

Poikkeavat havainnot epälineaarissa aikasarjaekonometriassa*

Jussi Tolvi
VTT

Turun yliopisto ja University of Leeds

Väitöskirjani otsikko koostuu kolmesta osasta: poikkeavat havainnot epälineaarissa aikasarjaekonometriassa. Käsittelen aluksi hieman näitä aiheita ja niiden välisiä suhteita, ja esittelen sitten lyhyesti väitöskirjani tärkeimmät tulokset. Lopuksi esitän muutamia työni herättämiä ajatuksia ekonometrian teorian ja käytännön välisestä suhteesta.

Ekonometria sijaitsee tieteiden maailmassa taloustieteen ja tilastotieteen välillä. Aluksi, viime vuosisadan alkupuolella, sillä tarkoitettiin jo olemassa olevien tilastollisten ja matemaattisten menetelmien soveltamista taloudellisten kysymysten tutkimukseen. Myöhemmin se on entistä selvemmin ollut aivan uusien tilastollisten menetelmien kehittäjä; menetelmien, joita on kaivattu taloudellisten havaintoaineiston erityispiirteiden vuoksi.

Epälineaariset mallit ovat myös taloustieteessä tärkeitä. Hyvin laaja yksimielisyyttä vallitsee nimittäin siitä, että taloudelliset aikasarjat selvästi ovat jollakin lailla epälineaarisia luon-

teeltaan. Näiden epälineaarisuuksien täsmällinen luonne sen sijaan ei ole ollenkaan selvää. Niinpä taloustieteessäkin on jo tutkittu suurta määrää erilaisia epälineaarisia malleja. Poikkeavat havainnot sen sijaan ovat tulleet taloustieteelliseen tutkimukseen mukaan vasta aivan viime aikoina. Poikkeavien havaintojen ja niihin liittyvien menetelmien soveltaminen on myös huomattavasti kiistellympää kuin monien muiden uusien ajatusten. Tilastotieteessä poikkeavat havainnot ovat jo pitkään olleet hyvin suosittu tutkimusaihe, ja niiden historia ulottuu itse asiassa aivan ensimmäisiin tilastotieteellisiin tutkimuksiin, muun muassa tähtitieteen alalla, jossa aineistoon liittyvät mittausvirheet saattoivat olla hyvin merkittäviä. Tilastotieteellisessä aikasarja-analyysissäkin poikkeavia havaintoja on tutkittu vakavasti jo 1970-luvulta lähtien.

Poikkeavilla havainnoilla tarkoitan työssäni, löyhästi määritellen, kulloinkin käytetyn tilastollisen mallin suhteen jotenkin erikoisia havaintoja. Yleensä, mutta ei aina, poikkeavat havainnot ovat huomattavasti suurempia tai

* *Lectio praecursoria Turun yliopistossa 8.6.2001.*

pienempiä kuin niitä lähellä olevat muut havainnot. Esimerkiksi työtaistelut, sodat ja pörssiromahdukset ovat tapahtumia jotka todennäköisesti synnyttävät poikkeavia havaintoja taloudellisiin aikasarjoihin.

Väitöskirjatyössäni tutkin siis joitakin epälineaarisen aikasarjaekonometrian menetelmiä. Erityisesti tarkastelen näiden menetelmien herkkyyttä havaintoaineistossa mahdollisesti esiintyvien poikkeavien havaintojen aiheuttamille ongelmille. Poikkeavien havaintojen olemuksesta ja jopa olemassaolosta on kuitenkin siis useampia mielipiteitä. Monissa tutkimuksissa on tosin löydetty taloudellisista aikasarjoista tietyssä mielessä poikkeavia havaintoja, mutta koska absoluuttista, kaikkialla pätevää määritelmää poikkeaville havainnoille ei ole olemassa, ei ole myöskään mitään varmuutta siitä että löydetyt poikkeavat havainnot todella ovat poikkeavia. Erityisesti epälineaaristen mallien maailmassa tämä voi olla epäselvää, sillä nämä mallit voivat tuottaa hyvin monenlaisia aikasarjoja, jotka intuitiivisesti näyttäisivät selvästi sisältävän poikkeavia havaintoja, vaikka näin ei todellisuudessa olisikaan. Käytännössä voitaneen kuitenkin olla melko yksimielisiä siitä, että tietyt havainnot ovat selvästi erilaisia kuin muut, ja nykyisin käytössä olevissa malleissa siis voidaan luokitella poikkeaviksi havainnoiksi.

Väitöskirjani sisältö on lyhyesti seuraavanlainen. Kirja koostuu johdannosta ja neljästä itsenäisestä artikkelista. Ensimmäinen artikkeli käsittelee erityyppisten poikkeavien havaintojen vaikutuksia kahteen epälineaarisuustestiin. Testit ovat Lagrangen kerrointestit autoregressiivistä ehdollista heteroskedastisuutta ja bilineaarisuutta vastaan. Tulosten mukaan yksittäiset poikkeavat havainnot voivat olla ratkaisevan tärkeitä. Jopa yksi havainto 500 havain-

non aineistossa voi muuttaa testien perusteella tehtävät johtopäätökset päinvastaisiksi kuin olisi havaittu ilman kyseisen poikkeavan havainnon vaikutusta. On siis mahdollista, että sarjassa oleva epälineaarisuus jää havaitsematta, tai vastaavasti tulkitaan todellisuudessa lineaarisen sarjan olevan epälineaarinen. Nämä tulokset perustuvat teoreettisiin simulointiko-keisiin.

Toisessa artikkelissa tutkitaan yksityiskohtaisesti useita pitkiä suomalaisia taloudellisia aikasarjoja. Sarjoista etsitään ensin tietyssä mielessä poikkeavia havaintoja, joita löytyykin jokaisesta sarjasta. Usein nämä löydetyt poikkeavat havainnot esiintyvät ryppäissä, ja useammissa sarjoissa samoihin aikoihin. Löydetuille poikkeaville havainnoille etsitään seuraavaksi mahdollisia selityksiä historiallisista taloudellisista ja poliittisista tapahtumista. Lisäksi tarkastellaan löydettyjen poikkeavien havaintojen vaikutuksia tutkittujen aikasarjojen tilastollisiin ominaisuuksiin. Käy ilmi, että poikkeavat havainnot ovat ratkaisevan tärkeitä myös näissä todellisissa aikasarjoissa, sillä esimerkiksi kolmen tässä artikkelissa käytetyn epälineaarisuustestin perusteella tehtävät johtopäätökset voivat olla riippuvaisia vain muutaman, ääritapauksessa jopa yhden ainoan, havainnon mukanaolosta aineistossa.

Kahdessa viimeisessä artikkelissa keskitytään niin sanottuihin pitkän muistin malleihin, joilla voidaan mallittaa toisistaan ajallisesti kaukana olevien havaintojen välisiä yhteyksiä. Teoreettisesti selvitetään ensin miten poikkeavat havainnot voivat vaikuttaa näiden mallien parametriestimaatteihin. Simulointikokeiden tulosten mukaan poikkeavat havainnot voivat joko kokonaan estää tutkituissa aikasarjoissa olevan pitkän muistin havaitsemisen, tai merkittävästi muuttaa tilastollista arviota sarjan pit-

kän muistin ominaisuuksista. Lisäksi näissä kahdessa artikkelissa tarkastellaan käytännön mahdollisuuksia suojautua poikkeavien havaintojen vahingollisilta vaikutuksilta. Näitäkin tuloksia havainnollistetaan useiden taloudellisten esimerkkien avulla, muun muassa pörssikursien, inflaation ja valuuttakurssien aikasarjoilla. Käytännössä on melko helppoa suojautua poikkeavien havaintojen aiheuttamilta ongelmilta, kunhan ongelmallisimmat havainnot jotenkin otetaan huomioon mallin valinnassa ja estimoinnissa.

Entä sitten teorian ja käytännön välinen suhde? Samassa hengessä kuin kansantaloustiedettä joskus kritisoidaan liiasta keskittymisestä pelkkään teoretisointiin, voisi pohtia myös ekonometrisen tutkimuksen nykytilaa. Siinä missä taloustiede kriitikoiden mukaan kehittää jatkuvasti uusia teorioita vertaamalla niitä mitenkään todellisuuteen, myös hyvin suuri osa julkaistusta ekonometrisesta tutkimustyöstä kehittää yhä uusia menetelmiä. Valitettavan vähän sen sijaan on olemassa eri menetelmiä soveltavaa, vertailevaa ja arvioivaa tutkimusta. En toki olisi kokonaan kieltämässä uusien menetelmien kehittelyä, mutta toivoisin että enemmän panostettaisiin myös näiden menetelmien tarkempaan tutkimiseen, sekä teoriassa että käytännössä.

Metodologian suhteen olen taipuvainen pragmaattisuuteen. Ekonometrian menetelmien mediaanikäyttäjät eivät pysty soveltamaan kaikkia uusia menetelmiä. Ja toisaalta eniten vaikutusta käytännön tutkimukseen on sitten-

kin melko yksinkertaisilla, intuitiivisesti ymmärrettävillä menetelmillä. Sama ilmiö näkyy ehkä siinäkin, että usein yleisimmin käytetyt menetelmät tiedetään jo tavalla tai toisella puutteellisiksi. Yksi esimerkki tästä on väitöskirjassanikin mainitun GPH-estimaattorin käyttö pitkän muistin testauksessa, vaikka se tiedetäänkin harhaiseksi mikäli aineistossa on myös lyhyttä muistia. Käytännön työhön tarvitaan siis helposti sovellettavia menetelmiä, joiden olisi kuitenkin syytä olla mahdollisimman hyvin tarkoitukseensa sopivia. Lisäksi näiden menetelmien mahdolliset puutteet eri tilanteissa tulisi perusteellisesti selvittää. Jonkinlainen valinta on kuitenkin tehtävä menetelmien käytännöllisyyden ja hyvyyden välillä. Yleisesti optimaalisia menetelmiä ei ehkä ole olemassa, mutta monissa tilanteissa kohtuullisia, tai ainakin riittävän hyviä menetelmiä voitaneen kehittää. Tämänäsuuntaisia ajatuksia esiintyy myös joidenkin aivan uudentyypisten menetelmien perusteluissa. Näistä esimerkiksi geneettiset algoritmit ovat varmasti tulossa yhä laajemmin myös ekonometrian piiriin.

Taloudellisista havaintoaineistoista joudutaan aina tekemään suuri määrä oletuksia. Yksi tällainen oletus on, että tutkitussa aineistossa ei ole poikkeavia havaintoja. Työni perusteella voidaan siis kuitenkin ajatella, että poikkeavia havaintoja itse asiassa esiintyy hyvin yleisesti, ja että ne voivat myös olla hyvin haitallisia. Toivoa sopii, että taloustieteilijät tulevaisuudessa ottavat nämä ongelmat entistä vakavammin. □