

# Ekonometria ja tilastotiede

## Kommentti professori Heikki Kaupille

Seppo Laaksonen  
*Tilastotieteen professori*  
Helsingin yliopisto

Ensiksi haluan onnitella uutta Turun yliopiston professoria. Kiitoksen ansaitsee myös valaiseva tämän lehden numerossa 1/2009 julkaistu virkaanastujaisesityelmä, joka käsittelee ekonometriaa (Kauppi 2009). Kirjoittaja kuvasi yleisellä tasolla ja myös havainnollisilla esimerkeillä ekonometrian ominaisuuksia. Sekä makro- että mikroekonometria tuotiin kiinnostavasti esille. Toivon, että uusi professori kykenee toimimaan niin, että opiskelijat eivät koe ekonometriaa painajaisena. Näin on ilmeisesti ennen ollut asianlaita, kuten Kauppi alkujohdannaansa esittää. Hyvillä havainnollistuksilla tämänkin alan oppiminen helpottuu.

Kirjoituksen koin siis pääosin myönteisenä, mutta erityisesti yksi kohta särähti. Siteeraan: ”Vaikka ekonometria soveltaa yleisiä tilastotieteellisiä periaatteita, se eroaa muista tilastotieteen sovellusaloista merkittävästi. Tärkein ero syntyy siitä, että taloudelliset tutkimusaineistot eivät muodostu kontrolloiduissa kokeissa vaan reaali maailmassa, jossa ne syntyvät monimutkaisten taloudellisten vuorovaikutus- ja tasapainosuhteiden seurauksena. Taloudellisten aineistojen analyysi kohtaa kaikki ne ongelmat,

jotka tyypillisesti poistuisivat kontrolloiduissa kokeissa.”

Ensimmäisen lauseen tuosta voin hyväksyä, vaikka erot joihinkin sovellusaloihin voivat olla vähäisiä. Sen sijaan ihmettelen, mistä Kauppi on saanut käsityksen, että tilastotieteen aineistot syntyvät vain kontrolloiduissa kokeissa. Luulen, että sellainen suuntaus on vähemmistönä. Kontrolloituihin kokeisiin perustuva tilastotiede oli ehkä vallitsevaa viime vuosisadan alkupuolella (tunnetuin tällä alalla lienee Ronald A. Fisher). Rinnalle syntyi 1930-luvulta alkaen otanta- tai surveypohjainen lähestymistapa (ehkäpä Jerzy Neyman oli alkuaikojen keskeinen vaikuttaja). Nämä kaksi lähestymistapaa myös kohtasivat, ajoittain jopa henkilötasolla menneissä kiistoissa, joissa Fisher ja Neyman olivat eri puolilla.

Tuloksena on ollut molempien hyväksyntä. Riippuu paljon tutkimuksen tavoitteista, kummalla linjalla on paras toimia. Jos on mielessä hyvin selkeä hypoteesi, voi kontrolloitujen kokeiden eli konfirmatorinen lähestymistapa tuottaa vastauksen helpommin. Esimerkiksi maatalouden tai terveysalan tutkimuksessa täl-

laisia tilanteita on paljon, mutta harvoin taloustieteissä. Sen sijaan otosaineistoja, joihin lasken myös rekistereihin tai vastaaviin kattaviin tietoihin perustuvat, harvoin kerätään yhden tai muutaman hypoteesin tutkimiseen. Niitä voidaan käyttää hyvin monissa tutkimusasetelmissä. Tilanne ei muutu vaikka mikroaineistoista tuotetaan parametrien estimaatteja eri ajankohdille tai väleille. Näiden makroaineistojen jatkokäsittelymahdollisuuksia on yhtä lailla paljon.

Taloustieteet ovat yksi enemmän tai vähemmän kontrolloidusti kerättyjen mikro- ja makroaineistojen hyväksi käyttäjä, mutta näin ovat melkein kaikki muutkin tieteet, joissa empiirisiä aineistoja käytetään. Mainitsen tässä viimeaikaisista omista harrastuksistani monikansalliset PISA- ja European Social Survey -aineistot. Edellisiä käyttävät ehkä eniten kasvatustieteilijät, jälkimmäisiä yhteiskuntatieteilijät. Molempia ovat myös taloustieteilijät käyttäneet mutta en tiedä, ovatko suomalaiset.

Itse olen erilaisten mikroaineistojen suurluttaja ja toimin taloustieteilijöiden lisäksi useiden muiden kanssa. Kokemuksen antamalla voimalla esitän lopuksi pari näkökohtaa, joihin toivon suomalaisilta ekonometrikoilta kontribuutioita.

Viime aikoina ovat yleistyneet työnantaja-työntekijäaineistot, ei vain yhdeltä vuodelta vaan pitkältä aikaväliltä. Kun aineistot yhdistetään sekä työnantaja- että työntekijätasolla yli vuosien, on mahdollisuus tutkia monia kiinnostavia aiheita ja myös tehdä vaikkapa vertailuja

maiden välillä. Aineistot näyttävät kattavilta ja virheettömiltä, mutta sitä ne eivät ole. Ilmiselvien virheiden karsiminen on yksi tehtävä, mutta vielä haastavampaa on mittausvirheiden vaikutuksen eliminointi. Sitä ei voi tehdä havainto havainnolta, koska tutkimusyksiköitä voi olla 10 miljoonaa. Tarvitaan hyvä teoreettinen mittausvirhemalli ja sen empiirinen sovelus tietokoneohjelmalla.

Toiseksi esimerkiksi ekonometrikoiden haasteeksi tuon esille aitojen otosphojaisten mikroaineistojen käytön. Kauppi ei ymmärtääkseni puhunut näistä mitään. Niiden oikeaoppinen käyttö vaatii sekä otanta-asetelman että yhä pahemmaksi menevän vastaamattomuuden huomioon ottamista. Luulen, että näitä tekniikkoja hallitsevat taloustieteilijät ovat harvassa Suomessa. Tosin tilanne ei ole erinomainen muissakaan tieteissä.

Lopuksi mainostan itse vuonna 1996 alkuun panemaani kansainvälistä verkostoa CAED (Comparative Analysis of Enterprise Micro Data), jonka puitteista löytyy kiintoisia empiirisiä esimerkkejä haastavista mikroekonometrisista tutkimuksista. Sen seuraava konferenssi, järjestyksessään yhdeksäs, pidetään Tokiossa lokakuun 2009 alussa. Lisätietoa saa nettisivulta: <http://gcoe.ier.hit-u.ac.jp/CAED/>. □

## Kirjallisuus

Kauppi, H. (2009), "Ekonometria", *Kansantaloudellinen aikakauskirja* 105, 51–54.