med dens fundamentalt urealistiske antagelser blokerede for forståelsen af en lang række vigtige fænomener som uligevægt, fundamental usikkerhed, andre institutioner end markedsinstitutionen, entreprenørskab, læring, etc., og at en forståelse af disse fænomener fordrede et brud med den walrasianske model. Særligt i Europa bed hovedstrømmens kritikere sig fast på en række universiteter. For eksempel var de Sraffa-inspirerede neo-ricar-

dianere en overgang meget indflydelsesrige på Cambridge University. Evolutionære økonomer var længe blandt de mest citerede italienske og tyske økonomer. I Danmark spillede kritikerne en rolle på universiteterne i Roskilde og Ålborg.

Lyttede hovedstrømmen så til kritikere? Næppe meget, ihvertfald ikke til de mere radikale kritikere. Mange af kritikerne marginaliseredes da også efterhånden, ofte meget direkte i form af de ikke fik tenure på de bedre universiteter eller måtte søge tilflugt på de mere imødekommende business schools. Selvom økonomer som Paul Davidson, Israel Kirzner, Geoff Hodgson, og Sidney Winter måske er stjerner inden for deres respektive sub-kulturer, er de relativt ukendte blandt flertallet af universitetsøkonomer.

Den fleksible hovedstrøm

Nu er det ikke her budskabet at alle gode ideer uproblematisk opsuges i hovedstrømmen, at der nødvendigvis er en perfekt selektion for teorier selv over lang tid, og at det er meningsløst at være kritisk over for hovedstrømmen. Mange af de kritikpunkter, der rejstes af for eksempel østrigske og evolutionære økonomer, havde væsentlig berettigelse, og nogle kritikpunkter er stadig valide (f.eks. den fortsatte negligeren af entreprenørens rolle i økonomisk udvikling). Men der er nok tale om, at kritikere af hovedstrømmen undervurderede, hvor tilpasningsdygtig og fleksibel denne i virkeligheden var og er. Den afgørende tilpasning var gennemslaget for spil- og informationsteoretisk analysemetodik fra 1980ernes begyndelse i efterhånden stort set alle økonomiske discipliner. Hermed blev den del af kritikken, der gik på atomisme, rigide ligevægtsbegreber, og ekstreme antagelser om information og usikkerhed, i høj grad demonteret, idet denne i meget høj grad havde rettet sig mod generel ligevægtsmodellen. At den såkaldte hovedstrøm gennem det seneste tiår i stigende grad har eksperimenteret med variable præferencer, altruis-

me, »intrinsic« motivation, forskellige typer af kognitive biases, etc. etc. og har undersøgt aggregerede konsekvenser af sådanne antagelser, blandt andet ved hjælp af en strøm af nye ligevægtsbegreber, har yderligere været med til at tage luften ud af kritikken af hovedstrømmen for manglende realisme i antagelserne. Økonomer har i stigende taget (formodet) robust evidence fra nabo-disciplinerne alvorligt, særligt fra psykologi/cognitive science, antropologi og sociologi (omtrent i denne rækkefølge).¹ Eksempler på det er Benabou og Tirole (2003) forsøg på at integrere psykologiske forskningsresultater om »intrinsic« motivation i principal-agent-teori, Akerlof og Kranton, f.eks. (2005) forsøg på at indkorporere sociologiske ideer om identitet, og ikke mindst de senere års fascination af neuroeconomics, f.eks. Camerer, Loewenstein & Prelec (2005).

Nogle vil måske mene, at den ret så strenge adskillelse mellem disciplinerne, som håndhævedes for bare få årtier siden, brydes ned i takt med, økonomer har opdaget, at der egentlig ikke er dét, der ikke kan formaliseres. Det er det nok for tidligt at sige noget mere definitivt om. Men økonomers nærmest opportunistiske beredvillighed til at formalisere, hvad formaliseres kan, betyder, at det begynder at give mening at betvivle, at der er nogle bestemte, uforandelige grundantagelser, der kendetegner »hovedstrømmen.« Treenigheden ligevægt-maksimering-stabile præferencer gør det næppe længere, ihvertfald hvis den seneste avantgardistiske forskning i førende tidsskrifter er pejlemærket. Hvor stort set al vigtig forskning i økonomisk teori i næsten hele 1900-tallet var baseret på denne treenighed, er det i et historisk perspektiv noget af et radikalt skifte, at økonomer nu synes langt mere åbne for mere realistiske op-

fattelser af agenters motivation og kognition, i højere grad modellerer processer uden for ligevægt, og interesserer sig for ikke-stationære nyttefunktioner, Bowles (2004). Det er bemærkelsesværdigt, at dette ikke synes drevet (ihvertfald ikke direkte) af den heterodokse kritik af hovedstrømmen, men snarere af opmærksomhed over for fundamentale indsigter i nabo-disciplinerne.

Man kan så indvende, at det måske er mindre relevant for en karakteristik af hovedstrømmen, hvad avantgardister som Matthew Rabin, Colin Camerer, eller Samuel Bowles beskæftiger sig med. Selvom deres forskning er vigtig, er det *eksplorativ* forskning, der af søger hovedstrømmens grænser, og som kun muligvis vil påvirke hovedstrømmen afgørende på længere sigt. Det er ikke forskning, der *definerer* hovedstrømmen, som snarere er karakteriseret ved hvad de resterende 95% mindre innovative forskere bedriver. Nelson & Winter (1982) skelner mellem »working paper« hovedstrøm og »textbook« hovedstrøm, hvor førstnævnte repræsenterer de nye innovative bidrag, som kun efter diskussion, udgivelse i tidsskrifter, endnu mere diskussion, etc. måske diffunderer ud i lærebøger, og med et betydeligt lag bliver den nye »lærebogs-hovedstrøm.« Sondringen synes at have ganske megen mening. Der er en væsentlig afstand mellem grundantagelserne i en gængs mikro-bog som Mas-Colell, Whinston & Green (1995) på den ene side og forskningsfronten i for eksempel eksperimentel økonomi. Men pointen er, at hovedstrømmen bevæger sig fremad trukket af den eksplorative forskning, der ikke mindst i disse år ret radikalt skubber til grænsen for hovedstrømmen. Det betyder, at man skal være varsom med sin anvendelse af hovedstrøms-etiketten og særligt med konklusioner der går på, at hovedstrømmen er fastlåst i bestemte antagelser der iboende begrænser den.

Udenfor hovedstrømmen

»Uden for hovedstrømmen – Alternative strømninger i økonomisk teori« består af ni kapitler, alle forfattede af forskere med til-

1. Og ofte for alvorligt. For eksempel accepterer Benabou & Tirole (2003) uden videre diskussion psykologiske ideer om intrinsic motivation, selvom deres status er ret uafklaret, ligesom Akerlof & Kranton (2005) empiriske basis for deres sociologisk inspirerede argumentation er særdeles »casual«.

knytning til Roskilde Universitet. Ambitionen er, som redaktøren skriver i sit forord (s. 11), at give en »indføring i nogle af de økonomiske skoler og retninger, som ligger i udkanten af og i mange tilfælde i opposition til hovedstrømmens standardmodeller.« Vi kommer vidt omkring, fra Jesper Jespersens kapitel om »post-keyensiansk« (sic!) makroteori, over fransk konventionsøkonomi (ved Søren Jagd) og institutionel økonomi (Martin Enevoldsen) til mere mikro-orienterede temaer som transaktionsomkostningsteori (Carsten A. Koch) og principal-agent-teori (Kirsten Bregm). Anders Christian Hansen, bogens redaktør, har bidraget med et kapitel om »økologisk økonomi.« Anders Hede skriver om »Fælles tankegods i økonomi og biologi,« og Hans Aage runder af med et kapitel om »Markedsideologi i hovedstrømmen.« Flere af kapitlerne giver fine og tankevækkende oversigter over virkelige eller postulerede alternativer til hovedstrømmen. Særligt Enevoldsen, Koch, Hede og Bregms kapitler er vellykkede introduktioner til deres respektive områder. Det skyldes blandt andet, at de er fri for den type overdreven polemik som alt for ofte har karakteriseret de heterodokse traditioner i økonomi. I det hele taget er tonen ganske urban i forhold til den hovedstrøm, man mener sig »udenfor.« Kun Aages kapitel er karakteriseret ved misvisende og unødvendig polemik (herom senere).

Man kan mene at bogens titel tildels er falsk varebetegnelse. Det er der flere årsager til. For det første er flere af bogens bidrag faktisk slet ikke »uden for hovedstrømmen« i nogen meningsfuld forstand. Eksempelvis gennemgår Kirsten Bregm i et kompetent og meget dækkende kapitel principal-agent-teorien, men denne teori er absolut ikke uden for hovedstrømmen og har aldrig været det; tværtimod er teorien helt i kernen af hovedstrømmen. Transaktionsomkostningsteorien i Williamsons udformning ligger vel også ret godt placeret i hovedstrømmen, hvilket selvfølgelig også gælder de i bogen mange gange omtalte ejendomsrets-argumenter (Coase, Demsetz, etc.). Martin Enevoldsens kapitel

om institutionel økonomi forsøger ganske vist (og ret så problematisk) at trække en linie fra den »gamle« og meget anti-mainstream institutionelle økonomi (Veblen, Commons, Ayres, etc.) til den helt overvejende mainstream »new institutional economics.« men refererer helt overvejende de grundlæggende bidrag til sidstnævnte (altså Coase, Williamson og North). Det er korrekt, at særligt Coase og North ofte har kritiseret »neoclassical economics,« se f.eks. Coases nobelprisforelæsning; Coase (1992), men det er ret klart, at målet for deres polemik er snævrere, nemlig traditionel generel ligevægtsteori à la Debreu. Anders Chr. Hansens forsøg på at konstruere fælles tematikker for det noget uklare felt »økologisk økonomi« benytter sig af meget mainstream miljøøkonomiske argumenter. Aages kritik af en påstået markedsideologi i hovedstrømmen benytter sig ligeledes af mainstream argumenter.

En anden årsag til at »uden for hovedstrømmen« er en lidt problematisk titel har at gøre med, at med en sådan titel kunne man forvente et nogenlunde komplet signallement af de økonomiske traditioner, der vitterligt har erklæret sig i opposition til hovedstrømmen. Det er ikke tilfældet: der er kun en ganske kortfattet diskussion af de »gamle institutionalister« og ingen diskussion af østrigerne eller de Schumpeter-inspirerede evolutionære økonomer i Nelson & Winter (1982)-traditionen. Den sidstnævnte udeladelse er måske den mest kritiske, for disse økonomer er nok dem, der mest succesfuldt har udfordret hovedstrømmen, blandt andet fordi de i modsætning til for eksempel østrigere og gamle institutionalister ikke har haft berøringsangst over for formelle metoder. Man kan også undre sig over, at bogen intet indeholder om eksperimentel økonomi, og at de mange forsøg på at inddrage begrænset rationalitet, neuro-economics, nuanceringer af hvad der antages om agents præferencer (f.eks. »sociale præferencer«), etc. kun er antydningvist diskuteret (nemlig i Anders Hedes udmærkede kapitel), når det nu er ambitionen at portrættere også forskning, der ligger »i ud-

kanten af ... i hovedstrømmens standardmodeller.« I denne »udkant« ligger ihvertfald de såkaldte »post-keynesianere«, som gennemgås af Jesper Jespersen: de er ikke mindst kendetegnet ved temperamentsfulde opgør med det konventionelle arvegods i makroøkonomien som IS-LM-modellen, og i det hele taget med forsøg på at udtrykke Keyneske indsigter ved hjælp af modificerede standardmodeller. Der er tale om en meget ortodoks Keynesianisme, hvis hellige skrifter er kapitel 12 (»The State of Long Term Expectations«) i *General Theory* og Keynes' 1937-artikel »The General Theory of Employment«, med deres understregning af fundamental usikkerhed, animal spirits, etc. Jespersen giver sin diskussion en metodologisk drejning ved at koble post-keynesianisme til den videnskabs-teoretiske retning »kritisk realisme.« Hans primære formål er at argumentere for nødvendigheden af realisme i de gjorte antagelser, samt at kontrastere post-keynesianere med hovedstrømmen ud fra hvor megen opmærksomhed, de relevante forklarende mekanismer vies. Her ryger kæden desværre af for Jespersens argumentation. For det første er diskussionen om antagelsers realisme nemlig helt generel og i øvrigt en ældgammel travet i økonomisk metodologi og ikke specifikt forbundet med kritisk realisme. For det andet er det misvisende, når Jespersen antyder, at post-keynesianisme bevæbnet med kritisk realisme adskiller sig fra anden makroøkonomi ved sin interesse for de generative mekanismer, der producerer makroøkonomiske fænomener. Stort set al makro-økonomisk debat, ja vel størstedelen af økonomisk diskussion overhovedet, handler om karakteren af sådanne mekanismer, og hvordan de bedst fanges i teorien. Der er en meget stor relevant litteratur i såvel videnskabsteori som økonomisk metodologi, der med fordel kunne have været inddraget, se f.eks. Mäki (2001).

Hvor Jespersens kapitel er mildt og implicit polemiserende, kammer Aages kapitel helt over i angreb på utallige fjender, navngivne som unavngivne. Baseret på en misvisende negativ definition af ideologi som »det blan-

dingsprodukt, som opstår, når man opgiver at sondre mellem, hvad man ved, hvad man formoder, og hvad man håber og ønsker« (s. 238), får miljøøkonomi, sundhedsøkonomi, kulturøkonomi, liberalisme, Buchanan, Hayek, Milgrom & Roberts, Møller & Nielsen, etc. etc. så hatten passer. Alle gør de sig skyldige i utidig ideologiseren. Aage derimod forklarer, at »Som Marx så rigtigt siger«: »Kapitalen driver af blod og snavs, fra top til tå, ud af alle porer« (s. 251) og er således ikke det mindste ideologisk, men strengt videnskabelig og objektiv. Lige så videnskabelig er den meget dybe påpejning af eksistensen af en »ideologisk bølgebevægelse« (s. 251), således at public choice-teori, privatisering, og andre vederstyggeligheder blot er en slags reinkarnationer af »den stærkt liberalistiske og ahistoriske Manchester-skole« (s. 251). Som det vel er antydnet, er kapitlet så rablende, at det burde have været udeladt.

Afrunding

Hovedstrømmens kritikere havde det relativt let for tre-fire årtier siden. Dengang kunne man ret entydigt karakterisere hovedstrømmen som optaget af ressourceallokering på et givent tidspunkt og baseret på et fokus på marginale tradeoffs, meget håndfaste rationalitetsantagelser, metodologisk individualisme, en version af utilitarisme og en grundlæggende generel ligevægt-opfattelse af økonomien. Denne version af hovedstrømmen (neoklassisk økonomi?) er nu kun uantastet i lærebøgernes univers. Den avancerede forskningslitteratur udfordrer derimod alle de nævnte kendetegn (måske undtaget metodologisk individualisme). Det er givetvis at gå for vidt, når Colander (2000) på den baggrund ligefrem proklamerer »the death of neoclassical economics«: I langt den meste publicerede forskning er der stadig tale om rationalitetsantagelser, der ikke bryder afgørende med forventet nytte-modellen, der gøres stadig vidtgående og urealistiske antagelser om agents, processer uden for ligevægt vies ringe opmærksomhed, etc. Men der er tale om den meget markante holdningsændring, at de

nævnte »neoklassiske« antagelser ikke er sakrosankte, men i meget højere grad åbne for debat, ligesom konstruktivt arbejde baseret på alternative antagelser har en langt bedre chan-

ce for publicering i de bedste tidsskrifter. Og det burde være gode nyheder for de økonomer, der mener sig »uden for hovedstrømmen.«

Litteratur

- Akerlof, G. A. & R. E. Kranton. 2005. Identity and the Economics of Organizations, *Journal of Economic Perspectives*.
- Benabou, R. & J. Tirole. 2003. Intrinsic and Extrinsic Motivation, *Review of Economic Studies* 70: 483-520.
- Bowles, S. 2004. *Microeconomics*. New York: Russell Sage Foundation.
- Buchanan, J. 1978. *What Should Economists Do?* Indianapolis: Liberty Press.
- Camerer, C., G. Loewenstein, & Drazen Prelec. 2005. Neuroeconomics: How Neuroscience Can Inform Economics, *Journal of Economic Literature* 43: 9-64.
- Coase, R. H. 1988. *The Firm, the Market and the Law*. Chicago: University of Chicago Press.
- Coase, R. H. 1992. The Institutional Structure of Production, *American Economic Review*, 82: no 4, pp 713-19.
- Colander, D. 2000. The Death of Neoclassical Economics, *Journal of the History of Economic Thought* 22: 127-43.
- Mas-Colell, A., M. D. Whinston, & J. R. Green. 1995. *Microeconomic Theory*. Oxford: Oxford University Press.
- Morgenstern, O. 1972. Thirteen Critical Points in Contemporary Economic Theory: An Interpretation, *Journal of Economic Literature* 10: 1163-89.
- Mäki, U. 2001. *The Economic Worldview: Studies in the Ontology of Economics*. Cambridge: Cambridge University Press.
- Nelson, R. R. & S. G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: The Belknap Press.
- Robinson, J. 1971. *Economic Heresies*. London: Basic Books.

Niels Jørgen Mau Pedersen: *Den offentlige sektor i flere niveauer*. Jurist- og Økonomforbundets Forlag, 2007. Anmeldt af forskningsleder, ekstern lektor, cand.polit. Henrik Christoffersen.

Kommuner er som institutionsform en af de grundlæggende samfundsinstitutioner i den danske samfundsmodel. Fastslået i grundloven og med en lang historisk tradition bag sig. Men kommuneinstitutionen er også en institution i forandring. Ikke blot derved, at kommunesektorens udgifter og beskæftigelse vokser med betragtelig hast. Selve kommunernes rationale og vores måde at tænke på, hvorledes kommunerne på en optimal måde kan bidrage til velfærd i samfundet, ændrer sig grundlæggende i disse år. Her giver læs-

ning af Niels Jørgen Mau Pedersens bog om »Den offentlige sektor i flere niveauer« stof til eftertanke. I særdeleshed vil de læsere, som også kender til forfatterens tolv år gamle bog om samme emne, få understreget, hvor grundlæggende anderledes vi er kommet til at tænke omkring håndteringen af økonomien i forbindelse med kommunestyret. Og de, som er bekendt med udviklingen i kommunesektoren, vil få anledning til at fundere over, om vi ikke igen er på vej et ganske andet sted hen. Niels Jørgen Mau Pedersen har i en længere årrække både været leder af det kommunal-økonomiske kontor i Indenrigs- og sundhedsministeriet, nu Velfærdsministeriet, og ekstern lektor ved Økonomisk Institut på Københavns Universitet, hvor han har bedrevet undervisningen i kommunal- og regionaløkonomi. Både den nu fremkomne bog og

bogen fra 1995 med titlen »Decentralisering og kommunaløkonomi« er skrevet med henblik på universitetsundervisningen.

»Den offentlige sektor i flere niveauer« afspejler på en balanceret måde forfatterens dobbelte udgangspunkt i økonomisk viden- skab og i håndværket at administrere hånd- teringen af den danske kommunaløkonomi. Der er, som Pedersen selv beskriver det, tale om en primært teoretisk fremstilling af for- skellige modeller og tankegange, men der er også en del institutionelt stof og mere beskri- vende afsnit. Pedersen skriver godt og klart og med et ligefremt sprog, som fortæller læ- seren om Pedersens på samme tid engagerede og rolige temperament, og som gør teksten vedkommende for den kommunaløkonomisk interesserede. Teknisk særligt krævende pas- sager er angivet med petit og kan eventuelt i første omgang undgås af læsere uden det for- udsatte minimum af matematiske forudsæt- ninger. Enkelte detaljer kunne have været pudset lidt bedre. Hverken Madsen (1989) på side 47 eller Petersen (1981) på side 56 er ek- sempelvis til at finde i den ellers meget fyl- dige og kvalificerede litteraturoversigt. Det skal siges med det samme, at bogen er uomgæn- gelig i den forstand, at den giver en særdeles kompetent indføring i det formelle system, som udgør rammerne omkring kommunernes økonomi: Udligning, udgiftsbehov, tilskud og overførsler samt lokale skatter. I tilknytning hertil giver bogen en solid indføring i det øko- nomiske tankegods, som har ligget bag ord- ningen af det bestående system. Det bestæn- de system er jo imidlertid i oprud. Hvad det er for økonomiske problemer, som ligger bag dette oprud, og hvad det er for økonomiske forståelsesmåder, som kendetegner nydannel- serne i det kommunale økonomiske system, går bogen derimod i hovedsagen uden om. Vi får kun lige præsenteret, at vi ikke fuldt kan forstå, hvad der sker i den kommunale verden ved i teoretisk henseende alene at forfølge det traditionelle fiscal federalism-spor.

I forhold til 1995-bogen gøres dette der- ved, at tilgangen i Pedersens nye bog er ble- vet udbygget. Det som særligt falder i øjnene er, at Pedersen nu lægger langt større vægt på

det overordnede spørgsmål om den offentlige sektors organisering. Her forfølger Pedersen nu to spor. For det første er bogen på samme måde som dens forgænger en grundbog i disci- plinen fiscal federalism eller public finance, som i bogens forståelse stort set er sammen- faldende discipliner. Centralt står her en nor- mativ undersøgelse af, hvorledes en given of- fentlig opgaveportefølje bør fordeles på den centrale myndighed og på decentrale myndig- heder, og hvorledes de finansielle relationer mellem central myndighed og decentrale myndigheder bør indrettes. For det andet pe- ger bogen også på, at kræfter, som virker i den politiske økonomi, er afgørende for, hvorledes den offentlige virksomhed vil ud- vikle sig inden for en given form for organisering af den offentlige sektor. Her er det den i hovedsagen positive disciplin, public choice, som bringes ind i fremstillingen.

Placeringen af public choice-tilgangen i princippet sideordnet med fiscal federalism- traditionen repræsenterer en fornyelse af fremstillingen, når man sammenholder »Den offentlige sektor i flere niveauer« med dens forgænger. Efter kommunalreformen i 1970 og med den moderne velfærdsøkonomis ind- tog herskede en stærk tro på Den gode Vel- færdssektor, hvor en intelligent indretning af velfærdssektorens organisation i sig selv bor- gede for effektivitet. I dag vil vi i højere grad have syn for sagn. Et er, at velfærdssektoren i princippet kan præstere det bedst opnåelige. Noget andet er, om den rent faktisk drives af kræfter, som også gør det.

Selv om det i høj grad er med til at gøre »Den offentlige sektor i flere niveauer« til en væsentlig bog, at den eksponerer en ny for- ståelsesmåde, som har trængt sig på i forhold til kommunesystemet, så er det også i forhold til skredet i forståelsen af kommunesystemets økonomiske problematik, at vi finder bogens begrænsning. Når vi (på en overskuelig og kompetent måde) får gennemgået Niskanens legendariske model af den budgetmaksime- rende bureaukrat, er det eksempelvis med læsevejledningen, at her er der tale om en provokerende model. Måske er det også Niels Jørgen Mau Pedersens baggrund som en cen-

tralt placeret embedsmand, der får ham til at lægge bånd på sig selv. Det kommer i det hele taget til udtryk derved, at bogen nok påpeger den politiske økonomis relevans, men at bogen på den anden side er meget tilbageholdende i forhold til at gå ind i de nye former for organisering og økonomiske mekanismer, som udvikles i takt med den tiltagende opmærksomhed omkring kommunernes politiske økonomi.

I bogens behandling af public choice-tilgangen er det en central problematik, hvorvidt den offentlige sektors organiseringsmåde i sig selv kan bidrage til at forklare den offentlige sektors størrelse. I forlængelse af Niskanen-tænkningen er det med et grundlæggende værk fra 1980 af Brennan og Buchanan en nøgle-tese i public choice-litteraturen, at den offentlige sektor vil tendere at rumme en Leviathan-mekanisme. Politikere og bureaukrater vil i forbindelse med udbuddet af offentlige ydelser udvise monopoladfærd og søge at sikre et overskud til sig selv, som finansieres af skatterne, og som ikke afleder nytte til borgerne og vælgerne.

Her leverer Niels Jørgen Mau Pedersen en særdeles interessant og også i høj grad original analyse af de empiriske resultater i den internationale litteratur om denne problematik. En meta-analyse udtrykt med moderne jargon. I alt inddrager han 15 studier startende med Oates' grundlæggende 1972-studie i bogen, som lagde navn til disciplinen »Fiscal federalism«. Pedersens indsigt er, at de empiriske resultater alt i alt er blandede. Han finder nok en tendens til, at de seneste og gennemgående mest raffinerede studier i højere grad end de tidligere og mere enkle landesammenligninger finder belæg for, at decentralisering til kommuner binder Leviathan og tenderer at begrænse den offentlige sektors størrelse. Der er imidlertid den afgørende tilføjelse til dette resultat, at en sådan binding overvejende ser ud til at gælde, såfremt der også sker en decentralisering af skatteudskrivningen. Pedersen konkluderer sammenfattende: »Skal decentralisering således være en binding på Leviathan synes det væsentligt, at decentraliseringen indeholder en reel de-

centralisering af skattefastsættelsen. Hvis der derimod i højere grad baseres på centrale, statslige overførsler til de decentrale niveauer kan resultatet meget vel blive omvendt – en større offentlig sektor.« Dette er unægtelig en tankevækkende analyse i og med, at vi i Danmark netop har gennemført en kommunalreform, hvor kommunerne på den ene side er blevet yderligere tilført opgaver og hvor der på den anden side tegner sig konturerne til en fremtidig økonomisk styring af kommunerne, hvor den historiske frie kommunale skatteudskrivningsret i det mindste i praksis forsvinder, og hvor den kommunale udligning er blevet styrket. Kommunerne har i den henseende fået større magt og mindre ansvar, og det kan i lyset af Pedersens analyse forekomme at være problematisk.

Mau Pedersens eget bidrag til forståelse af problematikken omkring kommunernes indpasning i stabiliseringspolitikken på makro-niveau tager udgangspunkt i Musgraves specificering af de tre grene af public finance, hvor den stabiliseringspolitiske gren placeres som en opgave for den centrale politik med den begrundelse, at stabiliseringsopgaven oplagt har tydelige eksternalitetsvirkninger, hvis den udføres på lokalt plan. Nu er det meget tidligt at drage konklusioner vedrørende de danske kommuners indplacering i en samlet efterspørgselsstyring efter kommunalreformen, men de første indikationer er unægtelig frygtindgydende. Den aftaleinstitution, der er udviklet siden begyndelsen af 1970'erne, har gennemgående kun virket delvist. Den aftale, som blev indgået mellem Kommunernes Landsforening og regeringen for det første år efter at den nye kommunalreform er blevet indkørt, aftalen for 2008, har imidlertid vist sig ikke at virke overhovedet, selv om den endda gav plads til en pæn realvækst i den kommunale sektor. Det ligger helt i forlængelse af Mau Pedersens konklusion. Man kunne derfor godt have ønsket sig og forventet, at Mau Pedersen havde udviklet sin analyse yderligere, men bogen går ikke nærmere ind i den aftalemekanisme, som skal indpasse de danske kommuner i makropolitikken.

Den klassiske fiscal federalism vedrører

som udgangspunkt først og fremmest organiseringen af forsyningen med lokale offentlige goder. »Den offentlige sektor i flere niveauer« betoner i overensstemmelse hermed også teorien om produktion af offentlige goder. En konsekvens af sådan en tilgang er, at klubteori kommer til at spille en betydende rolle. Her udgør produktionsmæssige overvejelser et udgangspunkt for undersøgelse af implikationer af centralisering versus decentralisering. Den produktionsfunktion, som her umiddelbart giver sig, er en funktion hvor antal medlemmer af en klub bestemmer klubbens produktion. Eventuelt kan der i produktionsfunktionen dog indgå en fast produktionsfaktor, som f.eks. udtrykker anvendelse af jord. Sådanne forestillinger om teknologi i den kommunale produktion er nærliggende ensbetydende med et regime, hvor produktivitet er en given størrelse bestemt af de gældende institutionelle omstændigheder. I nutidig talemåde forholder det sig sådan, at kun »varme hænder« tæller. Sådant en forestilling om teknologi var da også enerådende ovenpå kommunalreformen i 1970. Det kom eksempelvis til udtryk i den kommunale budget- og regnskabsreform, der blev implementeret i 1976, og som lagde grunden til den økonomitænkning, som stadig har stor gyldighed i dag. Informationssystemet blev her organiseret med henblik på primært at understøtte en forestilling om, at den overordnede ledelsesopgave i en kommune består i at løse prioriteringsproblemet. Derimod gav informationssystemet praktisk talt ingen mulighed for at belyse produktivitet. Eller udtrykt på en anden måde. Det har været gængs latin i dansk kommunaløkonomi, at output er identisk med input.

I dag er det imidlertid en plausibel betragtning, at selve projektet om modernisering af kommunesektoren mere end noget andet handler om at bryde med den traditionelle forestilling om teknologien i kommunernes produktion. Der sættes eksempelvis voldsomt på at intensivere indsatsen af uddannelseskompetencer i den kommunale arbejdsstyrke, og bag sådan en satsning må vel ligge en forestilling om en produktionsfunktion, hvor human kapital som ny variabel produktionsfaktor

kan bidrage til at hæve arbejdsproduktiviteten. Ligeledes foregår organisatoriske udviklingsprocesser, hvor kommunernes produktion samles i stadig større produktionsenheder med henblik på at muliggøre en udbygning af specialiseringen i produktionen. I denne forbindelse er det bemærkelsesværdigt, at den statslige styring af kommunerne i stigende omfang kommer til at forholde sig direkte til kommunernes teknologianvendelse. Den nye seminarielov med specialisering af læreruddannelsen er et konkret eksempel. Trepartsforhandlingerne med en gennemgående betoning af uddannelsesniveaet hos servicesektorens medarbejdere peger mere i almindelighed herpå. Det ville også have været interessant, om »Den offentlige sektor i flere niveauer« åbnede for dette væsentlige perspektiv.

Mau Pedersen supplerer den traditionelle fiscal federalisms fokusering på opgaven at sikre tilvejebringelse i tilstrækkeligt omfang af lokale offentlige goder med en analyse af problematikken, at kommunerne også varetager produktion af, hvad bogen betegner semi-offentlige goder. Altså goder, som i det mindste er kendetegnet ved, at der gør sig eksklusion i forbruget gældende. Det er som om disse semi-offentlige goder udgør et særtilfælde. Efter gennemgangen af ordinære offentlige goder og offentlige goder i ikke-håndgribelig form som f.eks. målsætninger for lokalsamfundet anfører Mau Pedersen: »Endelig er der for det tredje en række goder, som indtager en mellemstilling mellem offentlige og private goder. Sådanne semi-offentlige goder«. Men virkeligheden er jo, som det i skandinavisk sammenhæng først og fremmest er den norske kommunaløkonom, Jørn Rattsø, der har gjort opmærksom på, at kommunerne netop altovervejende beskæftiger sig med produktion af, hvad han betegner private servicegoder, som ønskes underlagt omfordeling.

I og med, at de danske kommuner i realiteten har udviklet sig til at være mastodontkoncerner for produktion af privat service bliver det stedse mere relevant at overveje industriøkonomiens betragtningssmåde i forhold til kommunesystemet. I virkelighedens verden

holder tankegang fra markedsøkonomien da også tydeligt indtog. Herunder indgår det med vægt i regeringens politik for modernisering af kommunesektoren at udvirke, at markeds kræfter kommer til at have stærkere kraft inden for sektoren. Konkurrence og forbrugermagt spiller altså en rolle og er udset til at komme til at spille en stadig stærkere rolle. I »Den offentlige sektor i flere niveauer« findes ordet udlicitering refereret fire gange i stikordsfortegnelser, og i alle tilfælde er der tale om ganske korte notationer.

Den økonomiske teori, som udfoldes i »Den offentlige sektor i flere niveauer«, forholder sig altså sammenfattende som udgangspunkt til problemet at etablere en offentlig organisering af produktionen af lokale offentlige goder ved hjælp af en teknologi karakteriseret ved arbejdskraft som afgørende produktionsfaktor i kommuner, der agerer som fuldt ansvarlige samfundsinstitutioner. Men læseren af bogen kommer uvægerligt til at tænke på, at kommunerne i dag i hovedsagen forestår produktionen af private goder med en teknologi, som gerne skulle indoptage

stadig mere intensiv anvendelse af human kapital og kapital i nye former herunder ikke mindst i form af informationsteknologi, at kommunerne af regeringen i stadig stigende grad presses til markedsførelse, samt at kommunerne endelig i stadig mindre grad fremtræder som selvstændigt ansvarlige for egne økonomiske dispositioner.

Kommunerne i den danske samfundsmode befunder sig i et vadested. Den tænkning og samfundsteori, der blev gjort gældende, da et nyt kommunesystem blev skabt efter 1970, afspejles i dag stadig i høj grad i den måde kommunesystemet forstås på og håndteres på. Derfor er denne tænkning i dag stadig væsentlig at få trukket op, som det bliver gjort i »Den offentlige sektor i flere niveauer«. Men vi må også se i øjnene, at kommunesystemet er på vej et andet sted hen. Det er derfor ligeledes en særdeles væsentlig udfordring for den økonomiske videnskab at få etableret et grundlag for at forstå kommunesystemets aktuelle udfordringer. Det har vi til gode. Også efter Niels Jørgen Mau Pedersens særdeles solide og nyttige bog.