

Ylioppilastutkintojen pitkä varjo

Artturi Björk

Mitä varten abit pänttävät ylioppilaskokeisiin? Tietyksi suurimmaksi osaksi siitä syystä, että he uskovat, että kokeen arvosanoilla on väliä: että ne, elleivät täysin määrää, niin ainakin merkittävästi vaikuttavat siihen, millaiseksi loppuelämä muodostuu. Vastaava sitkeästi elävä uskomus on, että yläasteen arvosanoilla on iso merkitys, koska niiden perusteella voi päästää eliittilukioon ja eliittilukiostahan tulee ulos eliittiä! Taloustieteilijät ovat kuitenkin tutkineet eliittilukioiden vaikutusta ja lopputulos on, että eliittilukiot eivät ole muita lukioita parempia kouluttamaan oppilaitaan, vaan niihin otetaan oppilaita ennestään taitavimmat oppilaat. Uskomus ylioppilasarvosanojen merkityksestä voi perustua samanlaiseen virhepäätelmään: laudaturin saaneet ylioppilaat menestyvät elämässään upeasti, joten minäkin varmasti menestyisin, jos vain saisin laudaturin ylioppilaskokeesta. Ehkä itsekkin ajattelin jotain tällaista uuden vuosituhannen ensimmäisenä syksynä, kun menin tekemään matematiikan yo-kokeen kolmatta kertaa. Mutta eihän paremman

arvosanan saaminen muuta ihmistä muuten kuin että ”nyt sillä on se parempi arvosana”. Mutta tuon verran se muuttaa, ja tutkimme Työn ja talouden tutkimus LABOREn tutkimusohjaajan Hannu Karhusen kanssa, mikä on paremman yo-arvosanan vaikutus.

Ennen tätä tutkimusta ylioppilaskokeen arvosanojen vaikutuksista Suomessa ei ollut tutkittua tietoa. Vastaavan lukion päättävän kokeen vaikutuksia oli tutkittu Israelissa, ja tutkimustyömme aikana julkaistiin myös tuloksia Ranskasta. Näissä tutkimuksissa pystyttiin havaitsemaan vain lyhyen aikavälin vaikutukset. Pitkän aikavälin vaikutuksista ei siis ollut ollenkaan tutkittua tietoa. Ylioppilaskokeita edeltävistä ja niitä seuraavista vaiheista sen sijaan oli. Kuten jo mainitsinkin, niin eri lukioiden vaikutuksia oli tutkittu Suomessa (Tervonen ym. 2017). Korkeakoulujen vaikutuksista sen sijaan on runsaasti tutkimuksia, ja viime vuonna julkaistiin tutkimus erityisesti juuri korkeakoulujen kurssiarvosanojen vaikutuksista Singa-

Arkkitehti Artturi Björkin (artturi.bjork@gmail.com) taloustieteen aineen väitöskirja *Essays in Economics of Education, Aging and Trade Unions* tarkastettiin Aalto-yliopiston kauppa korkeakoulussa 10.2.2023. Vastaväittäjänä toimii professori Janne Tukiainen Turun yliopistosta ja kustoksena professori Marko Terviö Aalto-yliopiston kauppa korkeakoulusta. Kirjoitus perustuu Björkin väitöstilaisuudessa pitämään *Lectio praecursoriaan*. Se keskittyy väitöskirjan ensimmäiseen esseeseen. Muut kaksi esseettä käsittelevät väestön ikääntymisen vaikutusta varallisuus- työllisyys- ja tuloeroihin, sekä palkanmuodostusta ammattiliittojen ja yritysten dynaamisissa vuorovaikutuksissa, ks. <https://aalto.doc.aalto.fi/bitstream/handle/123456789/119184/isbn9789526411231.pdf>

poessa.¹ Tiedoissamme oli siis Suomen ylioppilaskokeen arvosanojen pitkän aikavälin vaikutusten mentävä aukko.

Tutkimme ruotsin ja englannin kielen ylioppilaskokeita vuodesta 1982 vuoteen 2013. Käytössämme oli Ylioppilastutkintolautakunnan tiedot yo-kokeiden pisteistä ja arvosanarajoista, joiden perusteella koepisteet muutetaan arvosanoiksi. Tutkimuksen kohteena oli noin 2 miljoonaa koetta. Ruotsin ja englannin kielen kokeista voi saada pisteitä nolasta kolmeensataan. Lukion omat opettajat pisteyttävät kokeet alustavasti ja sen jälkeen ne lähetetään Ylioppilastutkintolautakunnan sensoreiden pisteytettäväksi. Kun kokeet on pisteytetty, Ylioppilastutkintolautakunta asettaa arvosanarajat niin, että eri arvosanojen osuudet ovat ennalta päätetyn mukaiset. Arvosanojen jakauma on karkeasti ottaen sellainen, että keskimmäisiä arvosanoja B, C ja M saa kutakin noin viidesosa kokeen palauttaneista, kahta matalinta ja kahta korkeinta arvosanaa tätä pienemmät osuudet. Vaikka koe itsessään tuottaa tarkkaa tietoa siitä, miten hyvin opiskelija pärjasi suhteessa muihin kokeeseen osallistuneisiin, niin todistukseen tuleva arvosana M kertoo vain, että opiskelija oli jossain 60 ja 80 % välillä, mutta ei tarkemmin missä.

Tähän mennessä olen kirjoittanut epämääräisesti vain arvosanojen vaikutuksesta, mutta vaikutuksesta mihin: lukion jälkeiseen koulutukseen, tuloihin ja työllisyyteen, jotka ovat tyypillisiä kiinnostuksen kohteita taloustieteessä. Arvosanoilla voisi toki olla muitakin mielenkiintoisia vaikutuksia, esimerkiksi mielenterveyteen tai tuleviin perhesuhteisiin, mutta näiden tutkimisen jätimme myöhemmäksi, osataan siksi, että tämä tutkimus toteutettiin

Strategisen tutkimuksen neuvoston Taidot työhön-projektin sateenvarjon alla.

Ihanteellinen keino arvosanan vaikutuksen tutkimiselle olisi ottaa ryhmä opiskelijoita ja vetää heille arvosanat hatusta sen sijaan, että ne määräytyvät koepisteiden perusteella. Tämän ryhmän kesken arvosanan vaikutus pystyttäisiin tunnistamaan vertailemalla suoraan esimerkiksi C:n ja M:n saaneiden opiskelijoiden työuria. Tällainen koe ei tietenkään tule kysymykseen, mutta käyttämämme tutkimusmenetelmä perustuu samaan ideaan: tutkimme opiskelijoita, joiden yo-kokeen pisteet olivat lähellä arvosanarajaa. Opiskelijat jäivät joko juuri ja juuri arvosanarajan alle tai ylittivät sen täpärästi. Nämä opiskelijat pärjäsivät koepisteillä mitattuna yo-kokeessa käytännössä yhtä hyvin, mutta saivat siitä huolimatta eri arvosanan.

Tätä menetelmää kutsutaan regressioepäjatkuvuusmenetelmäksi (*regression discontinuity method*, ks. esim. Cattaneo ym. 2020). Menetelmä perustuu siihen, että tarkastellaan koepisteiden ja tulevien ansioiden yhteyttä ja havaitaan, että opiskelijoilla, jotka ovat saaneet yhden koepisteen enemmän on vähän korkeammat ansiot, kunnes tullaan arvosanarajalle, missä yhden koepisteen ero tarkoittaa eroa myös arvosanassa. Tällä rajalla havaitaan epäjatkuvuus ansioiden ja koepisteiden yhteydessä: ansiot kasvavatkin huomattavasti enemmän tässä kohtaa. Koska ainoa asia, joka rajalla muuttuu epäjatkuvasti, on arvosana, tiedämme, että rajan yli havaittu hyppäys ansioissa johtuu siitä, että arvosana muuttuu.

Tutkimuksen ansiosta meillä on nyt tietoa siitä, että arvosanat vaikuttavat merkittävästi opiskelijoihin. Nopeimmin vaikutus näkyy kokeen uusimisessa. Kokeen uusinnan todennäköisyys laskee 16 prosenttiyksikköä, jos opiskelija saa korkeamman arvosanan. Opiskelijan, joka on saanut hädin tuskin korkeamman arvosana-

¹ *Israelia koskevista tuloksista ks. Ebenstein ym. (2016), Ranskasta Canaan ym. (2018) ja Singaporesta Tan (2022).*

nan, pitäisi pärjätä uusinnassa huomattavasti paremmin, jotta hän voisi korottaa arvosanaan-
sa, kun taas juuri ja juuri arvosanarajan alapuo-
lelle jääneen tarvitsee menestyä vain hieman
edellistä kertaa paremmin. Kokeiden vilkas uu-
siminen ei kuitenkaan ole kokonaisuuden kan-
nalta hyvä asia. Koetulokset pistävät opiskelijat
paremmuusjärjestykseen, jolloin kokeen uusijan
saama parempi arvosana tiputtaa muut opiske-
lijat järjestyksessä alemmaksi.

Arvosanalla on pieni vaikutus lukion jälkei-
siin opintoihin. Korkeamman arvosanan saa-
neet kouluttautuvat 11 päivää kauemmin ja val-
mistuvat yliopistosta 4 prosenttia todennäköi-
semmin. Vastaavalla tavalla kuin arvosana vai-
kuttaa siihen, valmistuuko opiskelija ammatti-
koulusta, ammattikorkeakoulusta, yliopistosta
vai jääkö vaille lukion jälkeistä tutkintoa, arvo-
sana voi vaikuttaa siihen mihin koulutusohjel-
maan opiskelija sijoittuu koulutusasteen sisällä.
Esimerkiksi jatkaako hän lukion jälkeen yliopis-
to-opintoihin Helsingin yliopiston oikeustie-
teelliseen tiedekuntaan vai jatkuvatko opinnot
Vaasan yliopiston hallintotieteiden tiedekun-
nassa. Näitä koulutusasteen sisäisiä vaikutuksia
on hankalampi tunnistaa. Tutkimuksessa huo-
mioimme kaikki erot koulutuksessa vertaamal-
la koulutusohjelmien mediaanituloja, joista ar-
vosanarajan eri puolille jääneet opiskelijat val-
mistuivat ja havaitsimme, että arvosanarajan ylä-
puolella olevat oppilaat valmistuivat koulutus-
ohjelmista, joissa keskiansiot ovat noin prosen-
tin rajan alle jääneiden ansioita korkeammat.

Ylioppilastutkinnon arvosanojen vaikutuk-
set jatkokoulutukseen muodostuvat pian yliop-
pilaskirjoitusten jälkeen, mutta pitkän aikavälin
vaikutukset havaitaan työurista. Toisin kuin
Israelissa ja Ranskassa on havaittu, Suomessa
korkeampi ylioppilastutkinnon arvosana ei kas-
vata tuloja ensimmäisen 10 vuoden aikana ko-

keen suorittamisen jälkeen. Tämä voi selittyä
sillä, että Suomessa opiskelut kestävät pidem-
pään ja täysi siirtyminen töihin tapahtuu meillä
myöhemmin. Joka tapauksessa erot arvosanara-
jan alle jääneiden ja sen yli päässeiden ansioiden
välillä alkavat näkyä 30-vuotiailla ja kasvavat
sen jälkeen tasaisesti niin pitkälle kuin käytet-
ävissä olevat tiedot ulottuvat: 50-vuotiaana ero
vuosittaisessa ansiossa on noin 3 % vastaten
noin puolentoista viikon palkkaa. Työllisyydessä
emme havaitse eroja, eli erot ansiossa selittyvät
sillä, että arvosanarajan yläpuolelle päässeet saa-
vat korkeampia palkkoja. Koska käyttämämme
tutkimusmenetelmä mahdollistaa sellaisten
opiskelijoiden vertaamisen keskenään, jotka
poikkeavat toisistaan ainoastaan kokeen arvo-
sanauksen suhteen, tutkimuksessa havaitut erot vas-
taavat kysymykseen, miten 19-vuotiaana sattua-
man aiheuttama ero yo-kokeen arvosanassa vai-
kuttaa pitkälle tulevaisuuteen. Tutkimuksen tu-
loksot osoittavat, että tällaiset erot eivät tasoi-
tuisi myötä, vaan sen sijaan nuorena syntyneet
erot voimistuvat. Tämä puhuu sen puolesta, et-
tä tällaisten erojen synnyttämistä tulisi välttää.

Miten erojen syntymistä sitten voitaisiin
välttää? Arvosanojen aiheuttamissa eroissa on
kyse informaatiosta: mitä kokeen arvosana ker-
too opiskelijan osaamisesta. Mitä isommat hyp-
päykset tässä informaatiosta on arvosanasta toi-
seen, sitä suurempi vaikutus arvosanalla on.
Kun ylioppilaskirjoitukset aloitettiin 1800-lu-
vun puolivälissä, arvosanoja oli vain neljä.
Tällöin hyppy yhdestä arvosanasta seuraavaan
olivat huomattavan suuria: A:n saanut ylioppilas
oli keskimäärin pärjännyt paremmin kuin 20 %
kokeen suorittaneista, ja M:n saanut ylioppilas
taas oli pärjännyt paremmin kuin 60 %.
Hyppäys vastasi siis 40 % eroa keskimääräises-
sä sijoittumisessa. Toisaalta arvosanoilla itses-
sään oli vähemmän merkitystä, koska kaikilla

ylioppilailla oli pääsy yliopistoon. Arvosteluasteikkoa laajennettiin kahdella arvosanalla 70-luvulla ja vielä yhdellä lisää 90-luvulla. Nyt arvosanoja on siis 7, ja hyppäysten koko arvostanasta toiseen on puolittunut korkeintaan 20 % eroiksi sijoittumisessa ylioppilaskokelaiden suoriutumisjärjestyksessä.

Arvosanojen määrää voitaisiin kuitenkin kasvattaa edelleen. Tällöin eri arvosanojen väliset keskimääräiset osaamiserot olisivat pienempiä, ja sillä, kummalle puolelle arvosanarajaa kokelas sijoittuu, olisi vähemmän merkitystä. Tiheämpi arvostusasteikko myös todennäköisesti vähentäisi kokeiden uusimista, kun kokeen uusimisen vaiva ja kustannus pysyisi entisellään, mutta hyöty korkeamman arvosanan saamisesta olisi pienempi. Vaihtoehtoisesti nykyisten arvosanojen rinnalla ylioppilastodistuksessa voitaisiin raportoida, mihin kokelaiden persentiiliin eli sadasosaan opiskelijan koetulos sijoittuu. Esimerkiksi Yhdysvalloissa yliopistojen sisänotossa käytetyissä testeissä tulokset raportoidaan sekä koepisteinä että persentiileinä.

Pari viikkoa sitten tuloksensa saaneita abiturienteja muutokset kokeiden tulosten raportointiin eivät kuitenkaan ehdi auttaa, vaan heidän ei auta muu kuin toivoa hyvää onnea niin, että koepisteet sattuisivat osumaan arvosanarajan yläpuolelle. Silläkin uhalla, että tieto lisää tuskaa: pelkästään yhtä pykälää paremman arvosanan kumulatiivinen vaikutus ansiotuloihin on 50-vuotiaaksi mennessä 12 000 euroa tai diskontattuna kahden prosentin vuosikorolla 7500 euroa. Nämä vaikutukset ovat sitä paitsi ajalta ennen korkeakoulujen sisäänpääsyuudistusta, jolloin ylioppilastutkinnon arvosanojen merkitys korkeakoulujen sisänotossa oli huomattavasti nykyistä pienempi. Kokeeseen valmistautumiseen käytetystä ajasta saa siis todennäköisesti ruhtinaallisen korvauksen! □

Kirjallisuus

- Canaan, S., ja Mouganie, P. (2018), "Returns to education quality for low-skilled students: Evidence from a discontinuity", *Journal of Labor Economics* 36: 395–436.
- Cattaneo, M., Idrobo, N. ja Titiunik, R. (2020), *A Practical Introduction to Regression Discontinuity Designs*. Cambridge University Press, Cambridge.
- Ebenstein, A., Lavy, V., ja Roth, S. (2016), "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution", *American Economic Journal: Applied Economics* 8: 36–65.
- Tan, B. J. (2022), "The consequences of letter grades on labor market outcomes and student behavior", *Journal of Labor Economics* (tulossa).
- Tervonen, L., Kortelainen, M. ja Kanninen, O. (2017), *Eliittilukioiden vaikutukset ylioppilaskirjoitusten tuloksiin*, VATT tutkimuksia 186. Valtion taloudellinen tutkimuskeskus, Helsinki.