

James J. Heckman

(1944–)

Miért van nagymértékű és tartós egyenlőtlenség az emberek között, és hogyan függnak össze az egyenlőtlenség különböző dimenziói? Hogyan alakulnak ki az egyenlőtlenségek, és milyen szerepet játszanak a folyamatokban az egyéni döntések? És hogyan hatnak ezekre a döntésekre a különböző társadalompolitikai intézkedések? Ha röviden akarnánk összegezni, James J. Heckman munkássága az egyének közötti heterogenitás tartalmi és főként mérési problémái körüli fáradhatatlan válaszkeresés.

2000-ben Daniel L. McFaddennel közösen kapott közgazdasági Nobel-díjat, a mikroökonometria területén elért eredményeik elismeréseként, Heckman maga „a szelektív minták elméletének és elemzési módszereinek a kidolgozásáért”. A két díjazott olyan becslési eljárások kifejlesztésében és tökéletesítésében ért el jelentős eredményeket, amelyek az egyének és háztartások viselkedésének statisztikai elemzése területén széles körben elterjedtek, nemcsak a közgazdaságtan, hanem a többi társadalomtudomány területén is.

A tudományterület

A mikroökonometria – más néven keresztmetszeti és panel-ökonometria – a mikroszintű (egyéni, háztartási, vállalati stb.) adatok elemzésének statisztikai módszereit foglalja magában. Ökonometriáról van szó, vagyis a statisztikának arról az ágáról, amely közgazdasági és szélesebb társadalomtudományi problémákat vizsgál, jellemzően megfigyelési (és nem kísérleti) adatokon. Az adatok nem kísérleti eredete jelenti az ökonometria központi problémáját. A mikroökonometria a döntéshozók szintjén gyűjtött információk felhasználásával vizsgálja azt, hogy milyen tényezők befolyásolják az egyes döntéseket. Olyan kérdések vizsgálatát teszi lehetővé, mint hogy milyen tényezők hatnak arra, hogy az emberek dolgoznak-e vagy sem, és ha igen, mennyit; vagy hogy milyen továbbtanulási döntést hoznak, milyen szakmát választanak, hol laknak, hogyan közlekednek, milyen javakat fogyasztanak stb. A döntéshozói szintű adatok előnyei nyilvánvalóak: szemben az aggregált idősorokkal vagy területi összehasonlításokkal, itt közvetlenül vizsgálhatók az egyének, nem kell az aggregálás okozta információvesztéssel vagy

esetleges torzításokkal számolni. Ingyen ebédek azonban nincsenek: az egyéni adatokat nemcsak költségesebb összegyűjteni és elemezni (a modern számítógépek előtt ez utóbbi gyakorlatilag lehetetlen volt); elemzésük speciális technikai problémákat von maga után. Egyéni szintű adatok elemzése felszínre hoz ugyanakkor olyan alapvető tartalmi problémákat is, amelyek az aggregált elemzések során általában rejtve maradnak. A 2000. év közgazdasági Nobel-díjasai e problémák felismerésében és kezelésében értek el elévülhetetlen érdemeket.

Életrajz és intellektuális hatások

James Heckman 1944-ben született az Amerikai Egyesült Államokban, Chicagóban. 12 éves korában családjával a déli Kentucky, majd Oklahoma államba költözött. Önéletírása szerint¹ ez a néhány év és egy későbbi, a még tradicionálisabb déli államokba tett utazása egész életre szóló hatással volt rá. A déli államokban ekkor még a „normális” mindennapok része volt a faji diszkrimináció. A feketékkel való együttélést az „egyenlő, de elkülönített” bánásmód jellemezte, a fehérek számára fenntartott éttermekkel, strandokkal, buszokkal, utcai padokkal. Heckman számára a fekete amerikaiak sorsának a vizsgálata későbbi életében is fontos kérdés maradt, a legutóbbi években pedig különösen hangsúlyossá vált.

Középiskolába a Colorado állambeli Denver egyik kertvárosában járt. Kiemelkedő hatással volt rá fizikatanára, Frank Oppenheimer,² aki a fizika elméletét mindig a legfontosabb kísérleti eredményekkel együtt tárgyalta. Ez a szemlélet, amely alapelvként kezeli az elmélet és az empiria egységét, meghatározó eleme lett Heckman egész tudományos munkásságának. Colorado Springsben, a Colorado College-ban matematika szakon szerzett bachelor-fokozatot, ahol matematika mellett többek között a fejlődő országok gazdaságáról is tanult egy félévet. Itt ismerkedett meg a közgazdaságtannal mint tudománnyal – Adam Smith, David Ricardo, Arthur Lewis és Paul Samuelson munkáin keresztül –, és választotta hivatásul a matematika és a fizika helyett.

A Chicagói Egyetemen kezdte meg közgazdaságtani doktori tanulmányait, ahol Milton Friedman volt rá a legnagyobb hatással. Rövid idő után azonban átjelentkezett a Princetoni Egyetemre, követve a fejlődő országok közgazdaságtan-professzorát, Arthur Lewist. Princetonban viszonylag hamar átnyer-

¹ *James J. Heckman* – Autobiography. <http://nobelprize.org/economics/laureates/2000/heckman-autobio.html>

² *Frank Oppenheimer*et kommunista pártbeli múltja miatt elbocsátották professzori állásából a Minnesotai Egyetem fizikatanszékéről, így került középiskolai katedrára Coloradóban. *Frank Robert Oppenheimer*nek, a Manhattan project korábbi vezetőjének a testvére volt.

gelt a munkagazdaságtan és az ökonometria területére, amit az ott folyó ökonometriai és nagyszabású empirikus munkaerő-piaci kutatások motíváltak.

Heckman munkásságát egész életre meghatározta az a személet, amely hangsúlyozza az elmélet és a mérési módszerek összhangját, és különös figyelmet szentel a gazdaság- és társadalompolitika hatásainak elemzésére. Gary Becker és Jacob Mincer úttörő munkássága nyomán az 1960-as évek végétől a munkagazdaságtan lett e szemlélet legfontosabb területe. Különösen nagy hatású volt Mincer tanulmánya, amely az amerikai nők munkavállalását – „munkakínálatát” – vizsgálta.³ Heckman egyik témavezetője, Al Rees ebben az időben irányította az első nagyméretű tervezett társadalompolitikai kísérletet, a New Jersey állambeli negatívjövedelemadó-kísérletet.⁴ A tervezett társadalmi kísérletek kikerülnek a megfigyelési adatokkal járó problémák egy részét, ám rendkívül drágák, és nehéz politikai és erkölcsi problémákat vetnek fel, ezért nagyon ritkán végezhető el. A rendelkezésre álló (társadalmi kísérleti vagy megfigyelési) adatoknak egy ideális körülmények között elképzelt kísérlettel való szisztematikus összehasonlítása rendkívül sokat segít a problémák felismerésében. A negatívjövedelemadó-kísérletek ráirányították a figyelmet az ilyen összehasonlítások lehetőségére. A munkagazdászok mellett a princetoni évek alatt Heckmanra nagy hatással volt Richard Quandt ökonometriai munkássága is, aki nemlineáris modellt dolgozott ki a közlekedési döntések – a „közlekedés iránti kereslet” – empirikus elemzéséhez.⁵

Graduális tanulmányait 1971-ben fejezte be. Ezután a Columbia Egyetemre került, ahol nagy hatással volt rá Kelvin Lancaster és Edmund Phelps munkássága, akik a reprezentatív döntéshozó feltételezésének problémáit feszegették, és hangsúlyozták a makroökonómia mikroalapjainak fontosságát. Az egyetem mellett az akkor New York városában található NBER⁶ munkatársa lett. Kollégái, Jacob Mincer, Finis Welch és abban az időben közeli barátja, Robert Willis mellett megtermékenyítő hatásúak voltak az interakciók az NBER gyakori vendégeivel, Gary Beckerrel, Reuben Gronauval és Sherwin Rosennel. Egy nyári konferencia során ismerkedett meg Daniel McFaddennel, aki – részben Quandt munkásságát kiteljesítve – nemlineáris ökonometriai módszert dolgozott ki diszkrét fogyasztói döntések mikroökonometriai vizsgálatára.

³ Mincer (1962).

⁴ Rees (1974). Lásd még Widerquist (2002) összefoglalóját a negatívjövedelemadó-kísérletek eredményeiről és problémáiról.

⁵ Quandt (1956).

⁶ National Bureau of Economic Research: az empirikus munkát folytató amerikai közgazdászok elitjét tömörítő, igen kiterjedt kutatói hálózat és annak részeként működő kutatóintézet.

James Heckman 1973 óta a Chicagói Egyetem Közgazdaságtan Tanszékének professzora. Itt természetesen sok kollégája volt rá hatással, köztük a legfontosabbak Gary Becker mellett William Brock, Lars Hansen, Jose Scheinkman, és a szociológus James Coleman voltak. PhD-diákok tucatjai nevelkedtek és nevelkednek ma is az irányítása alatt; többükkel hosszú távú munkakapcsolatot alakított ki (néhány közülük: Bo Honoré, Thomas MaCurdy, Richard Robb).

A chicagói tanszék az azt körülvevő intézetekkel nemcsak a világ egyik legmagasabb színvonalú közgazdasági kutatóműhelye, hanem a tudományhoz való sajátos közelítése miatt különálló iskolának is tekinthető. Mindezt rendkívüli egyéniségek is biztosították: a hetvenes évek elején Milton Friedman emelkedett ki vezéralakként, hogy aztán a későbbiekben olyanok vegyék át ezt a szerepet, mint George Stigler és Gary Becker. A chicagói megközelítés lényege, hogy a közgazdaságtan a társadalmi kérdések megértésének a leginkább megfelelő kerete. Nem arról van szó, hogy a diszciplína átütő sikert ért volna el e feladatában: a világ rendkívül bonyolult, a legfontosabb jelenségeket és problémákat nem igazán értjük. A feladat ezért a szisztematikusan haladás, még ha ez gyakran lassúnak tűnik is, és az irányzatok egy része később zsákutcának bizonyul. A chicagói szemlélet szerint a társadalomtudományok közül egyedül a közgazdaságtan rendelkezik olyan belső koherenciával és az elmélet és mérés összhangjával, ami erre a feladatra alkalmazható. A közgazdaságtan e feladatából fakadóan nem egyszerűen egy izgalmas játék érdekes feladatok megoldására, hanem véresen komoly misszió. Ezt a missziót pedig csak komoly munka és elhivatottság árán lehet teljesíteni, aminek velejárója a határozott, gyakran doktrinernek tűnő véleményalkotás is. A chicagói kutatószemináriumok hangulata és a legerősebb egyéniségek (köztük Heckman) vitastílusa ezért gyakran erősnek, sőt esetenként kimondottan udvariatlannak tűnik. Mindezek mögött az áll, hogy a stílus, az elegancia nem szempont sem az elméletekben, sem a tudományos párbeszédben: annál sokkal komolyabb dolgokról van szó.

Munkássága

James Heckman tanulmányok százait írta, nagy részüket társszerzőkkel közösen, és fáradhatatlanul dolgozik napjainkban is. Munkásságán végigvonul az emberek közötti egyenlőtlenség vizsgálata és a társadalompolitikai intézkedések hatásának az elemzése, illetve ezek mérésének és magyarázatának módszertani problémái. A továbbiakban két, egymással erősen összefüggő területet mutatunk be részletesebben: a szelektált minták elméletét és a társadalompolitikai programok hatásvizsgálatának módszertani problémáit. Heckman rengeteg más területen is dolgozott és ért el fontos eredményeket – ezekről csak említést teszünk, azt is a teljesség igénye nélkül.

A szelekciós probléma és kezelése

Heckman első és talán legfontosabb eredménye a szelektált minták modelljének (röviden mintaszelekciós modell) kidolgozása, és annak alkalmazása volt a női munkakínálatra. A Svéd Királyi Akadémia ezért az eredményért ítélte meg számára a Nobel-díjat. Heckman ezzel a problémával a hatvanas évek végén és a hetvenes évek elején foglalkozott intenzíven, amikor a Columbia Egyetem és az NBER munkatársa volt. A legfontosabb tanulmányok az 1974-es és 1979-es *Econometrica*-cikkek és az 1976-os statisztikai módszertani cikk volt, de a Bo Honoréval közösen írt 1990-es cikke is ide tartozik.

A szelekciós probléma a megfigyelési adatok természetes velejárója. Ha például arra vagyunk kíváncsiak, hogy mennyit ér valakinek a munkája a munkaerőpiacon, csak azokra az emberekre lehet béradatunk, akik dolgoznak. Ha azt szeretnénk megmérni, mennyit ér az egyetemi diploma a munkaerőpiacon, csak azoknak a keresetéből indulhatunk ki, akik egyetemet végeztek. A szelekció ezekbe az „almintákba” (a dolgozók, illetve az egyetemi végzettségűek csoportjába) nem véletlenszerű. Szemben egy ideális kísérlettel, itt nem a kutató dönti el véletlenszerűen, hogy ki dolgozik, illetve ki szerez diplomát, hanem maguk az egyének, illetve olyan körülmények döntenek, amelyeket a kutató nem ismer pontosan. A probléma ebből az „önszelekcióból” fakad. Könnyen lehet, hogy azok, akik nem dolgoznak, alacsonyabb bér ajánlatokat kapnának, mint azok, akik dolgoznak, hiszen ez utóbbiak kaptak számukra elfogadható ajánlatot, míg akik nem dolgoznak, nem kaptak. Hasonlóképpen az is elképzelhető, hogy azok, akik nem végeztek egyetemet, kevesebbet nyertek volna az egyetemi diplomával, mint azok, akik el is végezték. A szelektált alminta (dolgozók, illetve diplomások) ekkor nem reprezentálja a többieket, így torz eredményekhez vezethet, ha rájuk korlátozzuk az elemzést.

Tekintsük az eredeti heckmani kérdést: hogyan befolyásolja az egyének munkavállalását, ha keresetük (potenciális keresetük) vagy egyéb, nem munkából származó jövedelmük egzogén módon megnő? A még konkrétabb példa a férjzett nők munkavállalása, ahol nem munkából származó jövedelmeként a férj munkajövedelme kiemelkedően fontos tényező. Ekkor mind a két tényező egzogén növekedését idézheti elő például egy személyi jövedelemadó-csökkentés. A mérési problémát Gronau (1974) és Heckman (1974) egymástól részben függetlenül szorosán a munkakínálat elméletéből vezette le: a munkavállalás egyéni döntés, amit az ajánlott bér és a rezervációs bér különbsége határoz meg. (A rezervációs bér az a szint, ami felett az egyén elvállalná a munkát: az a megfontolás húzódik meg mögötte, hogy mindig van olyan, esetleg akár csillagászati összeg, ami mellett megérné munkába állni.) A rezervációs bért természetesen senkinél nem figyelhetjük

meg, az ajánlati bért pedig csak azoknál, akik azt el is fogadják, és munkába állnak.⁷ A mérési modellt két egyenletben foglalhatjuk össze: az első az ajánlati bér, a második az ajánlati és a rezervációs bér különbségének az egyenlete:

$$\ln W_i = \beta_1' x_{1i} + u_i$$

$$I_i^* = \beta_2' x_{2i} + v_i$$

$$I_i = 0 \text{ ha } I_i^* < 0$$

$$I_i = 1 \text{ ha } I_i^* \geq 0$$

Feltesszük, hogy x_1 és x_2 egzogének:

$$E(u | x_1) = E(u | x_2) = 0$$

$$E(v | x_1) = E(v | x_2) = 0$$

Az i index jelöli a különböző egyéneket, W az ajánlott bér, x_1 az ajánlati bért (az egyén határtermékét) befolyásoló megfigyelhető tényezők vektora, β_1 ezek hatása az ajánlati bérré, u pedig az ajánlati bért meghatározó nem megfigyelhető tényezőket sűríti magában. I^* az ajánlott és a rezervációs bér (W_R) különbsége, I pedig egy úgynevezett index változó: 0 értéket vesz fel, ha az egyén nem dolgozik (mert az ajánlott bér kisebb, mint a rezervációs bére), 1-et, ha igen. x_2 a munkavállalási döntést (az ajánlati és a rezervációs bér különbségét) befolyásoló megfigyelhető tényezők vektora, β_2 ezek hatása a döntésre (a különbségre), v pedig a döntést (a különbséget) meghatározó nem megfigyelhető tényezőket sűríti magában. Minthogy a munkavállalási döntésben szerepel az ajánlati bér, a második egyenlet megfigyelhető és nem megfigyelhető tényezői tartalmazzák az első egyenlet tényezőit, és hasonlóképpen, β_2 tartalmazza β_1 -et is.

A modell két kérdésre keresi a választ: mi és hogyan befolyásolja az ajánlati béreket; mi és hogyan befolyásolja a munkavállalást? Az első kérdésre β_1 ad választ, a másodikra pedig kétfajta válasz is van. β_2 adja meg a teljes hatásokat („redukált” válasz), ennél azonban érdekesebb a „strukturális” válasz, ami szeparálja az ajánlati és a rezervációs bérré gyakorolt hatásokat. A rezervációs bérré gyakorolt szeparált hatásokat azonban β_2 és β_1 megfelelő részeinek (x_2 és x_1 közös részére vonatkozó paramétereknek) a különbsége

⁷ Valójában természetesen esetükben sem az összes ajánlatot figyeljük meg, csak a legjobbat, amit el is fogadtak; kompetitív munkaerőpiac esetében azonban az ajánlott legmagasabb bér a munkavállaló határtermékértéke.

méri. Akár az első, akár a második kérdésre keressük is a választ, elengedhetetlen β_1 megfelelő mérése.

A legfontosabb mérési probléma abból fakad, hogy az első egyenlet csak azok esetében becsülhető, akik dolgoznak. Ezek pedig szelektált mintát alkotnak a teljes sokaságból, hiszen a munkavállalás nem véletlenszerű, azt a második egyenlet határozza meg:

$\ln W_i$ akkor és csak akkor megfigyelhető, ha $I_i^* \geq 0$.

β_1 becslése torz, ha csak a munkavállalókra korlátozott mintából történik:

$$E(\ln W_i | x_{1i}) = E(\beta_1' x_{1i} + u_i | x_{1i}) = \beta_1' x_{1i} + E(u_i | x_{1i}) = \beta_1' x_{1i}$$

de

$$E(\ln W_i | x_{1i}, I_i^* \geq 0) = E(\beta_1' x_{1i} + u_i | x_{1i}, I_i^* \geq 0) = \beta_1' x_{1i} + E(u_i | x_{1i}, I_i^* \geq 0) \neq \beta_1' x_{1i}$$

mert $E(u_i | x_{1i}, I_i^* \geq 0) \neq 0$. munkavállalást (az $I_i^* \geq 0$ feltételt) befolyásoló nem megfigyelhető tényezők (v) között szerepelnek az ajánlati bért meghatározó nem megfigyelhető tényezők is (u). A nem megfigyelhető ismerveik miatt magasabb ajánlatot kapók inkább fognak munkát vállalni, és ezért inkább lesznek részei annak a mintának, amelyről van megfigyelésünk:

$$E(u_i | x_{1i}, I_i^* \geq 0) = E(u_i | x_{1i}, \beta_1' x_{2i} + v_i \geq 0) = E(u_i | x_{1i}, v_i \geq -\beta_1' x_{2i}) \neq E(u_i | x_{1i}) = 0$$

Heckman (és Gronau) 1974-ben publikált felfedezése nemcsak az, hogy ez a torzító hatás („szelekciós torzítás”) létezik, hanem az is, hogy ha feltesszük a nem megfigyelhető tényezők normális eloszlását, akkor egy maximum likelihood problémaként megoldható a becslés, és konzisztensen becsülhető β_1 .

Az 1974-ben publikált megoldás azonban nem tárta fel a módszer mögött húzódó intuíción, és az akkori körülmények mellett rendkívül hosszú numerikus becslési eljárást kívánt. Heckman 1976-ban és 1979-ben publikált cikkeiben mutatta ki, hogy ugyanezek a feltevések egy egyszerűbb és intuitívabb megoldást is lehetővé tesznek, amit a mintaszelekció kétlépcsős eljárásának is szokás nevezni.⁸

⁸ Minthogy az alternatív becslési eljárás nem maximum likelihood, annál kevésbé hatásos; ez az ára az egyszerűbb és intuitív megközelítésnek. Az gyakorlatban igen nagy mintákon szokás ilyen jellegű becsléseket végezni, ezért a hatásosság kérdése kevésbé jelentős.

Feltesszük tehát, hogy u és v normális, u szórása σ , korrelációjuk ρ , és v szórása 1.⁹ Ekkor belátható, hogy

$$E(u_i | x_{1i}, v_i \geq -\beta_2' x_{2i}) = \rho\sigma\lambda_i$$

$$\lambda = \frac{\phi(\beta_2' x_{2i})}{\Phi(\beta_2' x_{2i})}$$

ahol ϕ és Φ rendre a standard normális eloszlás sűrűség- és eloszlásfüggvénye ρ u és v korrelációja, σ pedig u szórása. Minthogy pedig β_2 egy probit modelltől becsülhető, minden egyes i egyénre megbecsülhető λ_i , és az szerepeltethető az első egyenletben, ami által β_1 konzisztensen becsülhető:

$$\ln W_i = \beta_1' x_{1i} + \gamma\lambda_i + u_i$$

ahol $\gamma = \rho\sigma$.

A kétlépcsős eljárás tehát a következő: először megbecsüljük β_2 -t egy probit modelltől, majd azt behelyettesítve megbecsüljük λ_i -t minden egyénre. Ezt aztán hozzáadjuk az első egyenlethez, mint egy magyarázó változót, és egyszerű OLS-sel megbecsüljük. A standard hibákat korrigálni kell, ha $\gamma \neq 0$, de a paraméterbecslések konzisztensek.

A szelekciós torzítás abból fakad, hogy u és v korreláltak: ha nem azok, akkor nincs ilyen torzítás, és λ_i együtthatója 0 (nem kell szerepeltetni az egyenletben). A torzítás irányát u és v korrelációjának az előjele határozza meg: pozitív korreláció esetén pozitív a torzítás, vagyis a szelektált mintából becsült β_1 a valósnál magasabb lesz (γ előjele megegyezik a torzítás irányával, hiszen azt ρ előjele határozza meg).

A szelekciós torzítás tehát felfogható egy kihagyott változó problémájaként: ha nem szerepeltetjük λ_i -t az első egyenletben, akkor becslése torz és inkonzisztens, míg ha szerepeltetjük, akkor nem az. Ez a „kihagyott változó” analógia különösen jól mutat rá az intuíció lényegére: a probléma a szelekció endogenitása, amit az endogenitás általános kezelési lehetőségeivel kezelíthetünk meg. A triviális megoldás a problémára az, ha nem hagyjuk ki a változót.

Annyival bonyolultabb itt a helyzet, hogy a kihagyott változó nem rendelkezik közvetlen tartalommal, nem az egyének természetes tulajdonsága, és

⁹ Itt v szórását azért normalizáljuk, mert az ajánlati és a rezervációs bér különbségét nem figyeljük meg, csak azt, hogy nagyobb-e nullánál. A modell szerint ugyanis ez vezet munkavállaláshoz. I^* és így v szóródása azt mutatná meg, hogy mennyivel nagyobb vagy kisebb a különbség nullánál – ezt nem tudjuk megbecsülni, de nem is szükséges. A szórás 1-re normalizálása ezért nem jelent információvesztést. Ez egyébként nem más, mint a probit modellek mögötti standard érvelés.

közvetlenül nem is megfigyelhető. Értelmezése a következő: az adott megfigyelhető ismérvekkel (x_1 -ekkel) rendelkező emberek között várhatóan mennyivel magasabb bérajánlatot kapnak azok, akik azt elfogadják azt, mint azok, akik nem. E várható értékkülönbség meghatározásához két dolgot használunk fel: u és v eloszlását és a második egyenletből becsült I' értéket ($\beta_2' x_2$ -t). Az előbbi szerepe nyilvánvaló: ha nem tételezzük fel a normális eloszlást, nem vezethetjük le λ -t.

Technikailag λ akkor is meghatározható, és a becslés rendesen végigvihető, ha x_1 és x_2 ugyanaz, hiszen az első egyenletben e vektornak lineáris függvénye mellett λ ugyanezen vektor nemlineáris függvénye. Hihetőbb azonban az identifikáció akkor, ha x_2 tartalmaz olyan változót, amit x_1 nem, azaz van olyan tényező, ami befolyásolja a munkavállalást, de az ajánlati bért nem, vagyis befolyásolja a rezervációs bért, de az ajánlatit nem. Ekkor γ -t nagyrészt ezek az x_1 -ben nem szereplő változók identifikálják. Az ilyen változókat instrumentális változóknak (röviden instrumentumoknak) nevezzük; az instrumentális becslés általában is az endogenitás kezelésének az első számú módszere. Megfelelő instrumentumokat azonban igen nehéz találni, ráadásul külső információ nélkül nem dönthető el, hogy mi megfelelő, és mi nem. A kezdeti lelkesedés után az alkalmazott kutatóknak is fel kellett ismerniük: a mintaszelekció módszere nem csodaszer, a megoldás kritikus módon függ a megfelelő instrumentumoktól és az eloszlásbeli feltevésektől.

A modell a James Tobin által kidolgozott és róla elnevezett „tobit” modell továbbfejlesztésének is tekinthető (Tobin, 1998). A szakmai zsargon valószínűleg ezért is aggatta az eljárásra a „heckit” nevet.¹⁰ A mintaszelekciós modell pedig továbbfejleszhető egy úgynevezett „switching” regressziós modellé, ahol két „első” egyenlet van: ha az index függvény 1 értéket vesz fel, akkor az egyik figyelhető meg, ha 0-t, akkor a másik. Robert Willis és Sherwin Rosen 1979-ben publikált tanulmánya tartalmazza az első ilyen modellt: a szelekció a felsőfokú továbbtanulás modellje, az első két egyenlet a két állapotban (középiskolai végzettség vagy felsőfokú végzettség) elérhető keresetet modellezi, a harmadik, a szelekciós egyenlet a továbbtanulási döntést. E modell becslésének módszertani részleteit Rosen korábbi diákja, Lung-Fei Lee publikálta, szintén 1979-ben.

A probléma, kicsit másként megfogalmazva, Roy-modellként már korábban is ismert volt a szakma egy része számára, Andrew Roy 1951-es rövid tanulmánya nyomán. Heckman korábbi tanítványával, Bo Honoréval közö-

¹⁰ Az „it”-ezés eredete a logisztikus eloszlás-feltételezés melletti bináris függő változós modellben gyökerezik: a korábban logisztikus regresszióként ismert módszert kezdték el először rövidítve „logit”-nak nevezni. Ennek normális eloszlás-feltételezés melletti analógiájára a „probit” név terjedt el, kicsit félrevezető módon (a valódi analóg a „normit” lett volna). A Tobin-féle cenzorálási modell „tobit” elnevezését *Arthur Goldbergernek* tulajdonítják.

sen írt 1990-es cikke tisztázza a switching modell viszonyát az eredeti Roy-modellhez, és tárgyalja az eloszlásbeli és a különböző instrumentális változó feltevések jelentőségét. Részben e tanulmány, részben pedig a munkakínálati modellek nem kielégítő empirikus teljesítménye nyomán vált világossá, hogy e feltevések egyáltalán nem ártatlan technikai részletek, hanem az identifikálhatóság kulcsfontosságú feltételei. Az 1990-es években új lendületet kapott az endogén szelekció modellezési lehetőségeinek a kutatása, mégpedig a társadalompolitikai programok hatásvizsgálatának területén. James Heckman itt is úttörő szerepet játszott.

Társadalompolitikai programok hatásvizsgálata

Heckman ma is aktív kutatási területe a társadalompolitikai intézkedések („programok”) hatásvizsgálata módszertani kérdéseinek vizsgálata és különböző programok értékelése. Az aktív foglalkoztatáspolitikai eszközök talán a legalaposabban vizsgált társadalompolitikai programok és Heckman vonatkozó munkássága is jelentős mértékben ezek köré fonódik. Összefoglaló tanulmánya a 130 oldalas *Handbook of Labor Economics* cikk, amit Robert LaLonde-dal és Jeffrey Smithszel közösen írt.¹¹ Az aktív foglalkoztatáspolitikai eszközök munkanélküli, egyéb nem foglalkoztatott vagy alacsony keresetű (jellemzően hátrányos helyzetű) egyénekre fókuszálnak, és célul azok foglalkoztatási vagy kereseti helyzetének javítását tűzik ki.

A vizsgált kérdés az, hogy milyen hatása van az adott programnak a résztvevők foglalkoztatására (keresetére). A kérdés egész pontosan az, hogy *milyen eredményt értek el a résztvevők, összehasonlítva azzal, amit akkor értek volna el, ha nem vesznek részt a programban.* A megmérendő hatás tehát egy tényleges állapot (milyen eredményt értek el a program után) és egy úgynevezett *tényellentétes állapot* (milyen eredményt értek volna el a program hiányában) összehasonlításából adódik. Nyilvánvaló, hogy miért nem mérhető („identifikálható”) a program hatása pusztán a program résztvevőinek vizsgálatával: a tényellentétes eredmények így nem mérhetőek. A probléma felfogható a mintaszelekció egy eseteként; amiért külön tárgyalást érdemel, az a „heckittől” eltérő, úgynevezett nemparametrikus (normális eloszlást nem feltételező) megközelítés.

A tényellentétes eredmények mérése tehát elengedhetetlen a program hatásának azonosításához. A módszertani problémák pedig a tényellentétes eredmények mérésének nehézségéből fakadnak: valójában sosem tudhatjuk pontosan, hogy mi történt volna a résztvevőkkel, ha nem vesznek részt, hiszen egy egyén vagy részt vesz a programban, vagy nem. Az egyszerűség kedvéért

¹¹ Heckman–LaLonde–Smith, 1999.

tekintsünk egyetlen eredményváltozót (pl. foglalkoztatás két év múlva), tegyük fel, hogy a résztvevők ugyanazt a „kezelést” kapják, és hogy a program csak azokra hat, akik részt vesznek benne. Jelölje i egy adott egyén indexét, Y_i pedig az eredményt. Legyen D_i az a bináris változó, amely a részvételt jelöli: $D_i=1$, ha i részt vesz a programban, $D_i=0$, ha nem. A kísérleti statisztika nyelven $D_i=1$, ha az egyén kezelést kap (a kezelt csoportban van), 0, ha nem kap kezelést (a kontrollcsoportban van). Jelölje Y_{1i} az i -ik egyén kezelés utáni eredményét, Y_{0i} pedig a kezelés hiányában bekövetkező eredményét.

Gondolatkísérletünkben minden egyes egyén esetében szeretnénk tudni a program hatását az eredményváltozóra:

$$\Delta_i = Y_{1i} - Y_{0i}$$

Az identifikációs probléma miatt Δ_i természetesen nem mérhető: a résztvevők esetében Y_{0i} tényellentétes, a nem résztvevők esetében Y_{1i} tényellentétes. Tehát:

$$Y_{1i} \mid D_i = 1 \text{ mérhető, de } Y_{0i} \mid D_i = 1 \text{ nem mérhető; és}$$

$$Y_{1i} \mid D_i = 0 \text{ nem mérhető, de } Y_{0i} \mid D_i = 0 \text{ mérhető.}$$

Az első következménye ennek az, hogy Δ_i eloszlása önmagában nem identifikált még ideális kísérleti szituációkban sem.¹² Más szóval, egyéb, igen erős feltevések nélkül semmi esélyünk arra, hogy pusztán a megfigyelt eredmények alapján rekonstruáljuk a program hatásának az eloszlását. Ennek az az oka, hogy sosem tudhatjuk, hogy aki kezelést kapott, és a kezelt csoport eredményváltozó-eloszlásában mondjuk magasan helyezkedik el, az kezelés hiányában hol lenne az eredményváltozó eloszlásában.

Heckman egyik fontos felismerése volt, hogy a hatásvizsgálatok implicit módon kétféle várható értéket is definiálnak: a részvétel vagy „kezelés” várható hatását a népesség (vagy a célcsoport) egy véletlenül kiválasztott tagjára (ATE),¹³ vagy csak a résztvevőkre (TT).¹⁴ Elvi szempontokból a két várható hatás között jelentős a különbség: az ATE tartalmazza a programnak a közvetlen hatását azokra is, akik aktuálisan nem résztvevők, míg a TT nem. A várható hatások formális definíciója:

$$\begin{aligned} \text{ATE} &= E(\Delta_i) = E(Y_{1i} - Y_{0i}) = E(Y_{1i}) - E(Y_{0i}) \\ &= E(Y_{1i} \mid D_i = 1)\Pr(D_i = 1) + E(Y_{1i} \mid D_i = 0)\Pr(D_i = 0) \\ &\quad - E(Y_{0i} \mid D_i = 1)\Pr(D_i = 1) - E(Y_{0i} \mid D_i = 0)\Pr(D_i = 0) \end{aligned}$$

$$\text{TT} = E(\Delta_i \mid D_i = 1) = E(Y_{1i} \mid D_i = 1) - E(Y_{0i} \mid D_i = 1).$$

¹² Ezt a problémát Heckman–Smith–Clemens (1997) mutatta ki.

¹³ Average Treatment Effect.

¹⁴ Average Effect of the Treatment on the Treated.

A két várható hatás ezek alapján a következőképpen függ össze:

$$\begin{aligned} \text{ATE} &= E(\Delta_i \mid D_i = 1) \Pr(D_i = 1) + E(\Delta_i \mid D_i = 0) \Pr(D_i = 0). \\ &= \text{TT} \cdot \Pr(D_i = 1) + E(\Delta_i \mid D_i = 0) \Pr(D_i = 0). \end{aligned}$$

Az ATE esetében két, a TT esetében egy tényellentétes várható érték kell mérni. Tervezett (és ezeket imitáló természetes¹⁵) kísérletekben ezek a tényellentétes átlagos hatások egyszerűen mérhetők:¹⁶ a véletlenszerű hozzárendelés miatt

$$E(Y_{0i} \mid D_i = 1) = E(Y_{0i} \mid D_i = 0) \text{ és } E(Y_{1i} \mid D_i = 1) = E(Y_{1i} \mid D_i = 0)$$

és így a kezelt és a kontroll csoportban mért eredmények különbsége azonosítja mindkét hatást,

$$\text{ATE} = \text{TT} = E(Y_{1i} \mid D_i = 1) - E(Y_{0i} \mid D_i = 0).$$

Az esetek nagy részében azonban nem tervezett vagy természetes kísérletek határozzák meg a programban való részvételt: a résztvevők almintája szelektált. Ilyenkor is valamilyen kontrollcsoport választása az egyetlen járható – bár tökéletlen – út a hatások méréséhez: a tényellentétes várható eredményeket csak más egyének eredményeivel közelíthetjük. Rendkívül fontos kérdés e más egyének mintájának, a „kontrollcsoportnak” a meghatározása és az, hogy az összehasonlítás milyen módszerrel történik. Természetesen ilyenkor minden várható érték további feltételek mellett értendő: ugyanolyan megfigyelhető ismérvekkel rendelkezők átlagos eredményeit kell összehasonlítani. Helyhiány miatt itt nem bocsátkozunk bele az alkalmazható módszerek részleteibe, csak néhány általános megjegyzésre szorítkozunk.¹⁷

Három oka lehet annak, hogy a programban részt vevők tényellentétes eredményét nem jól közelítik a résztvevők kontrollcsoportjának az eredményei. Egyrészt változhatott az általános gazdasági helyzet mialatt a résztvevők – például átképzés miatt – nem voltak a munkaerőpiacon. Másrészt az érintett csoportok iránti kereslet eltérhet a többiekétől. A harmadik probléma kicsit bonyolultabb: ha a program résztvevőinek meg kell felelniük bizonyos kritériumoknak (pl. fél éve munkanélküli), akkor a veszélyeztetett (magas

¹⁵ Természetes kísérletek azok a helyzetek, amikor a programban való részvétel nem tudatos véletlenszerű hozzárendelésnek az eredménye, a kutató mégis úgy gondolja, hogy a részvételt a vizsgált probléma szempontjából irreleváns tényezők mozgatták, ezért az akár véletlenszerűnek is felfogható („ingorable treatment”). Azok a helyzetek, amelyeket valóban természetes kísérletnek tekinthetünk, rendkívül értékesek oksági hatások azonosítására szempontjából.

¹⁶ Technikailag azért azonosított az átlagos hatás (a hatás várható értéke) és azért nem azonosított a hatás eloszlása, mert a várható érték lineáris operátor: a különbség várható értéke nem más, mint a várható értékek különbsége.

¹⁷ Az érdeklődő olvasó *Heckman, LaLonde és Smith* 1999-es tanulmánya mellett magyar nyelven *Kézdi* (2004) műhelytanulmányából meríthet további ismereteket.

munkanélküliségi esélyű) csoportokon belül is azok vehetnek csak részt, akik éppen a szükséges időszakban voltak hátrányos helyzetben. Egy részük azonban tipikus esetben a program nélkül is kilábalna ebből, így, ha csak a résztvevők pályáját elemeznénk, ezt a potenciális önerős felemelkedést is a programnak tudnánk be. A megfelelő kontrollcsoport és a megfelelő mérési módszer megtalálása e problémák kiküszöbölését jelenti.

Heckman egyik rendkívül lényeges meglátása, hogy a kontrollcsoport és az alkalmazott mérési módszer megválasztásának a problémája szervesen összefügg egymással: a statisztikai-ökonometriai módszer nagymértékben meghatározza a megfelelő kontrollcsoportot is. Ráadásul az egymástól függetlenül definiált módszerek között szoros az átjárás: megfelelő feltételezések esetén a különböző módszerek egymással ekvivalensek. A mérni kívánt hatások pontos definíciója, a mérési módszerek és azok egymáshoz való viszonyának feltárása, és a mögöttük meghúzódó feltevések precíz számbavétele napjainkban is folyó intenzív kutatás tárgya. Ahogyan az ilyenkor lenni szokott, különböző szerzők különböző terminusokat használnak, egymás eredményeit, illetve azok értelmét és jelentőségét gyakran kétségbe vonják, és részben emiatt, a különböző eredmények egymáshoz való viszonya ma még nem teljesen tiszta. Ebben a munkában – sok esetben vitában – James Heckman talán a legaktívabb szereplő: munkássága nagymértékben hozzájárul a vitás kérdések tisztázásához, illetve a probléma új aspektusainak a feltárásához.

Egy utolsó megjegyzés a hatásvizsgálat és a mintaszelekció „heckit” modelljének a kapcsolatáról. A hatásvizsgálat problémája felírható az eredeti szelekciós modell terminusaiban is (az első egyenlet vagy egy diszkrét foglalkoztatási egyenlet, vagy egy folytonos keresleti egyenlet lenne, a második egyenlet pedig a programban való részvételt határozná meg). Az itt tárgyalt megközelítés abban különbözik az előző problémától, hogy nemparametrikus módon, vagyis függvényforma és eloszlásfeltevések nélkül modellezi mind a hatást, mind a szelekciós mechanizmust. A probléma itt tartalmilag szűkebb (társadalompolitikai program hatásvizsgálata), és ez teljesen általános („nemparametrikus”) módszertani megközelítést tesz lehetővé. Amennyiben ezen a szűkebb területen konszenzusra jut a szakma, az megkönnyítheti az általánosabb problémák mérési szempontból általánosabb (nemparametrikus) megközelítését is. James Heckman 2000. évben tartott Nobel-díjas-előadása ebbe az irányba tett jelentős lépésnek tekinthető.

Heckman szerepe az empirikus társadalomkutatásban

A fenti két nagy téma mellett Heckman rengeteg más területen dolgozott és ért el fontos eredményeket. A teljesség mindennemű igénye nélkül, röviden megemlítünk néhányat ezek közül. Az egyik ilyen terület (amelyen Burton Singerrel közösen dolgozott a nyolcvanas években) szintén az egyének kö-

zötti heterogenitás okozta mérési problémákkal foglalkozik, az időtartam-modellek (hazard modellek) keretein belül.¹⁹ Az időtartam-modellek olyan dolgok magyarázatát célozzák meg, mint a munkanélküliség hossza, az egy munkahelyen eltöltött idő vagy az emberi élet hossza stb. Az időtartamokat, ha már egyszer elkezdődtek, az határozza meg, hogy mikor milyen valószínűséggel érnek véget (az ember munkanélküliből foglalkoztatott lesz, kilép a munkahelyéről, meghal stb.). A közgazdasági alkalmazások egyik legfontosabb kérdése az, hogy a kilépési valószínűséget befolyásolja-e az, hogy mennyi ideje van az egyén az adott állapotban: például a munkanélküliség alatt csökken-e annyira az egyénnek a munkaerőpiacon értékelt tudása, ambíciói, esetleg pszichés egészsége, hogy minél többet tölt munkanélküliként, annál nehezebben helyezkedik el később. Ezt az időtartam-függést azonban megfigyelési adatokat elemezve nehéz megkülönböztetni az eredeti heterogenitástól: ha vannak eleve „jobbak” (könnyebben elhelyezkedők) és „rosszabbak”, akkor ugyanakkor munkanélkülivé lett emberek közül először a jobbak fognak elhelyezkedni, és minél inkább telik az idő, annál inkább a rosszabbak lesznek túlsúlyban a még munkanélküliek között, akik viszont nehezebben helyezkednek el. Ha az eleve jobbakat nem tudjuk teljes mértékben megkülönböztetni az eleve rosszabbaktól, akkor azt látjuk, hogy az idő múlásával egyre nehezebben helyezkednek el a munkanélküliek, ám ennek lehet, hogy semmi köze a munkanélküliség demoralizáló hatásához: a jelenséget minden ilyen hatás nélkül is megfigyelnénk. Heckman Singerrel írt tanulmányaiban azt kutatta, hogy hogyan lehet e két hatást úgy elválasztani egymástól, hogy az minél kevésbé függjön sokszor ad hoc eloszlásbeli feltevésektől. Legfontosabb eredményeik mára az alkalmazott kutatás szerves részei lettek.

A módszertani kutatások mellett Heckman több tartalmi kérdésben is jelentőset alkotott. Az aktív munkaerő-piaci programok megfelelően mért hatásait összegezve Heckman, LaLonde és Smith (1999) megállapítják, hogy a legtöbb országban az ilyen programok hosszú távú hatása nulla. A kivételek közé a gondosan megtervezett és célzott programok tartoznak, amelyeknek viszont a költségei is jóval magasabbak. A latin-amerikai országokban a munkahelyek biztonságát célzó megszorítások foglalkoztatási hatásait elemezve Heckman és Pages (2000) kimutatta, hogy azok – az intézményi háttér részleteitől függően – az egyenlőtlenségek növelését és konzerválását eredményezik, és jelentősen csökkentik a foglalkoztatást. Ez a megállapítás szemben áll a munkagazdászok jó részének a nézeteivel, melyek szerint a munkaerő-piaci „merevségeknek” nincs kimutatható hatékonysági hatása. Heckman és Pages a jövedelmi egyenlőtlenségek és a jövedelmi biztonság

¹⁹ Heckman–Singer, 1984.

területén más, a munkaerő-piaci mechanizmusokkal nem összeütköző intézkedéseket javasol. A legutóbbi időben Heckman az emberek közötti egyenlőtlenségek újratermelődésének egyik legfontosabb dimenziója, az etnikai különbségek és azoknak a családi környezeten és az iskolarendszeren keresztül történő továbbörökítésének a mechanizmusai felé fordult. A rendkívül élénk vita legfrissebb eredményeit összefoglaló kötetben Heckman empirikus eredmények tömegével alátámasztott álláspontja szerint az egyenlőtlenségeket az iskolarendszer tovább növeli, de nem önmagában felelős érte, ezért az egyenlő esélyeket célzó társadalompolitikai programoknak komplexebb feladataik vannak (Heckman és Krueger, 2003).

James J. Heckman kimagasló eredményeket ért el a mikroökonometria területén. Ezen teljesítmények egy része kollegáival közös munkából, informális interakciókból, esetenként pedig komoly vitákból származott. A szakma fejlődésének ma is természetes velejárója, és így volt ez a hetvenes években is, hogy a legfontosabb eredmények gyakran közös gondolkodásból származtak, sokszor „benne voltak a levegőben”. Bizonyos eredmények számavételénél az elsőségi kérdéseket egyesek néha megkérdőjelezték. Az azonban vitathatatlan, hogy Heckman mindig a mikroökonometria leginkább égető kérdésein dolgozott, és fáradhatatlan munkája sok eredeti megoldást és termékeny vitákra indító tanulmányt hozott. Elhivatottsága és rendszerező gondolkodása képessé tette arra, hogy egyedülálló módon szintetizálja a legfontosabb eredményeket, mindig szem előtt tartva azok gyakorlati, társadalompolitikai relevanciáját. Aktív jelenléte és munkássága nemcsak a mérési módszerek kidolgozásában és technikai tökéletesítésében hozott – és hoz ma is – a szakma egésze számára jelentős eredményeket, de az empirikus problémák általános megközelítésében is mindenki számára követendő mércét állított. A közgazdasági Nobel-díj egy ilyen területet emelt ki, de valójában az egész életművet állította minden társadalomkutató elé követendő példaként.

Kézdi Gábor

Irodalom

Főbb művei

- 1974 Shadow Prices, Market Wages, and Labor Supply. *Econometrica*, 42 (4), 679–694. o.
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*, 5, 475–492. o.
1979 Sample Selection Bias as Specification Error. *Econometrica*, 47 (1), 153–162. o.
1980 (Társszerző: *MaCurdy, Thomas*) A Life Cycle Model of Female Labor Supply. *Review of Economics and Statistics*, 47, 47–74. o.
1984 (Társszerző: *Singer, B.*) A Method of Minimizing the Impact of Distributional Assumptions for Duration Data. *Econometrica*, 52, 271–320. o.

- 1985 (Társszerző: *Singer, B.*) Econometric Analysis of Longitudinal Data. *Griliches, Z. és Intrilligator, M.* (szerk.): Handbook of Econometrics. Vol. 2, North-Holland, 1689–1763. o.
- 1986 (Társszerző: *MaCurdy, Thomas*) Labor Econometrics. *Griliches, Z. és Intrilligator, M.* (szerk.): Handbook of Econometrics. Vol. 3, North-Holland, 1917–1977. o.
- 1990 (Társszerző: *Honoré, Bo*) The Empirical Content of the Roy Model. *Econometrica*, 58, 1121–1149. o.
- 1997 (Társszerzők: *Smith, J. és Clements, N.*) Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts. *Review of Economic Studies*, 64, 487–535. o.
- 1999 (Társszerzők: *Lalonde, Robert, J. és Smith, Jeffrey, A.*) The Economics and Econometrics of Active Labor Market Programs. *Asbenfelter, A. és Card, D.* (szerk.): Handbook of Labor Economics. Vol. 3A, Elsevier, 1865–2097. o.
- 2000 Microdata, Heterogeneity, and the Evaluation of Public Policy. Nobel Prize Lecture.
- 2003 (Társszerkesztő: *Krueger, A.*) Inequality in America: What Role for Human Capital Policy? MIT Press.

Válogatott bibliográfia

- Angrist, J.–Imbens, G.–Rubin, D.* (1996): Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91, 444–455. o.
- Asbenfelter, O.* (1978): Estimating the effect of training programs on earnings. *Review of Economics and Statistics*, 6 (1), 47–57. o.
- Lee, L-F.* (1979): Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables. *Econometrica*, 47 (4), 977–996. o.
- Gronau, R.* (1974): Wage-Comparisons – A Selectivity Bias. *Journal of Political Economy*, 82 (6), 1119–1143. o.
- Kézdí, G.* (2004): Az aktív foglalkoztatáspolitikai programok hatásvizsgálatának módszertani kérdései. MTA Közgazdaság-tudományi Intézet.
- Mincer, J.* (1962): Labor Force Participation of Married Women. *Lewis, H. G.* (szerk.): Aspects of Labor Economics. Princeton.
- Rees, A.* (1974): An Overview of the Labor Supply Results. *Journal of Human Resources*, 9, 158–180. o.
- Roy, A.* (1951): Some Thoughts on the Distribution of Earnings. *Oxford Economics Papers*, 3, 135–146. o.
- Quandt, R.* (1956): A Probabilistic Model of Consumer Behavior. *Quarterly Journal of Economics*, 70, 507–536. o.
- Tobin, J.* (1958): Estimation of Relationships For Limited Dependent Variables. *Econometrica*, 26, 24–36. o.
- Wilderquist, K.* (2003): A Failure to Communicate: The Labour Market Findings of the Negative Income Tax Experiments And Their Effects on Policy and Public Opinion. Mimeo, Basic Income European Network.
- Willis, R.–Rosen, S.* (1979): Education and Self-Selection. *Journal of Political Economy*, 87 (5/2), S7–S36.