

Taloustieteen Nobel luonnollisten koeasetelmien kehittäjille ja hyödyntäjille

Mika Kortelainen

Vuoden 2021 taloustieteen Nobel-palkinto jaettiin David Cardin, Joshua Angristin ja Guido Imbensin kesken. Palkittujen työ on nostanut kausaalipäätelyn ja luonnollisten koeasetelmien merkitystä taloustieteessä. Heidän työllään on ollut myös laaja vaikutus taloustieteen tutkimukseen ja opetukseen sekä heijastumavaikutuksia muihin tieteenaloihin. Tärkeisiin kausaalikysymyksiin vastaaminen on myös lisännyt taloustieteellisen tutkimuksen yhteiskunnallista merkitystä. Keskityn kirjoituksessani Angristin ja Imbensin ansioihin empiirisen taloustieteellisen tutkimuksen kehittämisessä sekä erityisesti heidän merkittävimpiin kontribuutioihinsa syy-seuraussuhteiden arviointiin käytettävissä luonnollisissa koeasetelmissä ja niihin liittyvissä ekonometrisissa menetelmissä.

Nobel-palkittuja David Cardia, Joshua Angristia ja Guido Imbensia yhdistää systemaattinen pyrkimys vastata 1980-luvulla ja sitä ennen tehtyjen empiiristen taloustieteellisten tutkimusten keskeisiin heikkouksiin. Vaikka taloustiede on ollut jo hyvin pitkään empiirinen tieteenala, soveltavissa tutkimuksissa hyödynnetyt tutkimusasetelmat ja -menetelmät olivat pääosin hyvin erilaisia yli 30 vuotta sitten kuin ne ovat nykyisin. Vielä 1980-luvulla ekonometristen menetelmien keskeisiä oletuksia ei arvioitu sovelluksissa systemaattisesti ja tutkimuksissa tehdyt analyysit olivat lukijoille tyypillisesti varsin läpinäkymättömiä. Käytännössä empiirisiin tutkimustuloksiin suhtauduttiin yleensä skeptisesti, koska tulokset saattoivat olla hyvin herkkiä oletuksille ja erilaisille tutkijoiden analyyseissa

tekemille valinnoille (Angrist ja Pischke 2010). Tästä huolimatta keskeisimmät taloustieteen tutkimuskysymykset ovat olleet jo hyvin pitkään ns. kausaalikysymyksiä liittyen talouden erilaisiin syy-seuraussuhteisiin. Palkittujen tutkijoiden merkittävänä ansiona voidaankin nähdä uusien tutkimuskysymysten sijasta uudenlaisten, taloustieteessä aiemmin vähän hyödynnetyjen tutkimusasetelmien ja ekonometristen menetelmien kehittäminen ja hyödyntäminen syy-seuraussuhteiden analysointiin.

Kontrolloituja satunnaiskokeita on pidetty perinteisesti eri tieteenaloilla luotettavimpina tutkimusasetelmina syy-seuraussuhteiden tutkimiseen. Näin on myös taloustieteessä, ja tästä hyvänä osoituksena on satunnaistettujen koeasetelmien huomattava yleistyminen

KTT Mika Kortelainen (mika.kortelainen@utu.fi), on Turun yliopiston terveystaloustieteen professori ja johtava tutkija Valtion taloudellisessa tutkimuskeskuksessa (VATT).

tietyillä taloustieteen osa-alueilla kuten kehitystaloustieteessä ja koulutuksen taloustieteessä.¹ Satunnaiskokeiden merkittävimpana etuna voidaan pitää yksilöiden tai vertailtavien havaintojen satunnaistamista keskenään samanlaisiin koe- ja vertailuryhmään. Satunnaistamisen seurauksena koe- ja vertailuryhmä poikkeavat vain tutkittavan toimenpiteen (*treatment*) suhteen ja näin voidaan kontrolloida muut tutkittavaan ilmiöön vaikuttavat häiritsevät tekijät. Taloustieteessä hyvin moniin tutkimuskysymyksiin ei voida kuitenkaan hyödyntää kontrolloituja satunnaiskokeita muun muassa eettisistä tai käytännöllistä syistä johtuen. Esimerkiksi tuloverotuksen tai korkeakoulutuksen vaikutuksia tutkittaessa ei ole mahdollista arpoa erilaisia veroasteita tai pääsyä korkeakouluun eri ihmisille. Tästä johtuen tarvitaan muita mahdollisimman luotettavia tutkimusasetelmia syy-seuraussuhteiden tutkimiseen.

1. Luonnolliset koeasetelmat

Luonnollisia koeasetelmilla tarkoitetaan asetelmia, joissa yksilöiden (tai muiden havaintoyksikköjen) valikoituminen koe- ja vertailuryhmään määräytyy mekanismilla, joka tutkittavien yksilöiden näkökulmasta muistuttaa satunnaistamista. Usein luonnollinen koeasetelma muodostuu jonkin politiikkamuutoksen tai intervention seurauksena, mutta myös erilaiset sattumat (esim. syntymäajankohta) ja yhteiskunnan säännöt voivat luoda tutkimusasetelmia, jossa osa henkilöistä tai yrityksistä joutuu

tietyn toimenpiteen tai käsittelyn kohteeksi lähes satunnaisesti. Luonnollisia koeasetelmia on hyödynnetty muilla tieteenaloilla kuten epidemiologiassa ja psykologiassa jo paljon ennen 1980-lukua (esim. Snow 1855; Thistleweite ja Cambell 1960; Campbell 1969), mutta niin taloustieteessä kuin muillakin tieteenaloilla niiden rooli empiirisessä tutkimuksessa on ollut käytännössä varsin vähäinen ennen 2000-lukua. Tähän on osittain syynä menetelmien asteittainen kehittyminen, mutta tärkeä tekijä on ollut myös uskottavien empiiristen sovellusten yleistyminen vähitellen 1980-luvulta lähtien.

Ensimmäisten luonnollisia koeasetelmia hyödyntävien taloustieteellisten tutkimusten rooli oli keskeinen sekä menetelmien myöhemmän kehittämisen että niiden soveltamisen näkökulmasta. Yksi ensimmäisistä ja tärkeimmistä sovelluksista oli Angristin (1990) tutkimus Vietnamin sodan aikaisen asepalveluksen vaikutuksesta myöhempään työuraan ja ansioihin Yhdysvalloissa. Tutkimusasetelma perustui asepalveluksen todennäköisyyden satunnaiseen vaihteluun, joka syntyi syntymäajankohtaan perustuvan kutsuntajärjestyksen arvontan seurauksena. Kyseessä oli ensimmäinen sovellus, jossa luonnollista koeasetelmaa, tässä tapauksessa kutsuntajärjestyksen arvontatulosta, hyödynnettiin ns. instrumenttimuuttujana mielenkiinnon kohteena olevalle toimenpiteelle (asepalvelus)².

¹ *Abhijit Banerjee, Esther Duflo ja Michael Kremer saivat taloustieteen Nobel-palkinnon vuonna 2019 satunnaisten koeasetelmien kehittämisestä ja hyödyntämisestä erityisesti kehitystaloustieteessä.*

² *Angrist ym. (1996) käyttävät samaa tutkimusasetelmaa tarkastellessaan Vietnamin sodan aikaisen asepalveluksen vaikutusta kuolleisuuteen. Aiemmin Hearst ym. (1986) olivat myös hyödyntäneet tätä asetelmaa tutkiessaan arvontastatuksen vaikutusta kuolleisuuteen. Kyseisessä tutkimuksessa ei kuitenkaan hyödynnety instrumenttimuuttujamenetelmää, vaan mielenkiinnon kohteena oli arvontastatuksen suora vaikutus kuolleisuuteen (intention-to-treatment effect, ITT).*

Instrumenttimuuttujamenetelmää oli kehitetty ja käytetty taloustieteestä jo 1920-luvun lopulta lähtien, mutta aiemmat tutkimukset eivät olleet hyödyntäneet satunnaistamista tai luonnollisia koeasetelmia instrumentteina. Ennemminkin instrumenttimuuttujamenetelmää oli käytetty aiemmin varsin mekaanisesti kiinnittämättä tarpeeksi huomiota menetelmän varsin vaativiin taustaoletuksiin. Uudenlaisen sovelluksen ohella Angristin satunnaistamista hyödyntänyt tutkimus kannusti myöhemmin Angristia ja Imbensia kehittämään instrumenttimuuttujamenetelmän tulkintaa satunnaistettujen ja luonnollisten koeasetelmien yhteydessä.

Angrist ja Krueger (1991) tutkivat puolestaan koulutuksen vaikutusta palkkoihin hyödyntämällä syntymäajankohtaan liittyvää luonnollista koeasetelmaa. Yleisesti on hankalaa tutkia koulutuksen vaikutuksia tuloihin, koska korkeammin koulutetut ovat keskimäärin erilaisia ihmisiä kuin vähemmän koulutetut jo ennen koulutusta. Angrist ja Krueger pyrkivät ratkaisemaan tämän haastavan valikoitumisongelman vertaamalla alku- ja loppuvuodesta syntyneitä amerikkalaisia toisiinsa. He hyödynsivät havaintoa siitä, että yhdysvaltalaisessa järjestelmässä alkuvuodesta syntyneet voivat lopettaa koulunkäynnin 16-vuotissyntymäpäivänään. Tästä johtuen alkuvuodesta syntyneet pystyvät lopettamaan koulunkäynnin lähes vuotta aiemmin kuin loppuvuonna syntyneet ja syntymäajankohdalla on vaikutusta koulutuksen määrään. Tutkimuksessa syntymäajankohtaa hyödynnettiin instrumenttina tai kvasisatunnaisena vaihteluna koulutuksen määrälle, ja kyseisen asetelman avulla arvioitiin koulutuksen kausaali vaikutusta tuloihin. Vaikka kyseessä ei ole täysin ongelmaton instrumenttimuuttuja validilta instrumentilta vaadittavien oletusten

näkökulmasta, tutkimus kannusti muita tutkijoita arvioimaan instrumenttiehtoja tarkemmin ja tekemään aiheeseen liittyvää menetelmätutkimusta (mm. heikoista instrumenteista). Angristin ja Kruegerin tutkimus toimi tällä tavoin tärkeänä tienraivaajana luonnollisten koeasetelmien laajemman hyödyntämisen ja kehittämisen näkökulmasta.

2. Instrumenttimuuttujaan perustuva kausaalipäätteleminen

Angristin ja Imbensin keskeinen kontribuutio liittyy syy-seuraussuhteiden arviointiin käytävien ekonometristen tai tilastollisten menetelmien kehittämiseen. Heidän kehittämillään menetelmillä voidaan paremmin ymmärtää ja tulkita sekä kontrolloitujen satunnaiskokeiden että luonnollisten koeasetelmien tuloksia. Uraauurtavassa ja paljon viitatussa artikkelissa Imbens ja Angrist (1994) osoittavat miten instrumenttimuuttujamenetelmää voidaan hyödyntää tietyn toimenpiteen vaikutuksen estimointiin tilanteissa, jossa instrumenttimuuttuja toteuttaa tietyt minimaaliset ja usein empiirisesti uskottavat ehdot tai identifikaatiooletukset.

Artikkelissaan Imbens ja Angrist tarkastelevat yleistä viitekehikkoa, jossa toimenpiteen vaikutuksen sallitaan olevan heterogeeninen tutkittavassa populaatiossa. Esimerkiksi koulutuksen tuottojen tai sen vaikutuksen palkkaan voidaan monista eri syistä johtuen olettaa eroavan yksilöiden kesken. Heterogeenisuuden huomioimisella oli olennaista merkitystä myös Angristin (1990) ja Angristin ja Kruegerin (1991) tutkimuksissa, mikä osaltaan motivoi yleisemmän viitekehikon kehittämistä. Yleisemmässä tapauksessa koeasetelman tai luonnollisen

koeasetelman komplianssi eli toimenpiteen noudattaminen on myös todennäköisesti epätäydellistä (*incomplete compliance*). Esimerkiksi Vietnamin sotaan osallistui vain osa kutsuntajärjestyksen perusteella sinne arvoituista muun muassa terveyteen ja koulunkäyntiin liittyvien syiden takia (*never-takers*). Lisäksi osa sotaan osallistuneista oli vapaaehtoisia, ja he olisivat osallistuneet riippumatta arvonnän tuloksesta (*always-takers*).

Imbensin ja Angristin tutkimuksessa tarkastellaan kausaalipäätelyyn tarvittavia identifikaatioehtoja viitekehikossa, jossa toimenpiteen oletetaan vaikuttavan heterogeenisesti yksilöihin sekä sallitaan toimenpiteen epätäydellinen noudattaminen. Heidän yleinen viitekehyksensä perustuu ns. potentiaalisten tulemien malliin (*potential outcome model*), jota oli kehitetty ja hyödynnetty aiemmin tilastotieteessä sekä satunnaisten kokeiden (Neyman 1923) että havaintoaineistojen (Rubin 1974) tapauksissa. Potentiaalisten tulemien mallissa heterogeenisuus huomioidaan tarkastelemalla toimenpiteen vaikutuksia yksilötasolla. Kausaalivaikutuksia ei voida kuitenkaan koskaan arvioida yksilötasolla, koska yksilö havaitaan aina vain yhdessä maailmantilassa, joko toimenpiteen kohteena olevassa koeryhmässä tai vertailuryhmässä. Tämän ns. fundamentaalisen identifikaatio-ongelman takia toimenpiteen vaikutusten arviointi perustuu yksilötason sijasta populaation tai osapopulaation keskiarvoihin.

Potentiaalisten tulemien mallissa mielenkiinnon kohteena on mekanismi, joka määrittää toimenpiteeseen tai koeryhmään osallistumisen. Satunnaistamisen sijasta Imbens ja Angrist yhdistivät toimenpiteeseen osallistumisen instrumenttimuuttujaan tai yleisemmin IV-menetelmäkehikkoon. Tarkastellussa mallissa

instrumentin vaaditaan olevan (i) riippumaton potentiaalisista tulemistä (riippumattomuus), (ii) vaikuttavan toimenpiteeseen (relevanttius) sekä (iii) vaikuttavan tutkittavaan tulemaan vain tarkastellun toimenpiteen kautta (ns. pois-sulkurajoitus). Lisäksi instrumentin oletetaan (iv) vaikuttavan toimenpiteeseen yhdensuuntaisesti, ts. joko positiivisesti tai negatiivisesti (monotonisuus). Instrumenttina voidaan hyödyntää sekä satunnaistamisen että luonnollisen koeasetelman luomaa satunnaisvaihtelua mielenkiinnon kohteena olevaan toimenpiteeseen.

Imbens ja Angrist (1994) osoittavat, että yleisessä heterogeenisten vaikutusten tapauksessa, jossa sallitaan yksilöiden vapaaehtoinen osallistuminen toimenpiteeseen, tutkija voi estimoida yllä mainituilla oletuksilla (i)-(iv) ns. paikallisen keskimääräisen vaikutuksen (*local average treatment effect*, LATE). LATE-parametrin lokaalisuus johtuu siitä, että oletuksilla (i)-(iv) vaikutus voidaan identifioida vain niille henkilöille, joihin instrumentilla on vaikutusta tai jotka noudattavat asetelmaa (*compliers*)³. LATE-parametri ei ole informatiivinen muille henkilöille, koska arviointi perustuu instrumentin aiheuttamaan satunnaisvaihteluun toimenpiteessä. Esimerkiksi Angristin käyttämässä kutsuntajärjestyksen tulosta hyödyntävässä IV-asetelmassa asepalveluksen vaikutus voidaan arvioida vain niille henkilöille, joihin arvonnän tuloksella oli vaikutusta. Asepalveluksen vaikutus vapaaehtoisille tai keskimäärin voi kuitenkin poiketa LATE-parametrista.

Yleisemmin LATE-parametri voi heterogeenisten vaikutusten ja epätäydellisen

³ Englanninkielinen termi 'compliers' on yleisesti käytössä kirjallisuudessa ja viittaa osapopulaatioon, joihin instrumentilla on vaikutusta. Suomeksi voitaisiin puhua esimerkiksi noudattajista.

komplianssin tapauksessa poiketa toimenpiteen keskimääräisestä vaikutuksesta tutkittavassa populaatiossa (*average treatment effect*, ATE), josta ollaan yleensä kiinnostuneita esimerkiksi satunnaiskokeissa. Toisaalta Imbensin ja Angristin tulokset myös osoittavat, että tutkija voi estimoida satunnaiskokeista vain LATE-parametrin silloin kun toimenpiteen noudattaminen ei ole täydellistä tai kun toimenpide kohdentuu vain osaan koeryhmään arvotuista. Tämä voi olla rajoittavaa, kun tavoitteena on arvioida tietyn politiikan tai toimenpiteen keskimääräisiä vaikutuksia yhteiskunnassa. Monissa sovelluksissa politiikkarelevantti kysymys voi kuitenkin liittyä nimenomaan LATE-parametriin.

3. LATE-parametri ja kvasikokeelliset asetelmat

Imbensin ja Angristin (1994) esittämällä viitekehikolla sekä myöhemmillä IV-menetelmään liittyvillä menetelmäajennuksilla (esim. Angrist ym. 1996; Angrist ja Imbens 1995; Imbens ja Rubin 1997) on ollut keskeinen merkitys sovelluksissa, joissa on hyödynnetty erilaisia kvasikokeellisia instrumenttimuuttujia. Esimerkiksi tuomareiden ja sosiaalityöntekijöiden harkinnanvaraisiin päätöksiin perustuvat instrumentit ovat olleet suosittuja arvioitaessa erilaisten päätösten kuten rikostuomion tai lasten huostaanoton vaikutuksia mielenkiinnon kohteena oleviin tulemiin. Ideana näissä asetelmissa on hyödyntää yksilöiden kannalta lähes satunnaista vaihtelua tuomareiden (tai muiden päätöstekijöiden) määräytymisessä ja toisaalta tuomareiden välistä vaihtelua heille ominaisessa tuomioiden lievydessä (*judge leniency design*).

LATE-parametri on ollut keskeisessä roolissa näiden asetelmien kehityksessä ja tulkinnessa.

Imbensin ja Angristin viitekehikkoa tai LATE-parametria on hyödynnetty myös tutkittaessa esimerkiksi erilaisten kvasikokeellisten asetelmien kuten regressioepäjatkuvuusasetelman (*regression discontinuity design*, RDD) vaatimia identifikaatio-oletuksia (Hahn ym. 2001). Regressioepäjatkuvuusasetelmalla viitataan asetelmaan, jossa jonkin säännön tai tapahtuman seurauksena tutkittavat yksilöt päätyvät niukasti ja lähes satunnaisesti koe- ja vertailuryhmään. Erityisesti LATE-parametri on edesauttanut RDD-asetelmiin perustuvien tutkimusten tulointa ja toisaalta parantanut ymmärrystä menetelmään liittyvistä rajoitteista. RDD-asetelma on nykyisin yksi sovelletuimmista luonnollista koeasetelmista taloustieteessä ja tämän vuoden nobelistien työllä on ollut merkitystä myös sen käytön yleistymisessä⁴.

RDD ei ole kuitenkaan ainoa esimerkki nykyisin soveltavissa tutkimuksissa suosituista kvasikokeellista menetelmistä, jotka ovat hyödyntäneet potentiaalisten tulemien mallia tai Angristin ja Imbensin viitekehikkoa. Uudemmassa *difference-in-differences* (DID) -menetelmäkirjallisuudessa mielenkiinnon kohteena on myös toimenpiteen keskimääräisten vaikutusten arviointi yleisemmässä heterogeeniset vaikutukset sallivassa asetelmassa (de Chaisemartin ja D'Haultfoeuille 2020; Callaway and Sant'Anna 2021). Lisäksi vastaavanlaista potentiaalisten tulemien malliin perustuvaa

⁴ RDD-sovelluksissa hyödynnetään usein instrumenttimuuttujia tai ns. sumeaa regressioepäjatkuvuusasetelmaa (*fuzzy regression discontinuity design*). Angristin ja Lavyn (1999) tutkimus luokkakoon vaikutuksista oppimistuloksiin oli ensimmäinen kyseisen menetelmän ja yleensäkin ensimmäisiä RDD-menetelmän sovelluksia taloustieteessä.

menetelmäkehikkoa on sovellettu muun muassa synteettiseen kontrollimenetelmään (Abadie ym. 2010; Abadie 2021) ja ns. kulmapistemenetelmään (Card ym. 2015, 2017).

Kaiken kaikkiaan LATE-parametrin sekä Imbensin ja Angristin (1994) artikkelissa esitetyn menetelmäkehikon avulla on opittu paremmin ymmärtämään ja tulkitsemaan sekä satunnaistettujen että luonnollisten koeasetelmien tuottamia tuloksia erilaisissa sovelluksissa. Kyseisillä menetelmillä on ollut myös merkittävä vaikutus siihen, miten kvasikokeellisia asetelmia nykyisin hyödynnetään taloustieteellisessä tutkimuksessa.

4. Vaikutukset empiiriseen tutkimukseen taloustieteessä ja muualla

Miksi luonnollisten koeasetelmien sekä niihin liittyvien ekonometristen menetelmien kehittämisellä on sitten ollut niin merkittävä vaikutus empiiriseen tutkimukseen taloustieteessä? Taustalla on useita syitä, joita pyrin hahmottelemaan seuraavaksi.

Tärkein peruste liittyy soveltavien tutkimusten uskottavuuden selkeään parantumiseen luonnollisten koeasetelmien laajemman hyödyntämisen kautta. Vaikka monet luonnolliset koeasetelmat (kuten regressioepäjätkuvuusasetelma) olivat jo olemassa tai niitä oli kehitetty muilla tieteenaloilla, niiden soveltaminen oli hyvin vähäistä 1990-luvun alkuun asti. Angristin, Cardin ja Imbensin sekä monet muut luonnollisia koeasetelmia hyödyntäneet pioneeritutkimukset näyttivät esimerkkiä luonnollisten koeasetelmien eduista erilaisissa sovelluksissa. Vaikka samoja kysymyksiä, kuten koulutuksen vaikutusta tuloihin, oli tutkittu

aiemmin erilaisilla ekonometrisillä menetelmillä, tutkimusasetelmat alkoivat saada vähitellen enemmän painoarvoa empiirisissä sovelluksissa. Tästä muutoksesta taloustieteen empiirisessä tutkimuksessa on käytetty nimeä **uskottavuuden vallankumous** (*credibility revolution*), jonka tienraivaajina pidetään juuri Cardia, Angristia ja Imbensia. Yleensä sillä viitataan empiirisen tutkimuksen laadun parantumiseen uskottavampien tutkimusasetelmien seurauksena (Angrist ja Pischke 2010). Muutos alkoi työn taloustieteestä 1990-luvulla, mutta 2000-luvun aikana luonnollisia koeasetelmia on vähitellen alettu hyödyntää yhä enemmän taloustieteen eri sovellusalueilla, kuten julkisessa taloudessa, terveystaloustieteessä ja ympäristötaloustieteessä.

Luonnollisten koeasetelmien yleistymisen toisena tärkeänä syynä on niiden yllättävänkin laaja soveltuvuus moniin tilanteisiin. Aiemmin ajateltiin, että luonnollisia koeasetelmia voidaan hyödyntää vain hyvin rajatussa määrässä sovelluksia. Tähän ajatteluun oli syynä se, että luonnollisten koeasetelmien löytäminen vaatii usein tutkijoilta paljon instituutioihin ja järjestelmiin liittyvää selvitys- ja tutkimustyötä, eikä asetelmat ole useinkaan mitenkään ilmeisiä. Luonnollisten koeasetelmien laajentunut hyödyntäminen eri taloustieteen osa-alueilla on kuitenkin paljastunut, että asetelmat soveltuvat moniin taloustieteen sovelluksiin erittäin hyvin. Toisaalta niiden käyttöön liittyy rajoitteita esimerkiksi erilaisissa makrotaloudellisissa sovelluksissa, mikä osaltaan selittää asetelmien vähäisempää hyödyntämistä makrotaloustieteessä.

Yksi tärkeä käytännön syy luonnollisten koeasetelmien suosioon on myös se, että asetelmien perusidea on yleensä hyvin yksinkertainen ja selitettävissä myös muille kuin tutkijoille. Kuten

Angrist ja Pischke (2010) artikkelissaan toteavat, tyyppillisesti hyvät asetelmat mahdollistavat empiiristen menetelmien selittämisen yksinkertaisella tavalla sekä tuloksien suoraviivaisen esittämisen⁵. Esimerkiksi monissa luonnollisia koeasetelmia hyödyntävissä sovelluksissa on vain kaksi ryhmää, toimenpiteen kohteena oleva koeryhmä sekä verrokkina toimiva kontrolliryhmä. Tällainen asetelma mahdollistaa yksinkertaiset ja intuitiiviset graafiset vertailut, joiden ymmärtäminen ei edellytä syvällistä osaamista tilastotieteestä tai ekonometriasta. Tästä johtuen asetelmia ja niihin perustuvia tutkimustuloksia voidaan esitellä ymmärrettävällä tavalla niin ensimmäisen vuoden yliopisto-opiskelijoille, medialle kuin virkamiehillekin. Asetelmien yksinkertaisuus on tehnyt tutkimuksista myös paljon läpinäkyvämpiä ja siten osaltaan mahdollistanut tutkimusten toistettavuutta ja uudelleenanalysointia kustannustehokkaasti. Lisäksi asetelmien keskeisten oletusten läpinäkyvyys on mahdollistanut oletusten uskottavuuden arvioinnin sekä erilaiset herkkyystarkastelut.

Tärkeä tekijä luonnollisten koeasetelmien yleistymisessä taloustieteessä ovat myös kausaalipäätelyyn liittyvät oppikirjat. Erityisesti Angrist ja Pischke (2008) oli tärkeä teos ja se lisäsi luonnollisiin koeasetelmiin liittyvien ekonometristen menetelmien ymmärrystä niin opiskelijoiden kuin tutkijoidenkin keskuudessa. Nähdäkseni harvalla 2000-luvulla julkaistulla oppikirjalla on ollut niin merkittävä vaikutus sekä taloustieteen opetukseen että

tutkimukseen kuin kyseisellä kirjalla on ollut⁶. Varsinkin soveltavammat mikroekonometrikan kurssit perustuvat nykyisin yhä useammin Angristin ja Pischken kirjaan tai yleisemmin kausaalipäätelyssä hyödynnettäviin tutkimusasetelmiin ja -menetelmiin. Viime vuosina aiheesta on julkaistu lukuisia uusia kirjoja, jotka käsittelevät myös menetelmien kehitystä 2000-luvun aikana (Angrist ja Pischke 2014; Cunningham 2021; Huntington-Klein 2021). Ekonometrian kurssien ohella luonnollisia koeasetelmia ja kausaalipäätelyä hyödyntäviä sovelluksia on nostettu esiin yhä enemmän myös muilla taloustieteen kursseilla, jopa taloustieteen johdanto- tai peruskursseilla. Tällä tavoin asetelmat ja menetelmät ovat tulleet ainakin osittain tutuksi jo taloustieteen opintojen alkuvaiheessa tai ennen syventäviä opintoja.

Luonnollisten koeasetelmien hyödyntäminen on levinnyt myös taloustieteen ulkopuolelle. Niitä käytetään nykyisin laajasti esimerkiksi valtio-opissa, sosiologiassa ja kasvatustieteessä. Myös epidemiologiassa ja yleisemmin lääketieteessä luonnollisia koeasetelmia on hyödynnetty enenevässä määrin, mutta siellä kaikkea menetelmien potentiaalia ei ole vielä hyödynnetty (Moscoe ym. 2015; Craig ym. 2017; Hilton Boon ym. 2021). Esimerkiksi Imbensin ja Angristin kehittämälle instrumenttimuuttujamenetelmälle voisi olla paljon sovellusmahdollisuuksia erilaisissa lääketieteen satunnaistetuissa ja luonnollisissa koeasetelmissa, joissa toimenpiteen kuten lääkehoidon noudattaminen on epätäydellistä. □

⁵ Angrist ja Pischke (2010, s. 17): "Good designs have a beneficial side effect: they typically lend themselves to a simple explanation of empirical methods and a straightforward presentation of results."

⁶ Yksi osoitus kirjan vaikuttavuudesta on se, että kirjaan on jo nyt viitattu huomattavasti enemmän kuin yhteenkään Nobel-palkinnon tänä vuonna saaneiden tutkijoiden yksittäisiin tutkimuksiin. Google Scholariin mukaan kirjaan on viitattu jo lähes 19000 kertaa, mitä voidaan pitää korkeana viittaustasuna uudelle vuonna 2008 julkaistulle oppikirjalle.

Kirjallisuus

- Abadie, A. (2021), "Using synthetic controls: Feasibility, data requirements, and methodological aspects", *Journal of Economic Literature* 59: 391–425.
- Abadie, A., Diamond, A., ja Hainmueller, J. (2010), "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program", *Journal of the American Statistical Association* 105: 493–505.
- Angrist, J.D. (1990), "Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records", *American Economic Review* 80: 313–385.
- Angrist, J.D. ja Imbens, G.W. (1995): "Two-stage least squares estimation of average causal effect in models with variable treatment intensity", *Journal of the American Statistical Association* 90: 431–442.
- Angrist, J.D., Imbens, G.W. ja Rubin, D.B. (1996), "Identification of causal effects using instrumental variables", *Journal of the American Statistical Association*, 91: 444–472.
- Angrist, J.D. ja Krueger, A.B. (1991), "Does compulsory schooling attendance affect schooling and earnings?", *Quarterly Journal of Economics* 106: 976–1014.
- Angrist, J.D. ja Lavy, V. (1999), "Using Maimonides' rule to estimate the effect of class size on scholastic achievement", *Quarterly Journal of Economics*, 114: 533–575.
- Angrist, J.D. ja Pischke, J.-S. (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton, N.J.
- Angrist, J.D. ja Pischke, J.-S. (2010), "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics", *Journal of Economic Perspectives* 24: 3–30.
- Angrist, J.D. ja Pischke, J.-S. (2014), *Mastering 'metrics: The path from cause to effect*, Princeton University Press, Princeton, N.J.
- Callaway, B. ja Sant'Anna, P.H.C. (2021), "Difference-in-differences with multiple time periods", *Journal of Econometrics* 225: 200–230.
- Card, D., Lee, D., Pei, Z. ja Weber, A. (2015), "Inference of causal effects in a generalized regression kink design", *Econometrica* 83: 2453–2483.
- Card, D., Lee, D., Pei, Z. ja Weber, A. (2017), "Regression kink design: Theory and practice", *teoksessä Cattaneo, M. ja Escanciano, J.C. (toim.), Regression Discontinuity Designs, Advances in Econometrics* 38: 341–382, Emerald Publishing.
- Campbell, D.T. (1969), "Reforms as experiments", *The American Psychologist* 24: 409–429.
- Craig, P., Katikireddi, S.V., Leyland, A. ja Popham, F. (2017), "Natural experiments: An overview of methods, approaches, and contributions to public health intervention research", *Annual Review of Public Health* 38: 39–56.
- Cunningham, S. (2021), *Causal inference: The mixtape*, Yale University Press, New Haven, CT, saatavilla elektronisessa muodossa: <https://mixtape.scunning.com/>.
- de Chaisemartin, C. ja D'Haultfoeuille, X. (2020), "Two-way fixed effects estimators with heterogeneous treatment effects", *American Economic Review* 110: 2964–2996.
- Hahn, J., Todd, P. ja van der Klaauw, W. (2001), "Identification and estimation of treatment effects with a regression-discontinuity design", *Econometrica* 69: 201–209.
- Hearst, N., Newman, T.B. ja Hulley, S. (1986), "Delayed effects of the military draft on mortality. A randomized natural experiment", *New England Journal of Medicine* 314: 620–624.
- Hilton Boon, M., Craig, P., Thomson, H., Campbell, M. ja Moore, L. (2021), "Regression discontinuity designs in health: A systematic review", *Epidemiology* 32: 87–93.

- Huntington-Klein, N. (2021), *The effect: An introduction to research design and causality*, Routledge, Abingdon, saatavilla elektronisessa muodossa: <https://theeffectbook.net/>.
- Imbens, G.W. ja Angrist, J.D. (1994), "Identification and estimation of local average treatment effects", *Econometrica* 61: 467–476.
- Imbens, G.W. ja Rubin, D.B. (1997), "Estimating outcome distributions for compliers in instrumental variables models", *Review of Economic Studies* 64: 555–574.
- Moscoe, E., Bor, ja Bärnighausen, J.T. (2015), "Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: A review of current and best practice", *Journal of Clinical Epidemiology* 68: 122–133.
- Neyman, J. (1923/1990), "On the application of probability theory to agricultural experiments. Essays on principles, section 9", translated (with discussion) in *Statistical Science* 5: 465–480.
- Rubin, D.B. (1974), "Estimating causal effects of treatments in randomized and non-randomized studies", *Journal of Educational Psychology* 66: 688–701.
- Snow, J. (1855), *On the Mode of Communication of Cholera*. 2nd ed. John Churchill, London.
- Thistlewaite, D.L. ja Campbell, D.T. (1960), "Regression-discontinuity analysis: An alternative to the ex post facto experiment", *Journal of Educational Psychology* 51: 309–317.