



Information

Sveriges Riksbanks pris i ekonomisk vetenskap
till Alfred Nobels minne

(populärvetenskaplig information)

Sveriges Riksbanks pris i ekonomisk vetenskap till Alfred Nobels minne år 2000

James Heckman och **Daniel McFadden** har var för sig utvecklat teori och metoder inom området mikroekonometri; dessa har fått stor användning i statistisk analys av individers och hushålls beteende, inom såväl nationalekonomi som andra samhällsvetenskaper.

Mikroekonometri och mikrodata

Mikroekonometrin är ett gränsområde mellan ekonomi och statistik. Det omfattar ekonomisk teori och statistiska metoder som används för att analysera mikrodata, dvs. uppgifter om ekonomiska förhållanden för individer, hushåll och företag. Mikrodata förekommer dels som tvärsnittsdata, dvs. data avseende förhållanden vid en viss tidpunkt, dels som longitudinella data (paneldata), dvs. uppgifter för samma observationsenheter under en följd av år. Mikroekonometrin har haft en enastående stark utveckling under de senaste 30 åren tack vare tillkomsten av stora databaser med mikrodata.

Tillgången till mikrodata och allt kraftfullare datorer har givit helt nya möjligheter att empiriskt pröva mikroekonomisk teori. Forskarna har på individnivå kunnat studera en rad nya frågor: Vilka faktorer bestämmer om en individ väljer att arbeta och, i så fall, hur många timmar? Hur påverkar ekonomiska incitament individuella val av utbildning, yrke eller bostadsort? Vilka effekter på individers inkomster och sysselsättning har olika arbetsmarknads- och utbildningsprogram?

Samtidigt ger användningen av mikrodata också upphov till en rad nya statistiska problem. Problemen bottnar ytterst i begränsningar i dessa (icke-experimentella) data. Forskaren kan endast observera vissa variabler för vissa individer eller hushåll, vilket kan göra att urvalen inte blir slumpmässiga och därmed inte representativa. Även vid representativa urval förblir vissa av de egenskaper som påverkar individernas beteende icke-observerade, vilket gör det svårt eller omöjligt att förklara en del av variationen mellan individer.

Årets ekonomipristagare har, var för sig, visat hur man kan lösa fundamentala statistiska problem associerade med analys av mikrodata. Gemensamt för Heckmans och McFaddens metodologiska insatser är att de har en solid grund i ekonomisk teori och att de tillkommit i ett nära samspel med tillämpade empiriska undersökningar, där nya databaser varit en avgörande förutsättning. Heckmans och McFaddens mikroekonomiska bidrag har bildat skola, inte bara inom nationalekonomin utan också inom andra samhällsvetenskaper.

James J. Heckman

James Heckman har lämnat många viktiga metodologiska bidrag till mikroekonometrin, där den gemensamma nämnaren är selektionsproblem av olika slag. Hans metodologiska bidrag har utvecklats parallellt med empiriska undersökningar, i synnerhet inom arbetsmarknadsekonomi.

Mer information: www.nobel.se/announcement/2000

11 oktober 2000



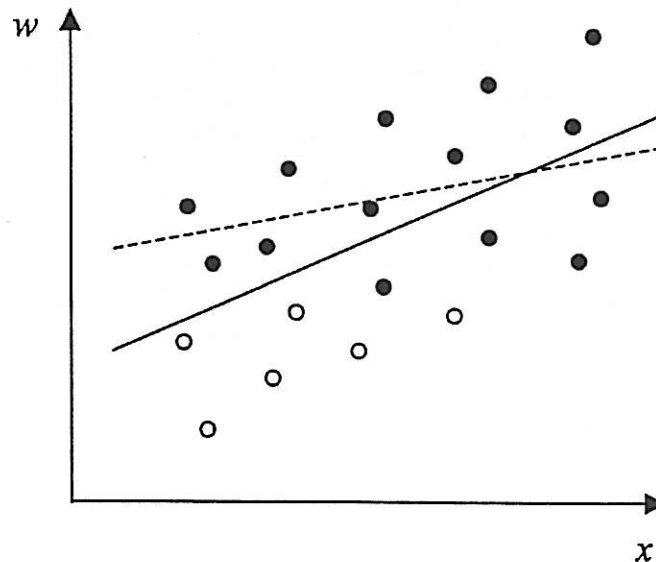
**KUNGL.
VETENSKAPSAKADEMIEN**
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Box 50005, 104 05 Stockholm. Tel: 08-673 95 25, fax: 08-15 56 70, e-post: rsas@kva.se, internet: www.kva.se

Problem vid selektiva urval

Selektiva urval är legio vid mikroekonometriska studier. De uppkommer när forskarens tillgängliga urval av data inte är ett slumpmässigt stickprov av den underliggande populationen. Selektiva urval kan vara resultatet av regler som styr datainsamlingen eller ett resultat av de ekonomiska aktörernas egna beteende. Man talar i det senare fallet om självselektion (*self selection*). Till exempel kan vi bara observera löner och arbetstider för de personer som valt att förvärvsarbeta; uppgifter om löner för universitetsutbildade finns endast för de personer som skaffat sig sådan utbildning. Avsaknaden av information om vilken lön som individen skulle ha haft om han eller hon valt annorlunda skapar statistiska problem vid många empiriska undersökningar.

Problemet med självselektion kan illustreras i figuren nedan, där w anger individens lön och x en faktor som påverkar lönen, t. ex. längden av individens utbildning. Varje punkt i figuren representerar individer med samma utbildnings- och lönenivå i ett stort, representativt stickprov av populationen. Den heldragna linjen visar det statistiska (och sanna) samband vi skulle kunna skatta om vi haft uppgifter om löner och utbildning för samtliga dessa individer. Anta nu – i enlighet med ekonomisk teori – att endast individer med marknadslöner över ett visst tröskelvärde (reservationslönen) väljer att arbeta. Individer med relativt höga löner och relativt lång utbildning kommer då att vara överrepresenterade i urvalet vi faktiskt observerar: de grå punkterna i figuren. Detta selektiva urval ger statistiska problem, eftersom vi kommer att skatta ett samband mellan lön och utbildning som ges av den streckade linjen i figuren. Vi finner alltså ett svagare samband än det sanna och underskattar därför utbildningens effekt på lönen.



Heckmans bidrag

Heckman gjorde sina metodgenombrott rörande självselektion vid 1970-talets mitt. De är nära kopplade till hans studier av hur individerna väljer arbetskraftsdeltagande och arbetad tid. Eftersom vi observerar variationer i arbetad tid endast bland personer som valt att arbeta, riskerar vi att urvalet av personer inte är slumpmässigt. Heckman anvisade en metod för att hantera sådana självselektionsproblem i en artikel från 1974 som behandlade gifta kvinnors utbud av arbetskraft. Arbetet är ett fint exempel på hur mikroekonomisk teori och mikroekonometrisk metod kan kombineras för att belysa ett centralt forskningsområde.

Några år senare föreslog Heckman ännu en metod för att beakta självselektion: den berömda Heckmankorrekturen (tvåstegs-metoden, Heckmans lamda eller Heckitmetoden). Denna metod har fått stort genomslag på grund av att den är enkel att praktiskt tillämpa. Anta att forskaren – som i exemplet ovan – vill skatta ett lönesamband på individdata men endast har löneobservationer för förvärvsarbetande personer. Heckmankorrekturen sker i två steg. Först formulerar forskaren en modell, baserad på ekonomisk teori, för sannolikheten att arbeta. Statistisk skattning av modellen ger resultat som kan användas för att förutsäga denna sannolikhet för varje individ. I ett andra steg korrigerar forskaren för självselektion genom att inkludera de individuella sannolikheterna som en ytterligare förklaringsvariabel vid sidan av utbildning, ålder, etc. Lönesambandet kan då skattas på ett statistiskt tillfredsställande sätt.

Heckmans bidrag har följts av ett mycket stort antal empiriska tillämpningar i nationalekonomi, men även i andra samhällsvetenskaper. Den ursprungliga metoden har sedermera generaliserats, både av Heckman och av andra forskare.

Varaktighetsmodeller

Varaktighets- eller durationsmodeller har en lång tradition inom tekniska och medicinska vetenskaper. De används också flitigt av samhällsvetare, bl. a. av demografer för att studera mortalitet, fertilitet och migration. Ekonomer använder dem t. ex. för att studera hur arbetslöshetstidens längd påverkar sannolikheten att få jobb. Ett problem vid sådana studier är att individer med dåliga arbetsmarknadsutsikter kan vara överrepresenterade bland dem som blir kvar i arbetslöshet. Denna urvalsselektion ger problem liknande dem vid självselektion: då urvalet arbetslösa vid en viss tidpunkt påverkas av icke-observerade individuella egenskaper kan vi få missvisande skattningar av arbetslöshetens "varaktighetsberoende". I samarbete med Burton Singer har Heckman anvisat ekonometriska metoder för att komma till rätta med dylika problem. Dessa har fått stor användning bland samhällsvetare.

Utvärdering av arbetsmarknadspolitik

När aktiv arbetsmarknadspolitik – som arbetsmarknadsutbildning eller anställningssubventioner – blivit allt vanligare i en rad länder har också behovet vuxit att utvärdera dessa program. Det klassiska problemet är att avgöra hur ett program påverkar deltagarnas inkomster eller sysselsättning jämfört med om de *inte* deltagit. Men då man inte kan observera samma individ i två roller samtidigt måste man använda information om icke-deltagare, vilket kan ge selektionsproblem av ovan diskuterat slag. Heckman är världens främste forskare på området mikroekonomisk utvärdering av arbetsmarknadsprogram. Tillsammans med olika medarbetare har han utförligt analyserat egenskaperna hos alternativa icke-experimentella metoder för utvärdering och utrett hur de förhåller sig till experimentella metoder. Heckman har också bidragit med egna empiriska undersökningsresultat. Studierna visar ofta att programmen har mycket små positiva – ibland negativa – effekter för deltagarna och att de ofta inte är samhällsekonomiskt lönsamma. Utfallen varierar dock starkt mellan olika deltagare och olika program.

Daniel L. McFadden

McFaddens viktigaste bidrag är att ha utvecklat ekonomisk teori och mikroekonomiska metoder för analys av sk diskreta val, dvs val mellan ett ändligt antal beslutsalternativ. Ett genomgående drag i McFaddens arbeten är samspelet mellan ekonomisk teori, statistisk metod och empiriska tillämpningar, där den yttersta drivkraften ofta är en önskan att lösa angelägna samhällsproblem.

Diskreta val

Mikrodata avspeglar ofta diskreta val. Uppgifter i en databas om individers yrke, bostadsort eller resätt, reflekterar sålunda de val dessa individer gjort bland ett begränsat antal alternativ. Den traditionella

efterfrågeanalysen i ekonomisk teori förutsätter att individens val representeras av en kontinuerlig variabel och kan därför inte användas för att studera diskreta val. Innan McFaddens prisbelönta insatser saknade empiriska studier av sådana beslut en underbyggnad i ekonomisk teori.

McFaddens bidrag

McFaddens teori för diskreta val utgår från mikroekonomisk teori, enligt vilken varje individ väljer det tillgängliga alternativ som maximerar nyttan. Forskaren observerar dock inte *alla* de faktorer som påverkar de individuella valen och tycker sig därför se en slumpmässig variation bland individer med samma observerade egenskaper. På basis av sin nya teori har McFadden utvecklat mikroekonometriska modeller som bl. a. kan användas för att förutsäga hur stor andel av en viss population som kommer att välja olika alternativ.

McFaddens viktigaste bidrag är hans utveckling av den s k betingade logitmodellen (*conditional logit analysis*) från 1974. För att beskriva modellen antar vi att individen har att välja mellan ett antal (säg J) olika alternativ. Låt X representera de egenskaper hos alternativen och Z de egenskaper hos individerna som forskaren kan observera i sina data. I en studie av resväg t. ex. kan de olika alternativen vara bil, buss och tunnelbana, X kan inkludera uppgifter om resornas pris och tidsåtgång, medan Z kan innefatta uppgifter om ålder, inkomst och utbildning. Utöver X och Z finns det emellertid ytterligare egenskaper hos individer och valalternativ som styr individens nyttomaximerande val, men som forskaren inte kan observera. Dessa summeras i en "slumterm", som antas ha en viss statistisk fördelning (extremvärdesfördelning) i populationen. Under dessa förutsättningar (och några tekniska antaganden) visar McFadden att sannolikheten för att individ i ska välja alternativ j kan skrivas på följande matematiska form:

$$P_{ij} = \frac{e^{X_{ij}\beta + Z_i\delta_j}}{\sum_{k=1}^J e^{X_{ik}\beta + Z_i\delta_k}}$$

I denna s k multinomiala logitmodell är e basen för den naturliga logaritmen och β respektive δ (vektorer av) parametrar. Eftersom forskaren i sin databas kan observera såväl variablerna X och Z samt det alternativ individen faktiskt väljer, kan han skatta parametrarna β och δ med kända statistiska metoder. Även om logitmodeller var kända sedan tidigare var McFaddens härledning helt ny och hans bidrag erkändes omedelbart som ett fundamentalt genombrott.

Modeller av detta slag är mycket användbara. De används t. ex. regelmässigt i studier av individers val av resväg. Man kan sålunda utnyttja modellerna i trafikplanering för att studera effekterna av åtgärder i trafiksystem eller andra omvärldsförändringar: Hur kommer prisförändringar, bättre tillgänglighet, eller förändringar i befolkningens demografiska sammansättning att påverka andelen resande med olika kommunikationsmedel? Men modellerna har också använts på en rad andra områden, bl. a. vid studier av val av bostad, bostadsort och utbildning. McFadden själv har tillämpat sina metoder för att belysa ett antal samhällsfrågor, som efterfrågan på energi till bostäder, på telefonservice och på äldreboenden.

Metodologisk vidareutveckling

En speciell egenskap hos modellen ovan är att de relativa valsannolikheterna för två alternativ, säg resa med buss eller bil, är oberoende av pris och kvalitet på andra transportalternativ. Denna egenskap, som benämns IIA (*independence of irrelevant alternatives*), är orealistisk vid många tillämpningar. McFadden har dels utvecklat statistiska test för att avgöra om IIA är uppfyllt, dels presenterat mer generella modeller. Ett exempel på en mer generell modell är den s k nestade logitmodellen, där

individernas val antas kunna ordnas i en bestämd sekvens. I en studie av val av bostadsort och typ av bostad t. ex. kan man anta att individerna först bestämmer sig för bostadsort och därefter väljer mellan tillgängliga bostadsalternativ.

Även efter dessa generaliseringar är modellerna känsliga för specifika antaganden om fördelningen av icke-observerade egenskaper i populationen. McFadden har under de senaste tio åren anvisat simuleringsmetoder (*method of simulated moments*) för statistisk skattning av diskreta valmodeller som tål betydligt generellare antaganden. Med allt kraftfullare datorer har dessa numeriska metoder nu blivit praktiskt användbara. Därmed kan individernas diskreta val beskrivas med större realism och deras beslut förutsägas med större precision.

Andra arbeten

McFadden har också gjort inflytelserika arbeten inom andra områden. Under 1960-talet anvisade han metoder för att empiriskt kvantifiera produktionstekniska samband och undersöka faktorerna bakom företagets efterfrågan på kapital och arbetskraft. Under 1990-talet har McFadden lämnat viktiga bidrag till miljöekonomisk forskning, bl. a. när det gäller metoder för att värdera miljöresurser. Ett exempel är en studie av välfärdsförlusterna som följde på miljöskadorna av utsläppen vid Alaskas kust från oljetankern Exxon Valdez år 1989. Denna studie ger exempel på McFaddens utomordentliga förmåga att knyta ihop mikroekonomisk teori och ekonometrisk metodologi i empiriska studier av viktiga samhällsproblem.

James J. Heckman föddes i Chicago år 1944 och bedrev universitetsstudier vid Colorado College med matematik som huvudämne. Han kom sedan till Princeton University för forskarutbildning i nationalekonomi och disputerade där år 1971. Heckman har därefter varit professor vid Columbia och Yale. Sedan 1995 är han Henry Schultz Distinguished Service Professor of Economics vid University of Chicago.

James Heckman
Department of Economics
The University of Chicago
1126 East 59th Street
Chicago, IL 60637
USA
<http://lily.src.uchicago.edu>

Daniel L. McFadden föddes 1937 i Raleigh, North Carolina, och fick sin grundläggande universitetsutbildning vid University of Minnesota med fysik som huvudämne. Han började sedan studera nationalekonomi vid samma universitet och disputerade år 1962. McFadden har varit professor vid University of Pittsburgh, Yale och MIT. Sedan 1990 är han E. Morris Cox Professor of Economics vid University of California, Berkeley.

Daniel McFadden
Economics Department
University of California at Berkeley
Berkeley, CA 94720
USA
<http://emlab.berkeley.edu/users/mcfadden/index.html>

Läs vidare

Advanced information on the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel, 2000, Kungl. Vetenskapsakademien, <http://www.nobel.se/announcement/2000>

Amemiya T. (1987), Discrete Choice Analysis, in P. Newman, M. Milgate and J. Eatwell (eds.), *The New Palgrave – A Dictionary of Economics*, Macmillan.

Heckman J.J. (1987), Selection Bias and Self-Selection, in P. Newman, M. Milgate and J. Eatwell (eds.), *The New Palgrave – A Dictionary of Economics*, Macmillan.

Heckman J. J. and J. Smith (1995), Assessing the Case for Social Experiments, *Journal of Economic Perspectives* 9, 85-110.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

McFadden D. L. (2000), Disaggregate Travel Demand's RUM Side: A 30-Year Retrospective, manuscript, Department of Economics, University of California, Berkeley
<http://emlab.berkeley.edu/users/mcfadden/index.html>.

Small, K. A. (1992), *Urban Transportation Economics*, Fundamentals of Pure and Applied Economics, Harwood Academic Publishers.