

Valikoitumisvirhe tulosevaluaatiotutkimuksessa

Tuijnman, Albert 1988. *Valikoitumisvirhe tulosevaluaatiotutkimuksessa*. Aikuiskasvatus 8, 4, 11–13.

— *Artikkelissa tarkastellaan analyysimenetelmiä valikoitumisvirheen poistamiseksi aikuisopetusohjelmien vaikutuksia estimoitaessa. Aikuisopiskelijat eroavat yleensä monissa suhteissa niistä, jotka eivät osallistu aikuisopiskeluun. Tästä aiheutuu metodologisia ongelmia estimoitaessa opetuksen vaikutusta yksilöihin ja yhteiskuntaan. Opiskeluun valikoitumisesta aiheutuva virhelähteitä ei voida olla ottamatta huomioon tutkimuksessa.*

Johdanto

Useimmat empiiriset tutkimukset koulutuksen mahdollisista sosiaalisista, ammatillisista ja taloudellisista hyödyistä ovat käsitelleet vain formaalin koulutuksen vaikutuksia. Nämä tutkimukset ovat tyypillisimmillään keskittyneet siihen, missä määrin perinteinen koulumuotoinen tai nuorisoasteen koulutus kausaalisesti määrittää kriteerimuuttujissa havaittuja eroja.

Viime vuosina yhä useammat ovat osallistuneet aikuiskasvatukseen ja -koulutukseen. Siksi on tullut välttämättömäksi kehitellä tutkimusmenetelmiä, jotka kykenevät erottamaan formaalin koulutuksen ja koulun jälkeisen aikuiskasvatuksen nettovaikutukset toisistaan. Nuorisoasteen koulutuksen vaikutuksia sosiaalitaloudelliseen menestymiseen on toisin sanoen liioiteltu, koska ei ole kyetty kontrolloimaan muita, havaitsematta jääneitä inhimilliseen pääomaan investoimisen puolia. Näin ollen eräs empiirinen kysymys on, miten tunnistaa, mitata ja arvioida aikuiskasvatuksen "todelliset" tuotot ilman identifikaatio-, mitaus- ja spesifikaatiovirheitä (Tuijnman et al., 1988).

Valikoitumisvirheen ongelma

Estimoitaessa koulutuksen pysyväisvaikutuksia joudutaan tekemisiin lukuisien käsitteellisten ja metodolisten ongelmien kanssa. Eräs tärkeä kysymys on se, miten voidaan kontrolloida ne mahdolliset tekijät, jotka eivät johdu koulutuksesta, mutta saattavat vaikuttaa havaitun muutoksen ja koulutuksen väliseen kausaaliyhtey-

teen häiritsevästi. Toinen vaikeus on siinä, miten tunnistaa ne koulutuskokemuksien osatekijät, jotka edistävät tai ehkäisevät tiettyjen koulutustuotosten esiintymistä. Kolmas ongelmaryhmä syntyy siitä, kun estimoitujen parametrien tai vaikutusta osoittavien kertoimien perusteella tulkitaan ja tehdään johtopäätöksiä empiirisiksi tuloksiksi. Kun koulutuksen tuotosten tutkimukseen liittyy yleisesti erittäin suuria vaikeuksia, niin aivan erityisesti tämä koskee niitä tutkimuksia, joissa estimoidaan oppivelvollisuusopintojen jälkeen tulevan koulutuksen vaikutuksia.

On monimutkainen tehtävä evaluoida ja arvioida hyötyjä, joita koituu yksilölle ja koko yhteiskunnalle aikuiskoulutukseen investoimisesta, koska näille useimmiten vapaaehtoiseen osallistumiseen perustuville kurseille on ominaista tavoitteiden, sisältöjen ja organisaatiomuotojen moninaisuus ja se että ne on tarkoitettu palvelemaan mitä erilaisimpia kohderyhmiä. On esimerkiksi erittäin ongelmallista yleistää tutkimustuloksia, koska ei ole mitään takeita siitä, että tutkimus voidaan asianmukaisesti toistaa. Tässä suhteessa olisikin muistettava, että on yleensä virheellistä verrata toisiinsa standardisoituja polkukertomia tai jopa efektin kokoa, jotka on estimoitu tutkimuksissa, joiden kohteet, otokset ja keskeiset analyysimenetelmät poikkeavat huomattavasti toisistaan (Kiefer, 1979; Bowman, 1987).

Ne reaali maailman mekanismit, jotka vaikuttavat yksilön valintojen muotoutumiseen yleisesti ja päätökseen osallistua aikuiskoulutukseen erityisesti, vaikuttavat todennäköisesti myös tutkimuksen kohteena olevan otoksen tai populaation ominaisuuksiin. Koska on luultavaa, että osallistujat eroavat systemaattisesti

ei-osallistujista useissa suhteissa, on yritys selittää osallistumismuuttujien varianssia ja saada lasketuksi luotettavat vaikutusta osoittavat estimaanit altis valikoitumisvirheelle, näennäisefekteille ja muille harhaisuuksille, jotka aiheutuvat virheellisestä mallin spesifioinnista. Valikoitumisvirhe syntyy silloin, kun jotkin latentit tekijät, jotka vaikuttavat aikuiskoulutukseen osallistumiseen, korreloivat merkittävästi sellaisten latenttien tekijöiden kanssa, jotka vaikuttavat tuotosmuuttujiin, kuten todennäköisesti tapahtuu juuri silloin kun osallistuminen on otokseen valinnan kriteerinä. Nykyisin ne mahdolliset valikoitumisvirheet, jotka aiheutuvat tavallisen pienimmän neliösumman (OLS) menetelmän käytöstä estimoitaessa vaikutusta osoittavia parametreja peruskoulutuksen, aikuiskoulutuksen ja työpaikkakoulutuksen sekä sellaisten kriteerimuuttujien kuin ammattiaseman, työtyytyväisyyden ja ansioiden suhteen, tunnetaan laajasti ja on hyvin dokumentoitu (Willis ja Sherwin, 1979; Björklund and Moffitt, 1987).

Eräässä tuoreessa tutkimuksessa, jossa mitattiin ja ennustettiin jaksottaiskoulutukseen osallistumista Malmö -tutkimuksen seuranta-aineiston pohjalta, Tuijman ja Fägerlind (1988) osoittivat, millä tavoin ja missä määrin ruotsalaiseen aikuiskasvatukseen osallistujat eroavat osallistumattomista. Osallistujien ja osallistumattomien väliset erot olivat selkeimmät kognitiivisten kykyjen (mitattu 10- ja 20-vuotiaina), vanhempien suhtautumisessa lastensa koulunkäyntiin, vastaajien saaman muodollisen koulutuksen ja ammattiaseman suhteen. Tämän tutkimuksen tulokset osoittivat selvästi, että ruotsalaisten aikuisten rekrytoimisessa aikuiskoulutukseen tapahtuu valikoitumista. Samanlaisia tuloksia on raportoitu myös muissa Skandinavian maissa. Tämä merkitsee sitä, että aikuiskoulutuksen vaikutuksia tutkittaessa on käytettävä tutkimusasetelmaa, jonka avulla valikoituminen voidaan mitata ja sen vaikutus tuloksiin ainakin osittain poistaa.

Valikoitumisvirheen tutkimusmenetelmät

Siinä inhimillisen pääoman traditiossa, joka on tutkinut koulutuksen vaikutuksia ansiotuloihin, on Mincerin (1974) ja kumppanien työtä seuraten enenevästi keskitytty tutkimaan hyötyjä, jotka aiheutuvat muista tuottavuuden kannalta relevanteista työvoiman laatuun kohdistuvista investoinneista. Erityisesti 1970-luvun puolivälistä lähtien on ollut käytettävissä runsaasti kirjallisuutta työpaikka- ja muun aikuis-

koulutuksen yhteiskunnallisten ja taloudellisten tuottojen arvioinnista. Taloustieteilijöiden ja sosiologien tekemän tutkimuksen tiimoilla on valikoitumisvirheen problematiikkaan kiinnitetty vakavaa huomiota. Sen tuloksena tutkimusmenetelmät ovat kehittyneet varsin nopeasti. Valikoitumisvirheen estimoimiseksi ja osittaiseksi poistamiseksi on ollut käytettävissä standardimetodologia siitä lähtien kun mm. Ashenfelter (1978), Heckman (1978 ja 1979), Laurence et al. (1979), Barnow et al. (1980)), ja Gay ja Borus (1980) esittivät perustekniikat, ja Dye ja Antle (1984), Ashenfelter ja Card (1985), Heckman ja Sedlacek (1985), Lee ja Maddala (1985) sekä Björklund ja Moffitt (1987) kehittivät joukon edistyneempiä menetelytapoja. Viime vuosina on julkaistu useita katsauksia ja tutkimuksia, joissa on käytetty näitä menetelmiä (katso Bassi 1985; Heckman ja Robb 1985; LaLonde 1986; Ashenfelter 1987; Barnow 1987; Fraker ja Maynard 1987).

Perusmalli aikuiskasvatuksen vaikutuksen estimoimiseksi muodostuu koulutusohjelman toteuttamista ja ohjelman vaikutuksia esittävästä rakenneyhtälöstä (Cavin ja Maynard, 1985). Vaikutusta osoittavien kertoimien estimoiminen on ongelmallista, koska lukuisat muuttujat, jotka vaikuttavat koulutusohjelman tuotoksiin, ovat vahvassa riippuvuussuhteessa myös koulutusohjelman toteuttamista mittaaviin muuttujiin. Ongelmanratkaisun avain on siinä, että saadaan luotettavat estimaatit sekä koulutusta että sen toteuttamista ilmaisevista muuttujista, ja kytetään identifioimaan koulutukseen osallistumista esittävä yhtälö erillään koulutuksen vaikutusta kuvaavasta rakenneyhtälöstä. Valitettavasti valikoitumisesta johtuvaa virhettä ei voida kokonaan eliminoida kontrolloimalla sosiaalinen tausta, kognitiiviset kyvyt, koulutustausta jne. Tarvitaan sellainen tutkimusasetelma, joka suoraan mahdollistaa korreloivien virhevariانسsia aiheuttavien tekijöiden estimoinnin.

Kirjallisuudessa on esitetty erilaisia tutkimusasetelmia, jotka kykenevät jossain määrin ottamaan huomioon valikoitumisesta johtuvat mallien identifikaatio- ja spesifikaatiovirheet. Aikuiskasvatuksen evaluaatiotutkijat ovat käyttäneet sekä kokeellisia että ei-kokeellisia tutkimusasetelmia koulutusohjelmien vaikutuksia arvioitaessa. Kokeellisissa menettelytavoissa käytetään tavallisesti useita randomisoituja ryhmiä, jotka eroavat selkeästi toisistaan koe-käsittelyn suhteen. Tämän jälkeen ryhmien ominaisuuksia verrataan toisiinsa alku- ja loppumittauksien osalta. Eräs vaikea ongelma liittyy vertailuryhmän käyttöön. Monissa tutkimuksissa on tutkittu vain osallistujia. Koska jokainen ryhmä, joihin naiset ja miehet vapaaehtoisuuteen perustuvassa aikuiskasvatukseen osallistuvat, on tietyn tyyppinen, on esitetty

epäilyjä saatujen tulosten merkityksellisyydestä. Siksi jopa sellaisissa kokeellisissa asetelmissä, joissa on käytetty useita tutkimusryhmiä, koulutusohjelman vaikutusten estimointi saattaa jäädä harhaiseksi, ellei voida käyttää sopivaa kontrolliryhmää. Toisin kuin niissä järjestelyissä, jotka ovat ominaisia kokeelliselle asetelmalle, ei-kokeellisessa tutkimuksessa saatavat ohjelman vaikutusta osoittavat estimaatit riippuvat siitä tavasta, miten osallistumista ja vaikutusta esittävät rakenneyhtälöt on identifioitu.

Useimmissa ei-kokeellisissa evaluaatiotutkimuksissa käytetään alku- ja loppumittausta tai sille vaihtoehdoisena aikasarja-asetelmia. Ekonometrisissa tai rakenneyhtälömalleissa tutkimusaineisto on usein hankittu poikkileikkausasetelmalla. Harvemmissä tapauksissa on käytettävissä seuranta-aineistoja. Jos on käytettävissä sekä ennen koulutusohjelmaa että sen jälkeen tehdyt mittaukset, on teoriassa olemassa hyvät mahdollisuudet valikoitumisvirheen poistamiseen. Jos nimittäin tutkimuksen otos on valittu siten, että osallistumiskriteeriä ei ole käytetty, voidaan valikoitumisvirhe välttää. Mallien spesifikaatiovirheet, jotka johtuvat relevanttien muuttujien pois jättämisestä, edellyttävät sellaiset olisivat olleet käytettävissä ja mitattavissa virheettä, voidaan ottaa huomioon sellaisessa ei-kokeellisessa asetelmassa, jossa käytetään alku- ja loppumittausta.

Tämän lisäksi on olemassa vielä eräs spesifikaatiovirhe, joka johtuu siitä, että latenttien muuttujien virhetermien korreloidessa keskenään merkitsevästi, tavallinen pienimmän neliosumman menetelmä (OLS) on taipuvainen tuottamaan epäluotettavat koulutusohjelman vaikutusta ilmaisevat estimaatit. LaLonden (1986) mukaan tavanomaiset ekonometriset ja polkuanalyttiset menettelytavat aikuiskoulutuksen vaikutuksia estimoitaessa aliarvioivat näitä vaikutuksia silloin, kun latentit muuttujat korreloivat negatiivisesti keskenään ja yliarvioivat silloin kun latentit muuttujat korreloivat positiivisesti. Menetelmiä, joilla virhetermien korreloimisesta aiheutuvat vääristymät voidaan poistaa, ovat esitelleet muiden muassa Jöreskog ja Sörbom (1988). Valitettavasti näiden menetelmien käyttö tuottaa tavallisesti epäluotettavat keskivirheen estimaatit. Lyhyesti sanoen koulun jälkeisen opiskelun ei-kokeellinen evaluaatio on vaativa tehtävä sekä tutkimuksessa tarvittavan aineiston ominaisuuksien että niiden vaikeuksien vuoksi, joita liittyy asianmukaisten analyysitekniikoiden menestykselliseen käyttöön.

Johtopäätös

Koska aikuiskoulutusohjelmiin osallistujat yleensä eroavat systemaattisesti useissa suhteissa osallistumattomista, aiheutuu tästä metodologisia ongelmia estimoitaessa näiden ohjelmien vaikutuksia yksilöihin ja yhteiskuntaan. Enää ei voida olla ottamatta huomioon niitä mahdollisia virhelähteitä, jotka aiheutuvat valikoitumisesta ohjelmiin. Tässä artikkelissa tarkasteltiin suhteellisen pientä osuutta siitä nykyisestä kirjallisuudesta, jossa esitellään analyysimenetelmiä valikoitumisvirheiden poistamiseksi ohjelmien vaikutuksia estimoitaessa.

Kirjallisuus

- Ashenfelter, O. 1978. Estimating the Effect of Training Programs on Earning. *Review of Economics and Statistics*, 60, 47—57.
- Ashenfelter, O. 1987. The Case for Evaluating Training Programs with Randomized Trials. *Economics of Education Review*, 6, 333—338.
- Ashenfelter, O. & Card, D. 1985. Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *Review of Economics and Statistics*, 67, 648—640.
- Barnow, B.S. 1987. The Impact of CETA Programs on Earnings: A Review of the Literature. *Journal of Human Resources*, 22, 157—193.
- Barnow, B., Cain, G. & Goldberger, A. 1980. Issues in the Analysis of Selection Bias. In E. Stromsdorfer & G. Farcas (Eds.), *Evaluation Studies Review Annual*, Vol. 5, pp. Beverly Hills: Sage.
- Bassi, L.J. 1985. Evaluating Alternative Job Creation Strategies. *Economic Inquiry*, 23, 671—690.
- Behrman, J.R. 1987. Schooling and Other Human Capital Investments: Can the Effects be Identified? *Economics of Education Review*, 6, 301—305.
- Björklund, A. & Moffitt, R. 1987. The Estimation of Wage Gains and Welfare Gains in Self-Selection Models. *Review of Economics and Statistics*, 69, 42—49.
- Bowman, M.J. 1987. The Importance of Examining Cohort Uniqueness in the Formulation of Human Investment Policies". *Economics of Education Review*, 6, 67—79.
- Cavin, E. & Maynard, R. 1985. Short-term Indicators of Employment Program Performance: Evidence from the supported Work Program. *Journal of Human Resources*, 20, 331—345.
- Dye, R.A. & Antle, R. 1984. Self-selection via Fringe Benefits. *Journal of Labor Economics*, 2, 388—411.
- Fraker, T. & Maynard, R. 1987. The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs. *Journal of Human Resources*, 22, 194—227.

- Gay, R.S. & M.E. Borus (1980). Validating Performance Indicators for Employment and Training Programs. *Journal of Human Resources*, 15, 29—48.
- Heckman, J.J. 1978. Dummy Endogenous Variables in a Simultaneous Equations System. *Econometrica*, 46, 931-959.
- Heckman, J.J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*, 47, 153—161.
- Heckman, J.J. & Robb, & Robb, R. 1985. Alternative Methods for Evaluating the Impact of Interventions. In J.J. Heckman & B. Singer (Eds.), *Longitudinal Analysis of Labor Market Data*, pp. Cambridge: Cambridge University Press.
- Heckman, J.J. & Sedlacek, G. 1985. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93, 6: 1077—1125.
- Jöreskog, K.G. and Sörbom, D. 1988. LISREL. Analysis of Linear Structural Relationships. Version 7.2. Uppsala: University of Uppsala.
- Kiefer, N. 1979. Population Heterogeneity and Inference from Panel Data on the Effects of Vocational Training. *Journal of Political Economy*, 87, S213—S26.
- LaLonde, R.J. 1986. Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review*, 76, 604—620.
- Laurence, K., Lee, L-f., Maddala, G.S. & Trost, R. 1979. Returns to College Education: An Investigation of Self-Selection Bias Based on Project Talent Data. *International Economic Review*, 20, 775—789.
- Lee, L-F. and Maddala, G.S. 1985. The Common Structure of Tests for Selectivity Bias, Serial Correlation, Heteroscedesticity and Non-Normality in the Tobit Model. *International Economic Review*, 26, 1—20.
- Mincer, J. 1974. *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research and Columbia University Press.
- Tuijnman, A. and Fägerlind, I. 1988. Measuring and Predicting Participation in Lifelong Education Using Longitudinal Data. *Scandinavian Journal of Educational Research*, 32 (winter).
- Tuijnman, A., Chinapah, V. & Fägerlind, I. 1988. Adult Education and Earnings: A 45-Year Longitudinal Study of 834 Swedish Men. *Economics of Education Review*, 7 (winter).
- Willis, R.J. & Sherwin, R. 1979. Education and Self-Selection. *Journal of Political Economy*, 87 (5, Part 2), S7—S36.

AIKUISKASVATUS

The Finnish Journal of Adult Education

Vol. 8,4/88

ISSN 0358-6197

Summary

Tuijnman, Albert. 1988. Selection bias in program impact evaluation reserach.

— The article deals with methods of analysis to eliminate selection bias when estimating the effects of Adult Education programs. Adult students generally differ from nonparticipants in many ways. This causes methodological problems when estimating the effect that teaching has on individuals and society. The sources of error that may result from the operation of selection bias on study can not be ignored in research.