

■ ■ ■ JOHN K. DAGSVIK:

James Heckman og Daniel McFadden: To pionerer i utviklingen av mikroøkonometri

Begrepet mikroøkonometri benyttes ofte til å betegne kombinasjonen av mikroøkonomisk modellering, tilhørende metoder for statistisk inferens, samt empiriske anvendelser basert på mikrodata. James Heckman og Daniel McFadden har hatt helt sentrale roller i utviklingen av mikroøkonomiet de siste 30 årene. Heckman har arbeidet mye med seleksjonsproblemer som oppstår i ulike varianter når observasjoner fra et utvalg av aktører avhenger av aktørenes valg. McFadden er mest kjent for sitt bidrag til utviklingen av metoder for analyse av aktørers valg blant diskrete alternativer, dvs. når settet av tilgjengelige alternativer er endelig. Denne artikkelen gir en oversikt over de sentrale idéene i deres arbeid.

I løpet av de siste tre tiårene har det skjedd en voldsom vekst i mengden av empiriske analyser basert på mikrodata. Det er flere grunner til dette. For det første var tilgjengeligheten av mikrodata for økonomisk analyse inntil slutten av 1960 tallet temmelig begrenset. Videre var tilgangen på brukervennlige datamaskiner og EDB programmer for statistisk bearbeiding og estimering nærmest ikke-eksisterende. Som kjent har computer-og informasjonsteknologien hatt en eventyrlig utvikling, noe som har gjort det mulig å bearbeide og analysere store datamengder på en effektiv måte uten å inneha spesiell ekspertise i programmering. Hånd i hånd med IT-utviklingen har det skjedd en rask utbredelse av mikrodata-baser som har bidratt til at man har sett muligheten av å analysere nye problemstillinger. Det har i sin tur inspirert forskere til å utvikle nye økonomiske modeller og estimeringsmetoder, tilpasset disse problemstillingene. Begrepet *mikroøkonometri*



John Dagsvik er forsker i Statistisk sentralbyrå.

benyttes ofte i denne sammenheng, hvor det med dette forstås kombinasjonen av mikroøkonomisk modellering, tilhørende metoder for statistisk inferens, samt empiriske anvendelser basert på mikrodata og innsamling av slike data. Typiske temaer for mikroøkonomisk forskning er analyse av arbeidstilbud, varighet av arbeidsløshet, betydning av utdanning, etterspørsel etter varige og ikke-varige konsumgoder, valg av transportmiddel (bytransport), effekten av arbeidsmarkedstiltak og husholdningers energiforbruk.

Årets Nobelprisvinnere i økonomi, James J. Heckman og Daniel L. McFadden, har hatt helt sentrale roller i utviklingen av mikroøkonomiet de siste 30 årene. Heckman har arbeidet mye med ulike varianter av seleksjonsproblemet, spesielt i forbindelse med utdannings- og arbeidsmarkedsanalyser, mens McFadden er mest kjent for sitt bidrag til utviklingen av metoder for analyse av aktørers valg blant diskrete alternativer, dvs. når settet av tilgjengelige alternativer er endelig. Hans mest kjente anvendelser er studier av personers valg av bytransport. Imidlertid har Heckman også arbeidet med diskrete valgproblemer, og McFadden har likeså studert seleksjonsproblemet. I denne artikkelen vil jeg først og fremst legge vekt på å belyse hva jeg mener er de sentrale idéene i deres forskningsarbeid, og i mindre grad dekke alle ulike typer anvendelser de har vært involverte i. Til slutt vil jeg gi en omtale av hvordan idéene som Heckman og spesielt McFadden har vært eksponenter for, har påvirket min egen forskningsaktivitet.

Siden begge prisvinnerne har beskjeftiget seg med problemstillinger innen teori og metoder for diskret valg-handling, og dette feltet ikke er så godt kjent utenfor spesialistenes rekke har jeg valgt først å gi en summarisk oversikt over feltet. En utmerket og mer detaljert oversikt over diskret valg-handlingsteori er gitt i kapittel 2 i Anderson et al. (1992).

Diskret valg-handlingsteori; litt historikk

I mange viktige beslutnings situasjoner er alternativene diskrete eller kvalitative. Typiske eksempler er valg mellom ulike typer jobber, valg av sted å bo, hustype, transportalternativ, osv. Slike fenomener kan ikke uten videre analyseres på samme måte som valgsituasjoner der godene er uendelig delbare. Blant annet er det ofte slik i diskrete valg-

situasjoner at alternativene ikke er representert ved kvantum, men i stedet ved en vektor av kjennetegn (attributter). Historisk har modeller for diskrete valg røtter innen psykologi og psykofysikk. Psykofysikk er et felt innen psykologi med fokus på studiet av sammenhengen mellom fysiske stimuli og følelsesmessig og opplevelsesmessig respons. Grunnleggeren av psykofysikk er G. T. Fechner. For å diskriminere mellom to stimuli, postulerte Fechner (1860) modeller av typen

$$(1) \quad P(j,k) = F(v_j - v_k)$$

der $P(j,k)$ er sannsynligheten for at aktøren skal vurdere stimulus j som sterkere enn stimulus k , F er en kumulativ fordelingsfunksjon og $\{v_j, j=1,2,\dots\}$ er en skala (uobserverbar) som representerer styrken av de respektive stimuli. Han drøfter med dette utgangspunkt hvordan skalaen $\{v_j\}$ kan tolkes og konstrueres fra data. Det er interessant å legge merke til at allerede så tidlig som i 1860 benyttes en stokastisk formulering for å ivareta det observerte fenomen at samme individ i repeterte eksperimenter ofte foretar ulik vurdering. I typiske «laboratorie-eksperimenter» i psykofysikk blir individer for eksempel bedt om å avgjøre hvilke av to objekter som er tyngst, hvilke av to lydsignaler som er sterkest, osv. En observerer da typisk at når for eksempel samme vektløftingseksperiment blir repetert så vurderer en gitt aktør ulikt fra eksperiment til eksperiment selv om tyngden på vektene er uforandret.

I 1927 introduserte psykologen Thurstone sin «*Law of comparative judgment*». Tilsvarende Fechner postulerte Thurstone en stokastisk modell for diskriminering mellom to stimuli. Thurstone antok at en aktør som står ovenfor en vurdering av stimuli j og k opplever at stimuliene gir henholdsvis nytte $v_j + e_j$ og $v_k + e_k$, der e_j og e_k er stokastiske restledd. De stokastiske restleddene er ment å representere aktørens problem med å presist evaluere verdien av stimuliene. Thurstone antok at disse restleddene var normalfordelte. Dette medfører at

$$(2) \quad P(j,k) = G(v_j - v_k)$$

der G betegner den kumulative normalfordelingsfunksjonen med forventning null. Dersom Thurstones modell generaliseres slik at $e_j - e_k$ har kumulativ fordelingsfunksjon F , impliserer dette at valgsannsynlighetene blir slik som i Fechners modell (1). Selv om Fechners modell er en stokastisk modell, er Thurstone den første til å formulere en eksplisitt stokastisk modell der selve *nyttefunksjonen* er stokastisk (Random Utility Model; RUM), og i motsetning til Fechner er Thurstones formulering mer i samsvar med tradisjonen i mikroøkonomisk nytte teori.

På 1950-tallet ble det publisert en rekke fundamentale arbeider i diskret valghandlingsteori, først og fremst av Marschak (1960) og Luce (1959). Luce (1959) postulerte en stokastisk variant av aksiomet, «Uavhengighet fra irrelevante alternativer» (IIA), som vi nå skal forklare. La B betegne mengden av alternativer (endelig) en aktør i står overfor, og $J_i(B)$ det mest attraktive alternativet i B , dvs. det alternativet som velges. Slik som hos Fechner oppfattes $J_i(B)$

som stokastisk. Vi forestiller oss at aktørens valghandling kan sees på som om den foregår i to (eller flere) trinn. I første trinn velges en delmengde som inneholder de mest attraktive alternativer i B . I trinn 2 velges det mest attraktive alternativet fra denne delmengden. Her er poenget at i trinn to forutsettes aktøren å velge fra delmengden som ble valgt i første trinn, som om denne var den opprinnelige valgmengden. Dvs. de alternativer som ble fjernet i første trinn er irrelevante i trinn to. Vi kan uttrykke dette formelt som

$$(3) \quad P(J_i(B) = j | J_i(B) \in A) = P(J_i(A) = j)$$

for $j \in A \subseteq B$. Likning (3) sier at sannsynligheten for at j skal velges fra B gitt at det mest attraktive alternativet befinner seg i delmengden A , er lik sannsynligheten for å velge j når A er *valgmengden*.¹

Luce kaller IIA et *probabilistisk rasjonalitetsprinsipp*. Det er et rasjonalitetsprinsipp i den forstand at bare de alternativene som er tilgjengelige på det aktuelle tidspunkt når valget skal foretas, skal influere på valget.² Luce (1959) viser at (3) holder hvis og bare hvis det eksisterer en skala, $\{v_j, j = 1, 2, \dots\}$, slik at

$$(4) \quad P(J_i(B) = j) = \frac{\exp(g(X_i, Z_j; b))}{\sum_{k \in B} \exp(g(X_i, Z_k; b))}$$

Denne skalaen er éntydig på addisjon av en konstant nær og den avhenger ikke av valgmengden B . Dette betyr at dersom skalaverdiene er kjente kan en skrive ned sannsynlighetene for en vilkårlig valgmengde ved hjelp av (4). Luce tolket $\{v_j\}$ som en nytteindeks som representerte attraktiviteten til alternativene. For binære sammenlikninger reduserer (4) seg til en variant av Fechners modell med $F(x) = (1 + \exp(-x))^{-1}$. I Luces arbeid ble det ikke drøftet om Lucemodellen (4) også kunne utledes fra en stokastisk nyttefunksjon slik som hos Thurstone. Holman og Marley (se Luce og Suppes, 1965) viste at dersom en antar at alternativ j har nytte U_{ij} gitt ved $U_{ij} = v_j + e_{ij}$, der restleddene, e_{ij} , $i = 1, 2, \dots, j = 1, 2, \dots$, er uavhengige og identisk fordelte stokastiske variable med en bestemt fordelingsfunksjon som kalles *ekstremverdifordelingen*³, så vil sannsynligheten for at U_{ij} skal være den høy-

¹ Venstre side av (3) er en betinget sannsynlighet slik den er definert i sannsynlighetsteorien. Det empiriske motstykke til venstresiden er andelen av aktører (i en observerbar homogen populasjon) med valgmengde B som har valgt alternativ j , blant alle aktører med valgmengde B som har valgt et alternativ i A . Høyre side av (3) er også en «betinget» sannsynlighet i den forstand at en her betinger med hensyn på at *valgmengden* er lik A . Det siste er (pr. definisjon) ikke det samme som betinging i sannsynlighetsteoretisk forstand. Det empiriske motstykke til høyresiden er andelen av aktører med valgmengde A som har valgt alternativ j . Generelt er derfor likning (3) ikke gyldig, fordi den uttrykker at to ulike definisjoner av «betinging» er like.

² Det er dermed ikke sagt at det ikke fins andre fornuftige rasjonalitetsprinsipper. For eksempel er maksimering av stokastisk nytte à la Thurstone et intuitivt rasjonalitetsprinsipp, men dette prinsippet er mye svakere.

³ Ekstremverdifordelingen er kjent fra statistisk teori som den asymptotiske fordelingen til maksimum av uavhengige identisk fordelte stokastiske variable. Det er tre typer av denne fordelingen. Den som er aktuell her har (standardisert) kumulativ fordelingsfunksjon lik $\exp(-\exp(-x))$, og kalles type I av noen forfattere og type III av andre.

este nytte være gitt ved (4). Dette betyr at v_j kan tolkes som gjennomsnittsnytten over repetisjoner av identiske valgeksperimenter for en gitt aktør, eller alternativt som gjennomsnittsnytten i en populasjon med identisk observerbare aktører, eventuelt en kombinasjon av begge deler. Det er imidlertid ikke lett å ha intuisjon på hvorfor denne fordelingsfunksjonen oppstår i sammenheng med IIA.

En av de tidligste arbeidene som diskuterte RUM i forhold til mikroøkonomiske anvendelser var Quandt (1956). Men trolig var det først og fremst McFaddens langvarige innsats og evne til å kombinere teori og statistisk metode, samt å demonstrere relevansen i empiriske anvendelser, som spilte en avgjørende rolle for den utbredelse dette feltet har i dag. Heckman har også bidratt vesentlig i denne sammenheng ved å demonstrere potensialet slike metoder har, for eksempel innen empiriske analyser av arbeidstilbud.

Daniel L. McFadden

McFadden ble født i Raleigh, North Carolina i 1937, og gjennomførte lavere grads studier ved universitetet i Minnesota med fysikk som viktigste fag. Senere gikk han over til økonomistudier og tok sin Ph.D. ved samme universitet i 1962. Han har vært knyttet til universitetene i Pittsburgh, Yale, Massachusetts Institute of Technology (MIT), og University of California, Berkeley (UCB). Siden 1990 har han vært professor ved UCB. Hans mest kjente forskningsbidrag er innen teori og metode for diskret valg-handling. Men han har også gjort betydelige bidrag innen andre områder. På 60-tallet og tidlig på 70-tallet utarbeidet han betydelige teoretiske og økonomiske bidrag innen produksjonsteori, se Fuss og McFadden (1978). I 1990 årene arbeidet McFadden mye med problemstillinger knyttet til miljøøkonomi/økonometri, slik som metoder for estimering av betalingsvillighet for naturressurser, se for eksempel Hausman et al. (1995). McFadden har videre analysert etter-spørsmål etter bolig blant eldre (McFadden, 1994).

Analyse av diskrete valg

Mens han studerte til doktorgraden, ble McFadden svært interessert i aksiomatiske valgteorier utviklet av matematisk psykologer, slik som de ble studert av Luce, Marschak og Marley. I 1973 publiserte han et berømt arbeid som inneholder både teoretiske bidrag, videreutvikling av Lucemodellen til et rammeverk som kan benyttes som utgangspunkt for økonometrisk analyse, samt korresponderende «skreddersydd» statistisk inferensteori. Kort oppsummert tar han utgangspunkt i en RUM der aktøren står ovenfor en valgmenge- de B (som kan være aktørspesifikk), med alternativer karakterisert ved alternativspesifikke kjennetegn (attributter) Z_j . Nyttien ved alternativ j for aktør i er gitt ved

$$(5) \quad U_{ij} = g(X_i, Z_j; b) + e_{ij}$$

der e_{ij} , $j = 1, 2, \dots, i = 1, 2, \dots$, er uavhengige og identisk ekstremfordelte restledd, og $g(\cdot)$ er en passende valgt funksjon som avhenger av en ukjent parametervektor b . Variabelen X representerer karakteristika ved aktøren. Som

forklart ovenfor, og vist av Holman og Marley (referert til i Luce og Suppes, 1965), gir dette at sannsynligheten for å velge alternativ j fra mengden B kan uttrykkes som

$$(6) \quad P(J_i(B) = j) = \frac{\exp(g(X_i, Z_j; b))}{\sum_{k \in B} \exp(g(X_i, Z_k; b))}$$

Dersom en har mikrodata med observasjoner av personers valg samt observasjoner av Z_j for alle j , og for X_i for alle personer i utvalget, kan en estimere b med standard statistiske metoder. McFaddens teoretiske bidrag i dette arbeide var å vise at under forutsetningen om at restleddene er uavhengige og identisk fordelte (og i tillegg noen regularitetsbetingelser), så er ekstremverdifordelingen den *eneste* fordelingen som er konsistent med IIA, dvs. den eneste fordelingen som gir modellen (6). McFadden kalte denne modellen den betingede logitmodellen (betingent med hensyn på attributtene). En mer vanlig betegnelse er den *multinomiske logitmodellen*. McFadden skrev allerede i 1967 et FORTRAN program som estimerer modellen ved sannsynlighetsmaksimeringsmetoden. Når modellen er estimert kan en predikere effekten av endringer i attributter (Z_j) og personkjennetegn (X_i). Men siden parametervektoren b ikke er alternativ-spesifikk kan en også predikere effekten av endringer i *valgmengden*, dersom en spesifiserer attributtene for eventuelle nye alternativer som inkluderes i valgmengden.⁴

I 1972 ble McFadden leder for *Travel demand forecasting project* ved UCB. En del av dette prosjektet var å predikere trafikkgrunnlaget for det nye tunnelbanesystemet BART (Bay Area Rapid Transit) som var under planlegging og bygging. Et utvalg på ca. 630 personer ble spurt om sine transportvalg (gitt de eksisterende transportalternativer). Det ble gjort betydelig arbeid med å kartlegge kostnader og reisetider, og en multinomisk logitmodell av type (6) ble estimert. I 1975 var BART ferdig og en kunne da sammenlikne prediksjonene av trafikkfordelingen på de nye alternativene med de faktiske realiserte. Tabellen nedenfor oppsummerer resultatene.

Pre-BART prediksjonsresultater og post-BART data

	Alternativer					
	Privatbil alene	Buss/Gå	Buss/ Privatbil	BART/ Buss	BART/ Privatbil	Kjøre bil sammen
Predikerte andeler	55,8%	12,5%	2,4%	1,1%	5,3%	22,9
(St.avvik)	(11,4%)	(3,4%)	(1,4%)	(0,5%)	(2,4%)	(10,7%)
Observerte andeler	59,9%	10,8%	1,4%	1,0%	5,2%	21,7%

Med «kombinasjonsalternativene», «BART/buss», og «Buss/Gå», menes det henholdsvis BART i kombinasjon med

⁴ Det benyttes også formuleringer av den multinomiske logitmodellen for diskrete valg som inneholder alternativ-spesifikke parametre. I tilfelle med alternativ-spesifikke parametre kan en naturligvis ikke uten videre predikere valgene når nye alternativer inkluderes i valgmengden.

matebuss, og Buss i kombinasjon med å gå til/fra bussholdeplassen. Tilsvarende for de andre kombinasjonsalternativene. De observerte andelene er beregnet for det samme utvalget som deltok i 1972 undersøkelsen. Standardavviket i parentes skyldes usikkerheten i parameterestimaterne. Som vi ser er prediksjonene for de nye alternativene BART/buss og BART/privatbil forbausende presise, men det kan skyldes «flaks», siden standardavvikene er forholdsvis store. De andre 4 alternativene er de som var tilgjengelige i 1972. For mer informasjon om denne undersøkelsen henvises det til McFadden (1978b, 2000).

Luce (1959) var klar over at IIA i mange situasjoner er urealistisk. Dette gjelder situasjoner der noen attributter er uobserverbare (for forskeren) og korrelert over noen alternativer. Følgende eksempel illustrerer dette: Sett at aktører står overfor valg av bosted. Det er tre alternativer tilgjengelige, nemlig Tønsberg, Fredrikstad og Ålesund. Det er nærliggende å tenke seg at værforholdene, og muligens lokal væremåte kan ha betydning for aktørens rangering av alternativene. Vær- og væremåte antar vi er uobserverbare for forskeren. Det er klart at vær- og væremåte er mer like i Tønsberg og Fredrikstad i forhold til i Ålesund. Følgelig er det nærliggende at restleddene i nyttefunksjonen knyttet til alternativene Tønsberg og Fredrikstad er sterkere korrelerte enn restleddene knyttet til henholdsvis Ålesund og Tønsberg, og Ålesund og Fredrikstad. Hvis dette er tilfelle vil IIA ikke holde, og den multinomiske logitmodellen vil ikke være en god modell. En annen Nobelprisvinner, Debreu (1960), diskuterte dette ved hjelp av et tilsvarende, om enn mer ekstremt eksempel. Dette ledet til intensiverende anstrengelser for å generalisere modellen (6), til en modell som tillater at restleddene er korrelerte over alternativer. Man ble nemlig fort klar over at det ikke først og fremst var formen på fordelingsfunksjonen til restleddene som var problemet, men at de er antatt uavhengige, eller eventuelt at korrelasjonen er den samme mellom alle restleddkombinasjoner. En opplagt kandidat var den *multinomiske probitmodellen* som er en generalisering av den binære probitmodellen foreslått av Thurstone, dvs. en RUM der restleddene er normalfordelte og kan være korrelerte. Imidlertid er den multinomiske probitmodellen ubekvem å arbeide med fordi uttrykkene for valgsannsynlighetene er ekstremt kompliserte og kostnads-krevende å beregne, noe vi skal komme tilbake til senere.

I 1977-78 lanserte McFadden den såkalte GEV (Generaliserte EkstremVerdi) modellen (McFadden, 1978). Som ofte er tilfellet ved nyvinninger i forskingen var det andre som også bidro til dette (se McFadden, 1981), men det er trolig riktig å si at McFadden hadde en avgjørende betydning for utforming, tolkning og implementering av ulike varianter av GEV modellen. GEV modellen var basert på nyttefunksjoner med simultant ekstremverdifordelte restledd, hvilket medfører at restleddene tillates å være korrelert over alternativene. I motsetning til den multinomiske Probitmodellen var det ikke opplagt hvordan denne generaliseringen skulle gjøres. Som spesialtilfelle av GEV er den såkalte *nøstede multinomiske logitmodellen* velkjent. Den nøstede multinomiske logitmodellen er mest naturlig i situasjoner der alternativmengden kan ordnes hierarkisk ifølge en «trestruktur». Et eksempel på en slik trestruktur har en når en skal

modellere arbeidstilbud og alternativene er «ikke jobb», «deltidsjobb» og «heltidsjobb». I dette eksemplet er «hovedgrenene» «jobb» kontra «ikke jobb». Gitt at valget er å jobbe så er det to mindre «grener», nemlig «deltid» og «heltid». I McFadden (1981) gis det en rigorøs behandling av diskret valghandlingsteori der blant annet disse problemene diskuteres. Denne artikkelen inneholder også en del nytt stoff; han diskuterer for eksempel betingelser som valgsannsynlighetene må oppfylle for at de skal være konsistente med RUM. I et arbeid sammen med Manski (Manski og McFadden, 1981) utvikles «skreddersydd» statistisk inferensteori for diskrete valgmodeller under ulike typer utvalgsdesign.

Selv om den nøstede multinomiske logitmodellen gir en tilfredstillende modellramme i mange situasjoner, har den multinomiske probitmodellen visse fordeler modelleringsmessig (når en ser bort fra det beregningstekniske problemet) siden denne modellrammen er så fleksibel. McFadden(1989) løste langt på vei det beregningstekniske problemet ved å introdusere sin *method of simulated moments*. Dette er en Monte Carlo simuleringmetode som kan benyttes effektivt for å beregne og å estimere multinomiske probitmodeller og andre modeller der valgsannsynlighetene ikke kan uttrykkes på lukket form. Dette arbeidet førte til en intens forskningsaktivitet for å utvikle mer raffinerte og praktiske beregnings- og simulering algoritmer.

I et nylig publisert arbeid, McFadden og Train (2000), vises det at en generell RUM kan approksimeres så nært en ønsker til en såkalt *blandet multinomisk logitmodell*. Dette er en multinomisk logitmodell der parametrene antas individspesifikke og antas å være fordelt over populasjonen ifølge en delvis spesifisert fordeling med parametre som bestemmes av data. I dette arbeidet diskuteres videre praktiske problemer knyttet til estimering og beregningsalgoritmer.

Både den multinomiske probitmodellen, den nøstede-og blandete multinomiske logitmodellen har på den ene side den fordel at de er svært fleksible rammeverk for empirisk modellering. På den andre siden er dette også en ulempe ved disse modellene fordi restriksjonene som følger fra teorien er svake, og en risikerer derfor at egenskapene ved modellen blir lite robust overfor mer eller mindre ad hoc spesifikasjon av strukturen i nyttefunksjonen og korrelasjonsstrukturen i restleddene. Det gjenstår å se hvilke av disse modell-typene som vil vise seg å være den mest fruktbare i anvendelser- og i hvilken utstrekning den multinomiske logitmodellen er for restriktiv.

I en rekke interessante anvendelser er alternativsettet en blanding av diskrete- og uendelig delbare alternativer. Typiske eksempler finner en i arbeidstilbud der det diskrete valg er alternativene «jobb» kontra «ikke jobb», og det kontinuerlige valget er hvor mye arbeidstid som tilbys gitt jobb. Et annet eksempel er valg av oppvarmingsutstyr i boliger, samt intensitet i energiforbruk. Teori og metode som benyttes i denne typen situasjoner kalles teori for diskret/kontinuerlig valghandling, og på dette feltet er både Heckman og McFadden sentrale. Heckmans modell for arbeidstilbud er, som vi skal belyse nedenfor, en modell med diskret/kontinuerlig valg. I Dubin og McFadden (1984) vises det hvordan en

kan kombinere teorien for diskrete valg med konvensjonell dualitetsteori til et slagkraftig analyseapparat. Vi vil illustrere metoden ved valg av oppvarmingsutstyr og energiforbruk som eksempel. La V_j representere den betingete indirekte nytten gitt at oppvarmingsutstyr av type j er valgt, dvs. V_j er funnet ved å maksimere nytten med hensyn på varmebehov samt konsum av andre varer gitt j og gitt budsjettbetingelsen - inkludert brukerkostnaden ved anskaffelse av oppvarmingsutstyr av type j . Den betingete indirekte nytten V_j vil altså være en funksjon av denne brukerkostnaden, av energiprisene, prisene på andre varer samt husholdningens inntekt. Valg av type oppvarmingsutstyr vil derfor være bestemt ved maksimering av V_j med hensyn på j . Dersom for eksempel V_2 er størst blir type 2 er valgt, og energiforbruk finnes ved å avlede etterspørselsfunksjonen(e) ved hjelp av Roys identitet, anvendt på den indirekte nytten, V_2 .⁵

James J. Heckman

James J. Heckman ble født i Chicago i 1944. Etter lavere grads studier ved Colorado College med matematikk som viktigste fag, studerte han til doktorgraden ved Princeton University hvor han fikk sin Ph.D. i 1971. Heckman har undervist ved Columbia University, Yale University og University of Chicago. Siden 1995 har han vært professor ved University of Chicago. Han har arbeidet med teoretiske og anvendte økonometriske problemstillinger knyttet til seleksjon, diskrete valg, dynamiske analyser og varighetsanalyser. Heckman har også arbeidet med matematikk knyttet til effekten av utdanning (Heckman, 2000a), og har videre interessert seg for mer prinsipielle vitenskapsteoretiske spørsmål, se for eksempel Heckman (1992, 2000b).

Arbeidstilbudsanalyser

I mikroøkonometriske analyser oppstår ofte ulike varianter av *seleksjonsproblemet*. Dersom en for eksempel skal estimere en empirisk relasjon for personers timelønn i arbeidsmarkedet, står en ovenfor det problemet at en bare observerer timelønna for de personene som faktisk jobber. Dersom en bare er opptatt av en lønnsrelasjon gitt at personen jobber er dette et mindre problem. Men som regel er en interessert i en relasjon uavhengig av arbeidsmarkedstatus, slik at en vil være i stand til å predikere timelønna til de som ikke er i jobb dersom de begynner å jobbe. Å predikere timelønna til de som ikke jobber ut fra observasjonene for de som jobber kan bli galt fordi de to gruppene antakeligvis er ulike. Problemet er at disse forskjellene i stor grad er uobserverbare. Et eksempel på en uobserverbar variabel som kan ha betydning her er «motivasjon». Vi skal nedenfor belyse dette problemet mer eksplisitt ved hjelp av et eksempel hentet fra Heckman (1974).

Betrakt aktør i (gift kvinne) som står ovenfor timelønn W_i i arbeidsmarkedet. Det antas at aktøren fritt kan oppnå så mye arbeid hun vil i markedet. La $S_i(h)$ betegne timelønningen som får aktør i til å ville arbeide h timer (skyggelønn til arbeidstid h). Generelt vil skyggeprisen avhenge både av arbeidstid og ektefelles inntekt. Heckman antar følgende empiriske modellspesifikasjon:

$$(7) \quad \text{Log}S_i(h) = ah + X_{1i}b + e_{1i},$$

$$(8) \quad \text{Log}W_i = X_{2i}c + e_{2i},$$

der X_{1i} og X_{2i} er nærmere spesifiserte vektorer av forklaringsvariable som er antatt å påvirke henholdsvis skyggelønn og timelønn, e_{1i} og e_{2i} er normalfordelte restledd med forventning null, gitt X_{1i} og X_{2i} , og a , b og c representerer ukjente parametre. Vektoren X_{1i} inneholder ektefelles inntekt. I følge standard mikroøkonomisk teori vil aktør i velge å jobbe dersom reservaslønna $S_i(0)$ er mindre enn timelønna W_i . I motsatt fall vil aktøren ikke jobbe. Gitt at personen jobber (dvs. når $S_i(0) < W_i$), så bestemmes arbeidstiden h ved likningen $S_i(h) = W_i$. Dette betyr at den avhengige variabel i likningene bare observeres når $S_i(0) < W_i$. Dette er en variant av en såkalt diskret/kontinuerlig modelltype. Som antydnet ovenfor kan en ikke her uten videre estimere likningene ovenfor ved for eksempel minste kvadraters metode på grunnlag av data for de personer som jobber. Utvalgseleksjonsregelen, $S_i(0) < W_i$, medfører at restleddene i likningene (7) og (8) ikke nødvendigvis har gjennomsnitt lik null over utvalget av de som jobber, hvilket medfører skjeve estimater dersom en benytter standard regresjonsanalyse. I forhold til tidligere arbeider innen arbeidstilbud, skiller Heckmans tilnærming seg fra disse ved at han knytter seg sterkere opp til mikroøkonomisk teori, slik likningssystemet ovenfor illustrerer. Heckman (1974) drøfter inngående denne modellen og gjennomfører en simultan estimering av modellen på grunnlag av mikrodata for gifte kvinner, og ved å benytte sannsynlighetsmaksimeringsprinsippet. I 1974 var dette en avansert økonometrisk analyse. Senere viser han, Heckman (1979), at modeller av typen ovenfor kan estimeres ved en enklere men dog mindre effektiv metode. Vi skal kort belyse idéen i denne forenklete metoden ved å betrakte estimeringsproblemet for likning (8). La oss først betrakte beslutningen om å jobbe eller ikke. Fra (7) og (8) får vi ved passende innsetting at sannsynligheten for å jobbe kan uttrykkes på følgende måte;

$$(9) \quad P(S_i(0) < W_i) = H(X_i d),$$

der H representerer den standardiserte kumulative normalfordelingsfunksjonen, X_i er en vektor som består av alle variable i X_{1i} og X_{2i} , og d er en ukjent parametervektor. Modellen (9) er en (reduert form) binær Probit-modell for beslutningen jobb/ikke jobb, tilsvarende formuleringen til Thurstone, og en kan estimere d ved sannsynlighetsmaksim-

⁵ Under estimeringen vil en her ofte stå overfor et seleksjonsproblem fordi en bare kan estimere etterspørselsfunksjonene gitt valg av type oppvarmingsutstyr på grunnlag av de husholdningene som har valgt denne typen. Tilsvarende som i estimering av Heckmans arbeidstilbudsmodell (omtalt i kapitlet om Heckman) er grunnen til dette at restleddene i de avledede etterspørselsfunksjonene vil avhenge av de samme restleddene som inngår i den indirekte nyttefunksjonen, $\{V_k\}$, slik at utvalget av de som har valgt oppvarmingsutstyr av type j ikke vil være et tilfeldig trukket utvalg. For en mer uførlig drøfting av seleksjonsproblemet viser vi til avsnittet om Heckman. I Dubin & McFadden (1984) vises det hvordan en ved å bruke tilsvarende idé som Heckman (1979) kan korrigere for mulige seleksjonseffekter under estimeringen.

meringsmetoden. Formelt kan forventningen til restleddet i (8), gitt at aktøren jobber, uttrykkes som $E(e_{2i} | S_i(0) < W_i)$. Ved å benytte relasjonene ovenfor kan en beregne et analytisk uttrykk for denne betingede forventningen, gitt ved

$$(10) \quad E(e_{2i} | S_i(0) < W_i) = ks(X_i d),$$

der $s(X_i d) = H(X_i d) / H(X_i)$ og k er en ukjent parameter. Vi ser altså at funksjonen $s(\cdot)$ er kjent slik at dersom d er gitt kan en beregne $s(X_i d)$. Heckmans estimeringsprosedyre foregår nå i to trinn: I første trinn estimeres d basert på (9), ved å benytte observasjoner for hvorvidt personene jobber eller ikke (i tillegg til observasjonene for X_i). I trinn 2 benyttes regresjonsanalyse til å estimere c og k basert på relasjonen

$$(11) \quad \text{Log} W_i = X_{2i} c + ks(X_i d^*) + e_{3i}$$

på grunnlag av data for de som jobber ved å benytte $s(X_i d^*)$ som en ekstra forklaringsvariabel, der d^* betegner estimatet på d oppnådd i første trinn. Det er enkelt å vise at denne prosedyren kontrollerer for seleksjonsskjevheter, dvs. at forventningen til restleddet e_{3i} gitt at aktøren jobber er lik null.⁶ Den simultane metoden omtalt ovenfor, (Heckman, 1974), er nå implementert som standard kommando i mange program-pakker, slik at Heckmans to-trinns metode, (Heckman, 1979), etter hvert har fått mindre betydning. Et eksempel på en mer avansert arbeidsmarkedstudie der seleksjon spiller en viktig rolle finner vi i Heckman og Sedlacek (1985).

Selv om Heckman på begynnelsen av 1970 tallet ikke var alene om å studere seleksjonsproblemer, var han den som mest kraftfullt arbeidet med dette. Dette har etter hvert nærmest ført til en «industri» der det arbeides med metoder og empiriske analyser av ulike typer seleksjonsproblemer.

Evaluering av arbeidsmarkedstiltak

Heckman har beskjeftiget seg mye med å utvikle metoder for evaluering av effekten av arbeidsmarkedstiltak. Vi skal illustrere typiske problemstillinger ved å ta utgangspunkt i den såkalte Roy-modellen (Roy, 1951). La U_{0i} og U_{1i} være henholdsvis nytten til aktør i av ikke å delta og å delta i et arbeidsmarkedsprogram. Formålet med å delta er forventning om økt lønn i arbeidsmarkedet etterpå. La Y_{0i} og Y_{1i} være henholdsvis lønn gitt at aktøren ikke har deltatt, eller har deltatt i programmet. Anta følgende lønnsrelasjon,

$$(12) \quad Y_{ij} = aD_i + X_i b + e_{ij},$$

for $j=1,2$, der X_i er en vektor av relevante individuelle kjennetegn, D_i er en dummyvariabel som er lik én hvis aktør i har deltatt i programmet og null ellers. Restleddet e_{ij} har forventning null og er ukorrelert med X_i , a og b er ukjente parametre. Vi ser at

$$(13) \quad E(Y_{1i} - Y_{0i} | X_i) = a,$$

slik at a representerer gjennomsnittseffekten av arbeidsmarkedsprogrammet. Spørsmålet er nå hvordan man skal esti-

mere a . Vi observerer Y_{1i} bare for de som har deltatt i programmet og Y_{0i} bare for de som ikke har deltatt i programmet. Her står vi altså på nytt overfor et seleksjonsproblem. Tilsvarende estimeringen av timelønnsrelasjonen (8) vil gjennomsnittet av restleddet e_{1i} over de personer som har deltatt i programmet ikke nødvendigvis være lik null fordi det er grunn til å anta at uobserverbare faktorer slik som evne og motivasjon påvirker både restledd i (12) og nytten av å delta i programmet. Tilsvarende kan gjennomsnittet av restleddet e_{0i} over de som ikke deltar i programmet være forskjellig fra null. Formelt kan vi uttrykke dette som

$$(14) \quad E(e_{1i} | U_{1i} > U_{0i}) \neq 0 \text{ og } E(e_{0i} | U_{1i} < U_{0i}) \neq 0.$$

Altså må vi ta hensyn til dette i estimeringsmetoden. Heckman har diskutert en hel rekke metodiske problemer som en må tenke over i denne sammenhengen. Dette gjelder ikke minst robusthet med hensyn på forutsetninger om funksjonsform og fordelingssegenskaper til de stokastiske restleddene. Blant annet har han deltatt i en omfattende debatt med andre faggrupper (statistikere og sosiologer) om ulemper og fordeler ved bruk av eksperimentelle kontra ikke-eksperimentelle data. Han har i tillegg presentert omfattende empiriske resultater om effekten av slike program, se Heckman et al. (1989, 1999). En av konklusjonene som synes å framgå av disse studiene er at arbeidsmarkedsprogrammer har svært liten effekt.

Dynamisk analyse og analyse av varighet

I mange sammenhenger er en interessert i å studere hva som påvirker varighet, for eksempel varighet av arbeidsløshet, levetid for bedrifter, hvor lenge en er i et jobbforhold før en bytter jobb, osv. Heckman (1981) introduserer en flerperiode multinomisk probitmodell, dvs. en RUM for valg over flere tidspunkt der restleddene kan være autokorrelerte, og det i tillegg kan være såkalt strukturell tilstandsavhengighet som skyldes at preferanser og muligheter kan påvirkes av erfaringer fra tidligere valg. La oss belyse en problemstilling ved et aktuelt eksempel. I analyser av varighet av arbeidsledighet observerer man typisk at personer som har vært ledig lenge har lavere tilbøyelighet for å få jobb enn personer (med samme observerbare kjennetegn) som har vært arbeidsledig kortere tid (positiv varighetsavhengighet). Det kan tenkes to forklaringer på dette. Den enkle forklaringen er at preferanser og muligheter blir påvirket av ledighet. Den andre forklaringen er at positiv varighetsavhengighet skyldes uobserverbar heterogenitet i preferanser og kvalifikasjoner. Dette fører til en seleksjonsprosess der de med høyest uobserverbare kvalifikasjonsvariable får jobb først. Eventuelt kan det være en blanding av disse to effektene. Heckman (1981, 1991) påpeker at dette fundamentale identifikasjonsproblemet ikke kan løses rent statistisk. Man trenger ekstra teoretiske forutsetninger for å sortere ut disse effektene.

⁶ Det er et sentralt poeng at vektorene X_1 og X_2 ikke er like. I motsatt fall vil identifikasjon av c og k i (11) bare avhenge av funksjonsformen til $s(\cdot)$, og en risikerer dermed at estimatene blir sensitive overfor valg av sannsynlighetsfordeling for restleddene.

Heckman har i mange sammenhenger studert virkningen av uobserverbar heterogenitet både metodisk og empirisk. I samarbeid med MaCurdy har han utvidet arbeidstilbudsmodellen presentert ovenfor til en diskret/kontinuerlig livsløpsmodell som estimeres ved hjelp av mikrodata for gifte kvinner, se Heckman og MaCurdy (1980). I samarbeid med Burton Singer og andre forskere har Heckman studert identifikasjon og estimering i varighetsmodeller, se for eksempel Heckman og Singer (1984a,b). I Flinn og Heckman (1982) vises det blant annet at estimatene i den stasjonære modellen for jobbsøking er ekstremt sensitive med hensyn på spesifiserte funksjonsform.

Inspirasjonskilde til egen forskningsaktivitet

Heckman, og spesielt McFadden har hatt stor betydning for min egen forskningsaktivitet. Her vil jeg nevne noen av de problemstillingene jeg har blitt inspirert av disse forskerne til å arbeide med. Ved en tilfeldighet kom jeg i kontakt med litteraturen for diskrete valg og arbeidstilbud i siste halvdel av 1970 tallet, og jeg ble fascinert og inspirert av denne måten å modellere på. Jeg begynte tidlig å forsøke å generalisere den multinomiske logitmodellen for diskrete valg til situasjoner der (myopiske) valg foregår ved flere tidspunkter. Her kan valgene være korrelerte over tid pga. variable som er uobserverbare for forskeren og som påvirker preferansene og er korrelerte over tid. I tillegg kan det oppstå korrelasjon over tid fordi preferanser og rammebetingelser kan påvirkes av tidligere valg (tilstandsavhengighet). Dette er analogt til Heckmans flerperiode probitmodell, omtalt ovenfor, som generaliserer den multinomiske probitmodell til flere perioder. I Dagsvik (1983, 1988) etablerte jeg en RUM som avledes av en bestemt type stokastiske nyttefunksjoner for valg ved flere tidspunkter (i kontinuerlig tid), og som generaliserer henholdsvis den multinomiske logitmodellen og GEV modellen. I Dagsvik (1996, 2000b) viste jeg at denne flerperiode RUM er konsistent med en generalisering av IIA.

I Dagsvik (1994) drøfter jeg en generell klasse av RUM som er en utvidelse av GEV modellen til en versjon for «kontinuerlig» valg. Innen dette rammeverket er det mulig å modellere situasjoner der mengden av tilgjengelige alternativer er uobserverbar og stokastisk for forskeren (motivert ut fra uobserverbar heterogenitet i muligheter). Her ble jeg inspirert av medarbeidere til McFadden; Ben-Akiva og Watanatada (1981), som tidligere hadde arbeidet med tilsvarende idé. Innen for eksempel analyse av arbeidstilbud er det interessant å benytte dette opplegget. I den tradisjonelle måten å modellere arbeidstilbud på forutsettes aktøren bare å ha preferanse for fritid og konsum, og videre å kunne tilpasse seg fritt gitt den økonomiske budsjettbetingelsen. I det alternative opplegget antas aktørene, i tillegg til å ha preferanser over fritid og konsum, også å ha preferanser over type jobber, karakterisert ved arbeidstid og timelønn samt delvis uobserverbare jobbattributter. Valgmengden for en aktør består av settet av tilgjengelige jobber (aktørspesifikk), og er typisk uobserverbar for forskeren. Empiriske anvendelse innen arbeidstilbud som benytter dette rammeverket er utført i samarbeid med kolleger, og er dokumentert i blant annet Aaberge et al. (1995).

Som beskrevet ovenfor utviklet McFadden GEV modellen i siste halvdel av 1970 årene. Det var imidlertid uklart hvilke restriksjoner på valgsannsynlighetene som følger fra GEV strukturen utover det at den er en RUM. Nødvendige og tilstrekkelige betingelser for at en valgmodell skal være RUM ble etablert av Falmagne (1978). Dette problemet drøfter jeg i Dagsvik (1995), og viser at enhver RUM kan approksimeres så nært en ønsker med en GEV modell. Dette betyr altså at GEV modellen i praksis ikke impliserer noen restriksjoner utover de som følger av en RUM.

Senere har jeg arbeidet med diskret valg når valgmengdene aktørene står ovenfor er endogene, se Dagsvik(2000a). Dette gjelder i markeder med matching, dvs. markeder der aktørene søker etter en partner i en populasjon av potensielle partnere. Typiske eksempler der matching foregår er i tilpassningen på arbeidsmarkedet, ekteskapsmarkedet og utdanningsmarkedet.

Oppsummering

Jeg har i denne artikkelen drøftet noen av de idéer og bidrag som er sentrale i Heckmans og McFaddens forskning. McFaddens viktigste enkeltstående bidrag er hans artikkel fra 1973, generaliseringen av den multinomiske logitmodellen til GEV modellen, samt utviklingen av praktiske Monte Carlo simuleringsalgoritmer for estimering og beregning, McFadden (1989). McFadden har i disse arbeidene demonstrert sin spesielle evne til å etablere en tett forbindelse mellom aksiomatisk teori, empirisk modellformulering og statistisk inferensteori.

Heckmans viktigste bidrag er å vise hvor gjennomgripende seleksjonsproblemet er, og hvor ofte det oppstår i ulike varianter. Dette fører til problemer med tolkningen av resultatene i økonometriske studier. Det er typisk for Heckman at han «presser» tolkning og drøfting av empirisk relevans i det teoretiske og økonometriske rammeverket. Nå som vi har blitt vant til å tenke på seleksjonsproblemet er det lett å glemme hvor abstrakt og vanskelig det var å begripe i sin fulle konsekvens når en første gang begynte å tenke på dette.

Det som er et fellestrekk ved forskningsinnsatsen til Heckman og McFadden, og som imponerer, er kombinasjonen av teoretisk analyse, evne til å omsette denne til empiriske relasjoner, økonometriske metode, praktiske beregnings- og estimeringsalgoritmer samt grundige empiriske analyser innen en lang rekke felter. ■

Takk

Jeg takker for konstruktive kommentarer fra Leif Andreassen, Jørgen Aasness, Jarle Møen, Torbjørn Hægeland, Jo Thori Lind, Tor Jakob Klette og Kai Leitemo.

Referanser

- Aaberge, R., J.K. Dagsvik, og S. Strøm (1995): Labor supply responses and welfare effects of tax reforms. *Scandinavian Journal of Economics*, **97**, 635-659.
- Anderson, S., A. de Palma, og J-F. Thisse (1992): *Discrete choice theory of product differentiation*. MIT Press.

- Ben-Akiva, M. og T. Watanatada (1981): Application of a continuous spatial choice logit model. I C.F. Manski og D. McFadden (eds.), *Structural analysis of discrete data with econometric applications*. MIT Press.
- Dagsvik, J.K. (1983): Discrete dynamic choice: An extension of the choice models of Thurstone and Luce. *Journal of Mathematical Psychology*, **27**, 1-43.
- Dagsvik, J.K. (1988): Markov chains generated by maximizing components of multivariate extremal processes. *Stochastic Processes and their Applications*, **28**, 31-45.
- Dagsvik, J.K. (1994): Discrete and continuous choice, max-stable processes and independence from irrelevant attributes. *Econometrica*, **62**, 1179-1205.
- Dagsvik, J.K. (1995): How large is the class of generalized extreme value models? *Journal of Mathematical Psychology*, **39**, 90-98.
- Dagsvik, J.K. (1996): Dynamic choice, multistate duration models and stochastic structure. Discussion Papers, no.172, Statistisk sentralbyrå.
- Dagsvik, J.K. (2000a): Aggregation in matching markets. *International Economic Review*, **41**, 27-57.
- Dagsvik, J.K. (2000b): Discrete choice in continuous time: Implications of an intertemporal version of the IIA property. Kommer i *Econometrica*.
- Debreu, G. (1960): Review of Luce, Individual Choice Behavior: A Theoretical Analysis. *American Economic Review*, **50**, 186-188.
- Dubin, J. og D.L. McFadden (1984): An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, **52**, 345-362.
- Falmagne, J-C. (1978): A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, **18**, 52-72.
- Fechner, G.T. (1860): *Elemente der psychophysik*. Vol. I & II. Breitkopf und Hartel, Leipzig.
- Flinn, C. og J.J. Heckman (1982): New methods for analyzing structural models of labor force dynamics. *Journal of Econometrics*, **18**, 115-168.
- Hausman, J., G. Leonard og D.L. McFadden (1995): A utility-consistent, combined discrete choice and count data model: Assessing Recreational use losses due to natural resource damage. *Journal of Public Economics*, **56**, 1-30.
- Heckman, J.J. (1974): Shadow wages, market wages and labor supply. *Econometrica*, **42**, 679-693.
- Heckman, J.J. (1979): Sample selection bias as a specification error. *Econometrica*, **47**, 153-161.
- Heckman, J.J. og T. MaCurdy (1980): A life cycle model of female labor supply. *Review of Economic Studies*, **46**, 47-74.
- Heckman, J.J. (1981): Statistical models for discrete panel data. I C.F.Manski og D. McFadden (eds.), *Structural analysis of discrete data with econometric applications*. MIT Press.
- Heckman, J.J. og B. Singer (1984a): A method of minimizing the impact of distributional assumptions for duration data. *Econometrica*, **52**, 271-320.
- Heckman, J.J. og B. Singer (1984b): The identifiability of the proportional hazard model. *Review of Economic Studies*, **50**, 231-241.
- Heckman, J.J. og G. Sedlacek (1985): Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection. *Journal of Political Economy*, **93**, 1077-1125.
- Heckman, J.J. og V.J. Hotz (1989): Choosing among alternative nonexperimental methods for estimating the impact of social program: The case of manpower training. *Journal of the American Statistical Association*, **84**, 862-874.
- Heckman, J.J. (1991): Identifying the hand of the past: Distinguishing state dependence from heterogeneity. *American Economic Review*, **81**, 75-79.
- Heckman, J.J. (1992): Haavelmo and the birth of modern econometrics: A review of the history of econometric ideas by Mary Morgan. *Journal of Economic Literature*, **30**, 876-886.
- Heckman, J.J., R. LaLonde og J. Smith (1999): The economics and econometrics of active labor market programs. I O. Ashenfelter og D. Card (eds.), *Handbook of Labor Economics*, Vol. 3A. North-Holland.
- Heckman, J.J. (2000a): Policies to foster human capital. *Research in Economics*, **54**, 3-56.
- Heckman, J.J. (2000b): Causal parameters and policy analysis in economics: A 20th century perspective. *Quarterly Journal of Economics*, **115**, 45-97.
- Luce, R.D. (1959): *Individual choice behavior: A theoretical analysis*. Wiley.
- Luce, R.D. og P. Suppes (1965): Preference, utility and subjective probability. I R.D. Luce, R.R. Bush, og E. Galanter (eds.), *Handbook of Mathematical Psychology*, Vol. III, Wiley, pp. 249-410.
- Manski, C.F. og D.L. McFadden (1981): Alternative estimators and sample designs for discrete choice analysis. I C.F. Manski og D.L. McFadden (eds.), *Structural analysis of discrete data with econometric applications*. MIT Press.
- Marschak, J. (1960): Binary-choice constraints and random utility indicators. I K. Arrow, S. Karlin og P. Suppes (eds.), *Mathematical methods in the social sciences*. Stanford University Press.
- McFadden, D.L. (1973): Conditional logit analysis of qualitative choice behavior. I P. Zarembka (eds.), *Frontiers of econometrics*. Academic Press.
- McFadden, D.L. (1978a): Modeling the choice of residential location. I A. Karlqvist, L. Lundqvist, F. Snickars og J. Weibull (eds.), *Spatial interaction theory and planning models*. North-Holland.
- McFadden, D.L. (1978b): Quantitative methods for analyzing travel behaviour of individuals: Some recent developments. I D. Hensher og P. Stopher (eds.), *Behavioral Travel Modelling*. Croom Helm.
- McFadden, D.L. (1981): Econometric models for probabilistic choice. I C. F. Manski og D. L. McFadden (eds.), *Structural analysis of discrete data with econometric applications*. MIT Press.
- McFadden, D.L. (1989): A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, **57**, 995-1026.
- McFadden, D.L. (1994): Demographics, the housing market, and the welfare of the elderly. I D. Wise (eds.), *Studies in the economics of aging*. University of Chicago Press.
- McFadden, D.L. og K. Train (2000): Mixed MNL models for discrete response. *Journal of Applied Econometrics*, **15**, 447-470.
- McFadden, D. L. (2000): Disaggregate behavioral travel demand's RUM side. A 30-year retrospective. Mimeo, Department of Economics, UCB.
- Roy, A. (1951): Some thoughts on the distribution of earnings. *Oxford Economic Papers*, **3**, 135-146.
- Thurstone, L.L. (1927): A law of comparative judgment. *Psychological Review*, **34**, 273-286.
- Fuss M. og D.L. McFadden (1978): *Production economics: A dual approach to theory and applications*, Vol. I and II. North-Holland.
- Quandt, R.A. (1956): A probabilistic theory of consumer behavior. *Quarterly Journal of Economics*, **70**, 507-536.