

# Premi Nobel d'Economia de l'any 2000: Una elecció encertada entre una mostra seleccionada

Jaume Garcia i Villar<sup>54</sup>

## Abstract

*The 2000 Nobel Prize in Economics was awarded to professors Heckman and McFadden for their methodological contributions in the microeconomic field, an approach concerned with econometric problems in the context of the use of information on individuals (persons, family units, business firms). The contributions of James Heckman refer fundamentally to the implications of the use of selected samples to obtain correct conclusions which can be extrapolated to the general population, a question especially relevant to analysing models of the employment market and evaluating social programmes. The contributions of Daniel McFadden have centred on the development of models of discreet choice, their fit with the real world, their capacity to estimate and consistency as an explanatory model of the behaviour of the agents. Their contributions have important implications for the evaluation of how several economic questions affect demand for transport and residential location among others.*

## Resum

*El Premi Nobel d'Economia de l'any 2000 ha estat atorgat als professors Heckman i McFadden per les seves contribucions metodològiques en el camp de la microeconometria, disciplina que s'ocupa dels problemes econòmics relacionats amb la utilització de dades individuals (persones, unitats familiars, empreses). Les contribucions de James Heckman s'han centrat fonamentalment en les implicacions de l'utilització de mostres seleccionades per tal de extreure conclusions correctes extrapolables a la població, tema d'especial rellevància en l'anàlisi de models del mercat laboral i per a l'avaluació de programes socials. Les contribucions de Daniel McFadden s'han centrat en el desenvolupament del models d'elecció discreta tant en termes de la seva adequació a la realitat i la seva facilitat d'estimació com en termes de la seva consistència amb un model explicatiu del comportament dels agents. Les seves contribucions han tingut implicacions importants a l'hora de valorar com diferents aspectes econòmics o relacionats amb les alternatives disponibles afecten la demanda de transport o la localització de residència entre d'altres temes*

## Introducció

“A James Heckman per les seves contribucions teòriques i metodològiques en l'anàlisi de mostres seleccionades i a Daniel McFadden per les seves contribucions teòriques i metodològi-

54. Catedràtic d'economia aplicada de la Universitat Pompeu Fabra i conseller acadèmic de l'Escola Superior de Comerç Internacional.

ques en l'anàlisi d'eleccions discretes." Amb aquesta cita, la Reial Acadèmia Sueca de les Ciències feia saber la concessió del Premi Nobel d'Economia de l'any 2000, expressant d'aquesta manera el seu reconeixement als guardonats per les seves aportacions teòriques i metodològiques en el camp de la Microeconometria. Ambdós guardonats no tenen una formació acadèmica inicial en economia, Heckman és B.A. en Matemàtiques pel Colorado College, mentre que McFadden és B.S. en Física per la Universitat de Minnesota, encara que tots dos són doctors en Economia, Heckman per la Universitat de Princeton l'any 1971 i McFadden per la Universitat de Minnesota l'any 1962. El paral·lelisme de les seves trajectòries també s'ha traduït en el desenvolupament de carreres acadèmiques molt arrelades a una única universitat, amb un període d'interrupció per marxar a una altra institució. Heckman és Henry Schultz Distinguished Professor d'Economia a la Universitat de Chicago, on ha desenvolupat tota la seva carrera acadèmica amb un breu parèntesi a la Universitat de Yale (1988-1990), mentre que McFadden és actualment E. Morris Cox Professor d'Economia a la Universitat de Berkeley, amb un parèntesi, més llarg que en el cas de Heckman, a l'Institut Tecnològic de Massachusetts (1978-1991). Tots dos guardonats formen part del grup d'investigadors que en els darrers anys han contribuït a consolidar una disciplina dintre del camp més genèric de l'Econometria que ara té una entitat pròpia: la Microeconometria.

## Què és la microeconometria?

Entre els aspectes que han marcat el desenvolupament de l'econometria com a disciplina des dels seus inicis, tant oficials com oficials<sup>55</sup>, podem destacar-ne dos: el tipus d'informació a la qual es pot accedir i, per tant, utilitzar en el treball empíric, i la disponibilitat de mitjans de càlcul que permetin processar aquesta informació. En aquest sentit, en la dècada dels setanta i començament dels vuitanta assistim a una creixent accessibilitat a bases de dades provinents d'enquestes, en les quals les observacions corresponen a individus en sentit ampli, ja siguin persones físiques, unitats familiars o empreses<sup>56</sup>, les quals presenten unes singularitats marcadament diferents a les corresponents a les dades temporals. Així, podem destacar un major pes de la informació de tipus qualitatiu (per exemple, estat civil de les persones, comunitat autònoma de residència, sector d'activitat econòmica, tipus d'ocupació, tipus de transport emprat per desplaçar-se de casa a la feina, la marca comprada per a un producte concret), una major rellevància dels aspectes lligats a la manera com s'ha triat la mostra de les enquestes (normalment els procediments de mostreig acostumen a ser estratificats o estan referits a grups concrets no representatius de la població), una major freqüència de situacions de no-observabilitat d'alguna variable rellevant per a una part de la mostra o per a la seva totalitat (el salari que rebria un individu al sector privat si està treballant en el sector públic, els gustos o les preferències dels individus), o la major adequació d'aquesta informació per analitzar el comportament que els models teòrics de tipus microeconòmic postulen per als agents econòmics.

55. La *Econometric Society* es crea l'any 1931, però alguns historiadors de l'Econometria (Morgan, 1990) ja troben indicis d'aquesta disciplina en les opinions d'alguns economistes de finals del segle XIX com Jevons, que encaixen perfectament en els plantejaments que expressa Frisch (1933) en l'editorial del primer volum de la revista *Econometrica*.

56. Exemples d'aquest tipus d'informació per al cas espanyol serien, entre d'altres: la *Encuesta de Población Activa*, la *Encuesta de Presupuestos Familiares*, la *Encuesta Industrial*, la *Central de Balances*, la *Encuesta de Estructura Salarial* o la *Encuesta Nacional de Salud*.

La microeconometria centra l'atenció en tots els aspectes de tipus estadístic però també de tipus metodològic (especificació) relacionats amb els temes anteriorment esmentats, que són propis de l'anàlisi econòmica aplicada, encara que alguns d'ells també siguin rellevants en altres disciplines, sobretot de l'àmbit de les ciències socials.

Serveixi com a exemple l'anàlisi de la discriminació salarial per motius de gènere. L'ús de dades individuals permet comparar, mitjançant l'anàlisi de regressió, quina part de la diferència observada entre les mitjanes dels salaris d'homes i dones és deguda a les diferents característiques (variables explicatives) d'aquestes dues poblacions i quina part es deu a la diferent retribució de les mateixes (coeficients), aquesta última la que s'associa a la discriminació. L'especificació d'aquestes equacions d'ingressos acostuma a tenir una forma funcional de tipus semilogarítmic derivada de models senzills basats en la teoria del capital humà, i inclou com a variables explicatives les típiques d'aquest plantejament, entre les quals n'hi ha algunes de tipus qualitatiu, com per exemple l'educació, i altres que també són qualitatives i que fan referència a característiques del lloc de treball o al sector d'activitat econòmica. D'altra banda, la mostra d'individus que treballen, que són aquells per als quals s'observa el salari, no té per què ser representativa de la població, sobretot en el cas de les dones amb una menor participació, cosa que fa que les mostres que es poden utilitzar siguin mostres seleccionades, aspecte que s'ha de tenir en compte a l'hora d'estimar consistentment els paràmetres d'aquests models. Fins i tot recentment s'ha generalitzat l'anàlisi procedint a comparar no solament les mitjanes per a ambdues poblacions, sinó tota la distribució, a través d'un plantejament més flexible com és la regressió quantílica, que permet que els nivells de discriminació puguin ser diferents segons els punts de la distribució de salaris per a homes i dones que comparem.

### *Els models d'elecció discreta*

La concessió del Premi Nobel a Daniel McFadden està basada en les seves contribucions metodològiques a l'anàlisi empírica de models que miren d'explicar decisions dels individus: models d'elecció discreta. En aquest tipus de model la variable endògena és una variable qualitativa que fa referència a una decisió o a una situació. Per exemple, en l'anàlisi de la demanda de transport la variable a explicar acostuma a ser quin és el tipus de transport emprat pels individus per fer els seus desplaçaments (cotxe, tren, autobús, a peu). És a dir, la variable endògena és una variable qualitativa que fa que l'anàlisi estàndard del model de regressió no sigui aplicable.

Ja des dels anys cinquanta s'han desenvolupat models de tipus probabilístic per analitzar els determinants de les probabilitats de triar una alternativa o una altra. És el cas del model lògit, popularitzat per Berkson (1951) i del model pròbit (Aitchison i Silvey, 1957), que difereixen en la forma de la distribució usada per calcular les probabilitats associades a cada alternativa. La principal contribució de McFadden és el conegut model lògit condicional (McFadden, 1973), no tant perquè es tracta d'una generalització del model lògit multinomial, ja desenvolupat en la bibliografia en aquells moments, sinó perquè la seva derivació està basada en un model de comportament de maximització de la utilitat per part dels individus. Davant l'elecció per part d'un individu "i" entre J alternatives, McFadden suposa que la utilitat associada a cada alternativa depèn de les característiques d'aquest individu (X), de les característiques de l'alternativa, que poden ser diferents per a cada individu (Z), i d'un conjunt de característiques inobservables que representen les preferències o gustos, escollint aquella alternativa que té la utilitat més alta. Per exemple, en el

cas de la demanda de transport en el vector de característiques X s'inclourien variables com els ingressos de l'individu, la zona on resideix o la seva posició en la unitat familiar, entre d'altres, i en el vector Z s'inclourien variables com el temps emprat en el desplaçament o el cost del desplaçament, totes dependents de la modalitat de transport triada<sup>57</sup>.

Suposant que la utilitat és lineal en les característiques X, Z i els termes aleatoris no observables, els quals segueixen una distribució *extreme-value Type I* independent i idèntica per a cada alternativa, la probabilitat que l'individu "i" triï l'alternativa "j" ( $P_{i,j}$ ) serà:

$$P_{i,j} = \frac{e^{x_i'\beta_j + Z_i'\alpha}}{\sum_{k=1}^J e^{x_i'\beta_k + Z_i'\alpha}}$$

on  $\beta_j$  és el vector de coeficients de les variables corresponents a les característiques dels individus (X) i  $\alpha$  el vector de coeficients corresponents a les característiques de les alternatives (Z). La forma d'aquestes probabilitats comporta una propietat per a aquest model d'importants implicacions empíriques: la independència de les alternatives irrelevantes. Sobre la base de l'expressió anterior, la relació entre les probabilitats de triar dues alternatives "j" i "k" és la mateixa, sigui quin sigui el conjunt de les alternatives entre les quals es pot escollir, tal com es pot observar en l'expressió següent:

$$\frac{P_{i,j}}{P_{i,k}} = \frac{e^{x_i'\beta_j + Z_i'\alpha}}{e^{x_i'\beta_k + Z_i'\alpha}}$$

Aquesta propietat és poc realista en situacions on algunes de les alternatives són gairebé substitutives, atès que implica que la probabilitat d'escollir una alternativa es veu modificada en la mateixa proporció per a totes les alternatives per l'aparició d'una nova alternativa o la desaparició d'una de les existents. L'exemple habitualment emprat per il·lustrar aquesta propietat és el referit al cas del transport. Suposem que la probabilitat d'escollir d'anar amb cotxe de casa a la feina és 2/3 i que la d'anar amb autobús (vermell) és 1/3, és a dir, el quocient entre aquestes dues probabilitats serà 2. Si l'autoritat responsable decideix pintar de blau la meitat dels autobusos i entenem que ara hi ha tres alternatives disponibles (cotxe, autobús vermell i autobús blau), segons el model lògit (tant en la seva versió multinomial com condicional), suposant que les probabilitats d'anar amb autobús d'un o altre color són iguals, en el cas d'un model lògit la probabilitat d'anar amb cotxe passarà a ser 1/2 i les corresponents a anar amb autobús seran en tots dos casos 1/4. Aquest és un resultat contraintuïtiu, ja que esperaríem que la d'anar amb cotxe continués essent 2/3 i les d'anar amb autobús 1/6 en cada cas. Les prediccions del model lògit estan *forçades* per la propietat d'independència de les alternatives irrelevantes que es deriva de la forma de les probabilitats d'aquest model.

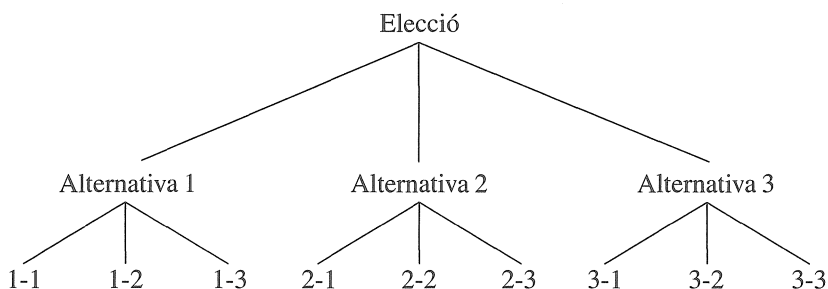
57. Cal destacar que el terme que recull els factors inobservables en el context d'aquest model té el paper de garantir que no necessàriament individus amb idèntiques característiques, tant personals com referides a l'alternativa, siguin observats fent la mateixa elecció.

Aquest és i ha estat un model àmpliament usat en la bibliografia per la seva facilitat de còmput enfront d'altres models alternatius, com el model pròbit, quan el nombre d'alternatives és elevat.

McFadden ha fet contribucions posteriors rellevants en aquest àmbit, la majoria dedicades a superar les limitacions del model lògit ja esmentades. Així, en el seu treball de l'any 1973 ja proposa una adaptació del coeficient de determinació del model de regressió per mesurar la capacitat explicativa d'aquests models d'elecció discreta basada en la comparació del valor de la funció de versemblança del model estimat amb la del model que només conté un terme constant per a cada alternativa. En treballs posteriors proposa tests d'especificació per a l'omissió de variables rellevants basats en el test dels multiplicadors de Lagrange però calculat sobre la base d'una regressió auxiliar (McFadden, 1987).

Per superar les limitacions del model lògit, en qualsevol de les seves dues versions condicional o multinomial, McFadden (1978) proposa generalitzar el model lògit mitjançant l'especificació d'una distribució per als termes aleatoris d'aquest model, que superi, en concret, la restrictiva independència d'aquests termes entre alternatives, causa fonamental de la propietat de la independència de les alternatives irrelevantes. Proposa d'usar una funció de distribució, la *generalized extreme value*, que inclou com a cas particular la distribució del model lògit, essent possible contrastar la hipòtesi nul·la de l'adequació del model lògit mitjançant un contrast d'hipòtesi referit a un paràmetre. Aquest model té una interpretació més senzilla en termes del que es coneix com el model lògit niat.

El model implica que les diferents alternatives es poden agrupar de manera natural segons el grau de similitud, tal com queda reflectit en el gràfic següent.



Les nou possibles alternatives estan distribuïdes en tres grups (alternatives 1, 2 i 3), de tal manera que aquesta especificació implica que dintre de cada grup el model que aplica és el logit, mentre que entre les tres grans alternatives la formulació trenca la independència de les alternatives irrelevantes. Així, si desapareix una de les alternatives, per exemple la 1-1, això no modificarà la relació entre les probabilitats de 1-2 i 1-3, però sí que pot modificar les seves relacions amb les probabilitats d'altres alternatives.

Com que aquests models niats són difícils d'estimar quan el nombre d'alternatives és gran, i com que en alguns casos la manera d'agrupar les alternatives no és única ni en la mateixa seqüència, s'han desenvolupat contrastos de la hipòtesi d'independència de les alternatives irrelevantes,

basats únicament i exclusivament en l'estimació del model sota la hipòtesi nul·la, és a dir, el model lògit (multinomial o condicional). McFadden ha contribuït també en aquesta línia de recerca. McFadden (1984) es proposa d'usar un test dels multiplicadors de Lagrange per contrastar aquesta hipòtesi, mentre que en Hausman i McFadden (1984) es procedeix a contrastar aquesta hipòtesi comparant les estimacions dels vectors de paràmetres per a un subconjunt d'alternatives en el cas d'usar tota la mostra i en el cas de fer servir la submostra d'observacions que han escollit aquestes alternatives<sup>58</sup>, mitjançant una versió adaptada del mètode general de contrastació de Hausman (1978) basat en la comparació de dos estimadors consistents sota la hipòtesi nul·la, però que en canvi sota l'alternativa només un d'ells preserva la consistència<sup>59</sup>. Posteriorment, McFadden (1987) mostra com es poden obtenir versions asimptòticament equivalents d'aquests dos tests mitjançant la significació de regressions auxiliars.

Finalment, cal destacar la contribució de McFadden a l'estimació de models d'elecció discreta en general, no únicament els models tipus logit, que presenten elevades dificultats de càlcul a l'hora d'avaluar les probabilitats corresponents. Aquest és el cas del model pròbit multinomial, proposat per Hausman i Wise (1978) com a alternativa al model lògit multinomial, però que comporta avaluar, per procediments numèrics, integrals corresponents a probabilitats d'una distribució normal multivariant d'ordre  $J-1$ , cosa que fa que el càlcul de les mateixes sigui factible fins a quatre alternatives però que ja no sigui factible per a un nombre elevat d'alternatives.

McFadden (1989) proposa el mètode d'estimació per simulació<sup>60</sup> basat en la idea de generar replicacions dels termes aleatoris inobservables i construir les probabilitats ajustades per a cada alternativa (en funció dels paràmetres a estimar) com la mitjana de les decisions (1 si l'elecció és la de l'alternativa en qüestió, 0 en cas contrari) que correspondrien a aquestes replicacions, per tal de procedir a usar-les en el context d'un estimador de moments basat en la no-correlació entre els errors (diferència entre la decisió i la probabilitat d'escollir una alternativa) i un conjunt d'instruments<sup>61</sup>.

## Estimació amb mostres seleccionades

La concessió del Premi Nobel a James Heckman està basada en les seves contribucions metodològiques a l'anàlisi empírica de models amb mostres seleccionades. Suposem que volguéssim estimar la mitjana dels ingressos de les famílies catalanes. Si seleccionem una mostra entre aquelles famílies en les quals el cognom del cap de família comença per A no sembla que la manera de seleccionar la mostra hagi d'afectar l'estimació de l'esmentat paràmetre. Donat l'objectiu que es

58. La idea implícita d'aquest contrast és que si les alternatives són independents en el sentit comentat, llavors la probabilitat de triar una alternativa entre  $J$  d'elles és igual a la probabilitat de escollir-la entre totes les  $J$  alternatives condicionat a estar entre les  $J$ .

59. Aquest test a nivell empíric presenta ocasionalment problemes perquè la diferència entre la matriu de variàncies i covariàncies del model estimat amb el subconjunt de les alternatives i la corresponent al model amb totes les alternatives no és sempre definida positiva, com es desprèn del marc teòric.

60. En McFadden i Ruud (1994) es troba una presentació detallada d'aquest mètode d'estimació en un context més ampli que el dels models d'elecció discreta.

61. Aquest procediment proposat per McFadden supera les limitacions de la proposta de Lerman i Manski (1981) consistent a utilitzar aquestes probabilitats simulades a la funció de versemblança i maximitzar-la, però la consistència del procediment requereix que el nombre de replicacions tendeixi a infinit amb la mida de la mostra.

pretén, la mostra serà representativa de la població. En canvi, si escollíssim la mostra entre aquelles famílies en les quals el cap de família té estudis superiors, la mostra estaria seleccionada d'acord amb un criteri relacionat amb la variable per a la qual volem estimar la mitjana i, consegüentment, sobreestimaríem la mitjana dels ingressos de la població a partir d'aquesta informació.

En essència, el model de regressió pretén estimar la mitjana de la variable endògena condicionada al conjunt de variables explicatives. Un dels supòsits habituals en el model de regressió és que la mostra amb què es pretén estimar el model és una mostra representativa de la població o, amb altres paraules, que si el model està correctament especificat, l'esperança del terme de pertorbació condicionada a la mostra amb la qual s'estima el model és zero. En canvi, si la mostra està seleccionada amb un criteri relacionat amb el terme de pertorbació, l'esperança del terme de pertorbació condicionat a la mostra deixa de ser zero i, per tant, els mínims quadrats ordinaris (MQO) són inconsistents. Heckman (1976) analitza aquest tema en el context d'un model més general que engloba altres models amb característiques similars al model de selecció de la mostra.

Suposem que pretenem estimar un model d'oferta de treball de característiques similars a l'estimat per Heckman (1974), on l'equació d'hores ( $h$ ) té l'especificació següent, atès que per als individus que no treballen  $h = 0$ ,

$$\begin{aligned} h_i &= X_i'\beta + \gamma w_i + u_i & \text{si } X_i'\beta + \gamma w_i + u_i > 0 \\ &= 0 & \text{si } X_i'\beta + \gamma w_i + u_i \leq 0 \end{aligned}$$

on  $X_i$  és un vector de variables explicatives,  $w_i$  és el salari net de l'individu i la variable  $u_i$  correspon al terme de pertorbació. L'estructura d'aquest model és la d'un model tipus Tobit<sup>62</sup>, on els salaris no són observables per a aquells que no treballen ( $h_i = 0$ ), per la qual cosa és necessari especificar una equació de salaris com la següent per completar el model.

$$w_i = Z_i'\delta + \varepsilon_i$$

on  $Z_i$  és un vector de característiques explicatives dels salaris i  $\varepsilon_i$  és el corresponent terme de pertorbació.

Heckman (1974) estima conjuntament per màxima versemblança les dues equacions, tenint en compte els problemes de selecció de la mostra de l'equació de salaris i també la potencial endogeneïtat dels salaris a l'equació d'oferta de treball, però en el seu article de 1976 proposa un procediment basat a utilitzar MQO per estimar l'equació de salaris. En concret la funció de regressió seria:

$$E(w_i | h_i > 0) = Z_i'\delta + E(\varepsilon_i | h_i > 0)$$

62. Aquest model va ser proposat per Tobin, Premi Nobel de l'any 1981 com a procediment per estimar equacions de demanda amb dades individuals en presència de zeros (Tobin, 1958).

la qual es pot expressar en termes de la forma reduïda de l'equació d'oferta de treball<sup>63</sup> com

$$E(w_i | e_i > -X_i^{**}\beta^*) = Z_i'\delta + E(\varepsilon_i | e_i > -X_i^{**}\beta^*)$$

on queda de manifest que l'esperança del terme de perturbació no és zero en la mesura que  $\varepsilon_i$  i  $e_i$  (el terme de perturbació de la forma reduïda de l'equació d'hores) no estiguin relacionades. De fet, sota el supòsit de normalitat per a les perturbacions, la funció de regressió seria

$$E(w_i | e_i > -X_i^{**}\beta^*) = Z_i'\delta + \sigma_{\varepsilon\varepsilon} \lambda_i$$

on  $\lambda$  és el conegut "terme de correcció de Heckman". Aquesta equació es pot estimar per MQO si disposem d'informació per a totes les observacions de totes les variables explicatives llevat del salari, en cas contrari s'haurà d'aplicar el mètode de la màxima versemblança. Heckman (1979) interpreta el biaix en l'estimació de  $\delta$  com un problema d'omissió de variables rellevants proposant un procediment bietàpic per estimar consistentment els paràmetres d'aquesta equació<sup>64</sup>, procediment que s'ha estès també a models de característiques similars al de selecció, com és el model de *switching* (Lee, 1978).

Aquest model de selecció de la mostra neix fonamentalment com a solució a un problema d'inobservabilitat, però posteriorment ha estat usat com una de les possibles generalitzacions del model Tobit, un model en què la variable endògena pren el valor zero per a un percentatge significatiu d'observacions. L'equació d'oferta de treball anteriorment considerada és un exemple d'aquest tipus de model, en el qual la variable endògena incorpora a la vegada una decisió (o una situació), en aquest cas participar en el mercat de treball, i una característica quantificable en cas de treballar: el nombre d'hores treballades.

La principal limitació del model Tobit<sup>65</sup> és que el mateix model explica tant la decisió com els valors positius de la variable endògena, en el cas considerat la situació de participar en el mercat de treball i les hores treballades, essent molts els exemples, inclòs aquest referit a l'oferta de treball, per als quals la restricció que comporta una especificació tipus Tobit no és realista, entre altres raons perquè els zeros poden ser causats per diferents tipus de raons.

Les dues generalitzacions del model Tobit més utilitzades en la bibliografia han estat el model de doble tanca, proposat per Cragg (1971) i el model de selecció de Heckman, entès no com un model per resoldre un problema d'inobservabilitat, sinó com un model que contempla especificacions diferents per a l'equació que explica la decisió i la que explica la quantitat per a valors positius.

63. El criteri de selecció expressat en termes de la forma reduïda és equivalent al del salari de reserva per sota del qual l'individu no treballa.

64. En García (1991) es realitza un estudi de simulació per valorar els biaixos de diferents mètodes d'estimació aplicats a un model d'oferta de treball que també inclou una equació de salaris. Els biaixos són importants, fins al 30 %, sobretot quan el percentatge d'observacions per a les quals s'observa el salari és petit.

65. L'altre tipus de limitació del model Tobit és la distribucional com a conseqüència de suposar la normalitat i l'homoscedasticitat de les perturbacions, l'incompliment de les quals comporta la inconsistència de les estimacions per màxima versemblança.



La metodologia d'anàlisi de mostres seleccionades és aplicable a un àmbit amb rellevància des d'un punt de vista social: l'avaluació dels programes socials (per exemple, de benestar o de formació). En aquest camp les contribucions de Heckman també han estat importants i han permès entendre millor què estem avaluant i com ho estem avaluant. Heckman (1990) presenta una anàlisi comparativa de les maneres d'avaluar els programes socials, ja siguin resultat d'experiments socials o no, utilitzades en la bibliografia, basades fonamentalment en el càlcul de l'efecte mitjana<sup>66</sup>. Però la qüestió, tal com apunta Heckman, és quin és l'efecte mitjana que volem mesurar. El més habitual ha estat valorar com variaria el resultat del "tractament" (per exemple, el nivell d'ingressos) si una persona aleatòriament escollida seguís un programa social (per exemple, de formació). En el cas més senzill de suposar que modelitzem aquest efecte mitjançant una variable fictícia, l'únic problema economètric seria controlar el possible biaix generat pel procés selecció dels participants en el programa en termes de la seva possible correlació amb el terme d'error. En canvi, sembla més interessant quantificar l'efecte mitjà per a aquells qui han decidit participar en el programa, és a dir, condicionat a participar en el programa. Això complica substancialment el procés d'estimació<sup>67</sup>, atès que el model resultant és un model amb paràmetres d'interès no fixos i correlacionats amb el terme d'error corresponent. Finalment, cal destacar que la identificació dels paràmetres d'interès en aquests models també és un tema rellevant que està tractat entre d'altres en Heckman i Honoré (1990).

A l'últim, encara que estrictament no correspon a la metodologia anteriorment considerada, relacionada amb els problemes d'estimació en mostres seleccionades, cal destacar les aportacions fetes per Heckman en l'anàlisi dels models de durada. És a dir, models que miren d'explicar el temps que un individu passa en un determinat estat, o alternativament, la probabilitat d'abandonar un estat després d'haver-hi estat un determinat temps. En la bibliografia econòmica aquest tipus de models s'han aplicat tradicionalment a l'anàlisi de la durada dels episodis d'atur o a la del episodis de vaga.

En aquests models la probabilitat d'abandonar un estat després d'haver-hi estat un determinat temps s'aproxima per la funció de *sortida* (*hazard function*), que té una forma funcional lligada a la forma de la funció de distribució de la durada. En aquest context, i donat el tipus de variable (durada), la distribució normal no ha estat usada, encara que sí la lognormal o unes altres de més habituals en aquest àmbit, com l'exponencial o la Weibull. La durada (o la funció de sortida) es fa dependre d'un conjunt de característiques observables per l'econòmetra i, llevat de la distribució exponencial, les altres alternatives permeten que es pugui donar un efecte dependència, en el sentit que la probabilitat d'abandonar un estat condicionada a estar-hi un determinat temps (alternativament la funció de sortida) depengui d'aquest temps.

Un dels problemes en aquest terreny és la importància dels factors no observats per l'econòmetra i que influeixen en l'esmentada probabilitat. La no-consideració d'aquests factors en l'especificació del model comporta inconsistències importants. D'una banda, se sobreestima la dependència negativa i, d'una altra, s'infravalora l'impacte de les variables explicatives que sí s'observen.

66. Heckman *et al.* (1999) expliciten un conjunt de mesures que serien d'interès més enllà de l'efecte mitjana i que comportaria l'anàlisi de canvis en la distribució, no únicament en la mitjana. En aquest aspecte cal destacar la contribució de Heckman *et al.* (1997).

67. En Heckman (1997) i Heckman *et al.* (1998) es discuteixen dos mètodes d'estimació per a aquest tipus de model.

Una part important de les aportacions de Heckman en l'àmbit dels models de durada té a veure amb el fet de mirar d'identificar separadament els efectes de l'heterogeneïtat no observada i els efectes de la dependència respecte del temps en aquest estat. En el cas de l'atur, la diferència esmentada entre aquests dos factors es podria reflectir si aquells que estan en atur tenen una major probabilitat d'estar en l'atur en el futur bé per la pèrdua de experiència (de capital humà) o bé per tenir senyals negatius de cara al mercat laboral (*market stigma*). Amb l'efecte heterogeneïtat les dotacions inicials incideixen d'una manera persistent en el fet d'abandonar un determinat estat, mentre que la dependència fa que els efectes de les dotacions inicials s'atenuïn o augmentin amb el pas del temps.

En un treball important, Heckman i Singer (1984) presenten les condicions d'identificació en models on la funció de sortida s'explicita, mentre que per a l'heterogeneïtat no observada no es proposa cap distribució, només requerint que compleixi les condicions per ser-ho. Les conseqüències de la incorrecta especificació funcional per mirar d'identificar els efectes dels dos factors esmentats es discuteixen en Heckman i Singer (1985), on es demostra que és possible identificar els dos efectes sense imposar formes funcionals concretes, encara que amb la imposició de supòsits restrictius poc encoratjadors, tal com destaca Heckman (1991).

## **El perquè de l'elecció dels guardonats d'aquest any**

Assar Lindbeck, prestigiós economista suec estretament lligat a la concessió del Premi Nobel d'Economia, escrivia a l'abril de 1989 un article on comentava les dificultats i problemes associats a l'elecció dels guardonats amb aquest premi, centrant l'atenció en quatre aspectes: com s'ha d'entendre l'àmbit de l'economia en la concessió del premi; quins criteris s'han d'aplicar per decidir si un candidat és mereixedor del premi; quin ordre s'ha de seguir en l'atorgament del premi; i quan el premi ha de ser compartit.

Tal com indica Lindbeck, la interpretació de l'Acadèmia del terme ciències econòmiques és ampli en la mesura en què la interdisciplinarietat s'ha considerat important sempre que comporti implicacions rellevants en termes econòmics. Els casos de Heckman i McFadden encaixen en moltes de les diferents accepcions que les contribucions d'anteriors guardonats hagin pogut tenir. Podríem destacar l'especificitat de les seves contribucions més importants, però al mateix temps han fet contribucions en camps diferents, han contribuït al desenvolupament de la Microeconometria com a disciplina i han fet aportacions interdisciplinàries, no únicament per l'ús de l'instrumental estadístic o la seva influència en altres camps, sinó també pel caràcter que tenen les contribucions empíriques dels dos guardonats.

Amb relació als criteris, Lindbeck esmenta l'originalitat, la rellevància científica i pràctica, la incidència en el treball científic i, fins a cert punt, l'impacte en la societat en sentit ampli, incloent les polítiques públiques. Tots aquests aspectes requereixen el pas del temps per contrastar la qualitat i la rellevància de les contribucions. Segons Lindbeck, per a l'Acadèmia prima la qualitat sobre la quantitat, ja sigui de nominacions o de cites. Les contribucions de Heckman i McFadden destil·len originalitat, han incidit de manera clara i manifesta en el treball empíric en moltes àrees, i també en el posterior treball metodològic en microeconometria, han estat contribucions

nombroses i freqüentment citades per la seva rellevància i han tingut incidència en la societat en un sentit ampli. Només cal recordar els estudis sobre la demanda potencial del sistema de transport ràpid (BART) a San Francisco, o l'aplicació del seu model de valoració contingent aplicat als recursos naturals emprat en la valoració dels danys al medi ambient generats pel desastre de l'Exxon Valdez a Alaska l'any 1989, pel que fa a les aportacions de McFadden. Les contribucions de Heckman han millorat els processos d'avaluació de programes socials o han contribuït a valorar l'impacte de les polítiques de discriminació positiva o sobre els drets civils en la situació econòmica dels afroamericans.

Les concessions d'aquest any corresponen al segon dels criteris d'ordenació cronològica dels candidats apuntat per Lindbeck: mantenir una visió pluralista de la recerca en economia, premiant diferents camps de l'anàlisi econòmica, amb enfocaments diversos i amb diferents visions del món. El premi d'aquest any és clarament un premi a l'econometria i, en concret, a la Microeconometria. Són els primers guardonats en aquest camp concret, la microeconometria, i han obtingut el quart premi concedit a un econòmetra al llarg dels trenta-dos anys d'existència del Premi Nobel d'Economia, després dels guardons a Ragnar Frisch i Jan Tinbergen (1969), Lawrence R. Klein (1980) i Trygve Haavelmo (1989) segons el camp destacat en la cita de la Reial Acadèmia Sueca de les Ciències<sup>68</sup>. Sens dubte, són dos dels més destacats participants en el desenvolupament en els darrers trenta anys d'aquesta disciplina. Altres també podien haver estat mereixedors del premi com a representants de la mateixa, però probablement són els dos econòmits amb unes contribucions específiques de més impacte i difusió entre la professió. Entre els qui es dediquen a l'anàlisi econòmica i fan ús de tècniques microeconòmiques, qui no ha utilitzat la *lambda* de Heckman a l'hora d'estimar un model, o el model lògic multinomial i ha vist freqüentment en els *outputs* de les estimacions la informació del *R<sup>2</sup> de McFadden* (nom de l'autor inclòs).

En relació amb l'últim aspecte lligat als problemes i a les dificultats a l'hora de decidir els guardonats, el dotzè premi compartit en Economia respon al caràcter complementari de les aportacions d'ambdós guardonats amb el comú denominador de l'anàlisi microeconòmica. Formen part d'un grup de premis compartits entre els quals, segons Lindbeck, també estan inclosos els d'Ohlin i Meade (1977), els de Lewis i Schultz (1979) i els de Fogel i North (1993).

Per acabar, unes quantes dades estadístiques que lliguen amb el camp dels guardonats. Destacar que Heckman i McFadden són els números 45 i 46 de la llista de guanyadors d'un Premi Nobel d'Economia, que formen part dels trenta atorgats a ciutadans dels Estats Units, i que els seus premis són el novè atorgat a un professor de la Universitat de Chicago i el tercer a un de la Universitat de Berkeley. Així mateix, el premi d'aquest any encaixa en una de les tendències en l'anàlisi econòmica durant aquesta segona meitat del segle xx, en opinió de Lindbeck, com és la creixent importància del mètodes quantitius i, en concret, de l'econometria.

68. Alguns altres guardonats han fet contribucions substantives al camp de l'Econometria, encara que no siguin les destacades per a la concessió del premi. És el cas de Tjalling C. Koopmans, guardonat l'any 1975, amb importants contribucions al camp de l'estimació de models econòmics, o James Tobin, guardonat l'any 1981, amb una aportació crucial en l'àmbit dels models microeconòmics per a l'anàlisi de la demanda, tal com s'ha destacat anteriorment.

## Comentaris finals

De tot el que hem dit, inclosa la cita de la Reial Acadèmia Sueca de les Ciències, no hauríem de concloure que la concessió del Premi Nobel d'aquest any a James Heckman i Daniel McFadden es deu exclusivament a les seves aportacions com a econòmetres. Encara que segurament sí que són les que poden haver tingut més impacte i les que justifiquen el caràcter compartit del premi, no són els únics camps en què han fet aportacions importants. Heckman, com ja s'ha esmentat, ha treballat en temes d'economia laboral (oferta de treball, models de durada, educació i ingressos), en temes relacionats amb l'avaluació de projectes socials, o amb les polítiques de drets civils i discriminació positiva. Per la seva banda, McFadden ha treballat també en temes de teoria de la producció, de medi ambient, o darrerament en temes relacionats amb el comportament econòmic de la gent de més edat.

Però encara més important, ambdós economistes són un excel·lent exponent del treball de recerca rigorós i ben fet amb implicacions rellevants per a l'anàlisi empírica i la presa de decisions. Precisament per aquest motiu, dintre del *seleccionat* col·lectiu d'investigadors que comparteixen aquestes virtuts, ells representen una *elecció* encertada.

## Referències bibliogràfiques

- Aitchison, J.; Silvey, S. (1957). «The Generalization of Probit Analysis to the Case of Multiple Responses», *Biometrika*, núm. 44, p. 131-140.
- Berkson, J. (1951). «Why I Prefer Logits to Probits», *Biometrika*, núm. 7, p. 327-339.
- Cragg, J. (1971). «Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods», *Econometrica*, núm. 39, p. 829-844.
- Frisch, R. (1933). «Editorial», *Econometrica*, núm. 1, p. 1-4.
- García, J. (1991). «Métodos de estimación de modelos de oferta de trabajo basados en la predicción de los salarios», *Investigaciones Económicas*, núm. 15, p. 429-455.
- Hausman, J. (1978). «Specification Tests in Econometrics», *Econometrica*, núm. 46, p. 1251-1272.
- Hausman, J.; McFadden, D. (1984). «Specification Tests for the Multinomial Logit Model», *Econometrica*, núm. 52, p. 1219-1240.
- Hausman, J.; Wise, D. (1978). «A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogenous Preferences», *Econometrica*, núm. 46, p. 403-426.
- Heckman, J. J. (1974). «Shadow Prices, Market Wages and Labor Supply», *Econometrica*, núm. 42, p. 679-694.
- (1976). «The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for such Models», *Annals of Economic and Social Measurement*, núm. 5, p. 475-492.
- (1979). «Sample Selection Bias as a Specification Error», *Econometrica*, núm. 47, p. 153-161.
- (1990). «Varieties of Sample Selection», *American Economic Review, Papers and Proceedings*, núm. 80, p. 313-318.
- (1997). «Instrumental Variables: A Study of Implicit Behavioral Assumptions in one Widely

- used Estimator», *Journal of Human Resources*, núm. 32, p. 441-461.
- Heckman, J. J.; Honoré, B. (1990). «The Empirical Content of the Roy Model», *Econometrica*, núm. 58, p. 1121-1149.
- Heckman, J. J.; Ichimura, H.; Todd, P. (1998). «Matching as an Econometric Evaluation Estimator», *Review of Economic Studies*, núm. 65, p. 261-294.
- Heckman, J. J.; Lalonde, R.; Smith, J. (1999). «The Economics and Econometrics of Active Labor Market Programs». A: O. Ashenfelter i D. Card [eds.]. *Handbook of Labor Economics*, vol. 3A, North-Holland, Amsterdam, p. 1865-2097.
- Heckman, J. J.; Smith, J.; Clements, N. (1997). «Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts», *Review of Economic Studies*, núm. 64, p. 487-535.
- Lee, L-F. (1978). «Unionism and Relative Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables», *International Economic Review*, núm. 19, p. 415-433.
- Lerman, S.; Manski, C. (1981). «On the Use of Simulated Frequencies to Approximate Choice Probabilities». A: C. Manski i D. McFadden [eds.], *Structural Analysis of Discrete Data with Economic Applications*, MIT Press, p. 305-319.
- Lindbeck, A. (1999). «The Sveriges Riksbank (Bank of Sweden) Prize in Economics Sciences in Memory of Alfred Nobel», disponible a [www.nobel.se](http://www.nobel.se)
- McFadden, D. (1973). «Conditional Logit Analysis of Qualitative Choice Behavior». A: P. Zarembka [ed.], *Frontiers in Econometrics*, Academic Press, Nova York, p. 105-140.
- (1978). «Modelling the Choice of Residential Location». A: K. Karlqvist, L. Lundqvist, F. Snickars i J. Weibull [eds.], *Spatial Interaction Theory and Planning Models*, Amsterdam, p. 75-96.
- (1984). «Econometric Analysis of Qualitative Response Models». A: Z. Griliches i M. Intriligator [eds.], *Handbook of Econometrics*, vol. 2, Amsterdam, p. 1395-1457.
- (1987). «Regression-Based Specification Tests for the Multinomial Logit Model», *Journal of Econometrics*, núm. 34, p. 63-82.
- (1989). «A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration», *Econometrica*, núm. 57, p. 995-1026.
- McFadden, D.; RUUD, P. (1994). «Estimation by Simulation», *Review of Economics and Statistics*, núm. 76, p. 591-608.
- Morgan, M. S. (1990). *The History of Econometric Ideas*, Cambridge University Press, Cambridge.
- Tobin, J. (1958). «Estimation of Relationships for Limited Dependent Variables», *Econometrica*, núm. 26, p. 24-36.