



Eliittilukioiden vaikutukset yliopisto-opintojen aloittamiseen ja koulutusalan valintaan

LASSI TERVONEN & MIKA KORTTELAINEN & OHTO KANNINEN

Lukioiden väliset erot oppimistuloksissa ja oppilasohjassa ovat Suomessa merkittäviä ja erot korostuvat isoimmassa kaupungeissa, joissa kilpailtujen ja suosittujen lukioiden sisäänpääsyrajat ovat hyvin korkeita. Tällaisia hyvin korkean sisäänpääsyrajan lukiota kutsutaan eliittilukioiksi. Artikkelissa tutkitaan eliittilukioiden pidemmän aikavälin vaikutuksia yliopisto-opintojen aloittamiseen ja koulutusalan valintaan vertaamalla eliittilukioiden sisäänpääsyrajan juuri ylittäneitä niukasti rajan alle jääneisiin.

.....
English summary at the end of the article

Johdanto

Kansainvälisten vertailujen mukaan peruskoulujen väliset erot oppimistuloksissa ovat olleet Suomessa vähäisiä, vaikkakin viime vuosina ne ovat kasvaneet. Sen sijaan toisen asteen koulutuksessa ja varsinkin lukioissa koulujen välillä on ollut jo pitkään isoja oppilasohjaa ja oppimistuloksia koskevia eroja. Nämä erot korostuvat varsinkin isoimmassa kaupungeissa, joissa kilpailtujen ja suosittujen lukioiden sisäänpääsyrajat muodostuvat hyvin korkeiksi. Kutsumme tällaisia hyvin korkean sisäänpääsyrajan lukiota eliittilukioiksi.

Tämän tutkimuksen tarkoituksena on tutkia eliittilukioiden vaikutusta yliopisto-opintojen aloittamiseen ja opintoalan valintaan. Tarkasteltavaksi on valittu viisi lukiota, joiden yleislinjojen keskiarvorajat ovat olleet pääkaupunkiseu-

dun korkeimpia vuosina 2000–2008. Lukiot ovat Etelä-Tapiolan lukio, Helsingin normaalilyseon lukio (Norssi), Helsingin yliopiston Viikin normaalikoulun lukio (Viikki), Mäkelänrinteen lukio (Märsky) ja Ressun lukio.¹

Aikaisemman tutkimuksen perusteella tarkastelemillamme eliittilukioilla ei ole vaikutusta ylioppilaskirjoitusten arvosanoihin (Tervonen & al. 2017). Kyseiset lukiot voivat kuitenkin tuottaa opiskelijoilleen hyötyjä, jotka eivät näy ylioppilaskirjoitusten arvosanoissa. Yksi tällainen mahdollinen hyöty on se, että eliittilukiossa opiskelut saattaa päästä todennäköisemmin yliopistoon tai muuhun korkeakouluun kuin eliittilukion ulkopuolelle jäänyt. Tässä tapauksessa kyse voisi olla siitä, että eliittilukioissa opitaan esimerkiksi yliopiston pääsykokeissa hyödyttäviä taitoja. Lisäksi eliittilukiossa opiskelu saattaa esimerkiksi johtaa muissa lukioissa opiskelevien valintoja kunnianhimoisempiin alavalintoihin jatko-opinnoissa.

Artikkeli on osa Suomen Akatemian Strategisen tutkimuksen neuvoston rahoittamaa "Osaaminen, koulutus ja tulevaisuuden työelämä" -hanketta (rahoituspäätös 303687). Haluamme myös kiittää Etelä-Tapiolan lukiota, Helsingin normaalilyseon lukiota, Helsingin yliopiston Viikin normaalikoulun lukiota, Mäkelänrinteen lukiota sekä Ressun lukiota luvasta julkaista kyseisiä lukioita koskevia tutkimustuloksia hankkeessamme.

1 Usein eliittilukioista puhuttaessa nostetaan esiin Helsingin Suomalainen Yhteiskoulu (SYK). Emme kuitenkaan tarkastele sitä tässä, koska SYK:ssa hakijoille on erikseen sisäänpääsyrajat eri pitkiä kieliä opiskelleille, eikä näiden linjojen erotteleminen ole aineistomme perusteella mahdollista.

Taulukko 1. Yliopistoissa ja korkeakouluissa aloittaneiden osuudet

	Eliitti	Lukio (PKS)	Lukio	Kaikki
Yliopisto tai AMK	0,878	0,635	0,733	0,442
Yliopisto	0,733	0,352	0,342	0,192
N	4 708	57 165	238 691	651 569
Kauppatieteellinen	0,223	0,219	0,123	0,142
Lääketieteellinen	0,095	0,045	0,051	0,048
Oikeustieteellinen	0,091	0,048	0,03	0,033
Teknillistieteellinen	0,238	0,226	0,223	0,215
N	3 451	20 088	81 531	114 477

Käsitlemiemme eliittilukioiden opiskelijoista suuri osa päätyy yliopistoon tai muuhun korkeakouluun. Taulukko 1 kuvaa yliopistoissa aloittaneiden osuutta eliittilukioiden (Eliitti), pääkaupunkiseudun lukioiden (Lukio [PKS]), kaikkien lukioiden (Lukio) ja kaikkien toisen asteen yhteisvalintaan osallistuneiden (Kaikki) joukossa. Taulukosta nähdään, että yliopistoissa aloittaneiden osuus eliittilukioiden opiskelijoista² on lähes neljä kertaa suurempi kuin kaikkien yhteisvalintaan osallistuneiden joukossa ja yli kaksi kertaa suurempi kuin koko maan tai pääkaupunkiseudun lukioissa keskimäärin. Erot pienenevät, jos myös ammattikorkeakoulut huomioidaan, mutta ne pysyvät silti selvinä. Tässä tutkimuksessa keskitymme kuitenkin yliopistoihin ja yliopistotasoisiin korkeakouluihin (joita myös kutsutaan tässä artikkelissa yliopistoiksi). Lääke-, oikeus-, kauppa- ja teknillistieteellisiä aloja pidetään joskus ”eliittialoina” (Nori 2011), ja teknillistieteellistä alaa lukuun ottamatta eliittilukioiden opiskelijat päätyvätkin opiskelemaan niitä muita useammin: kun tutkitaan yliopistossa aloittaneiden joukkoa, eliittilukioiden opiskelijat päätyvät lääketieteelliseen noin kaksi kertaa ja oikeustieteelliseen noin kolme kertaa todennäköisemmin kuin kaikki toisen asteen yhteisvalintaan osallistuneet. Samassa joukossa he myös päätyvät kauppatieteelliseen lähes kaksi kertaa todennäköisemmin kuin lukiolaiset keskimäärin, mutta eivät juuri todennäköisemmin kuin pääkaupunkiseudun lukiolaiset.

Taulukon 1 perusteella ei kuitenkaan voi pää-

tellä, että eliittilukiossa opiskelu kannattaa, koska se on tie todennäköisempään yliopistopaikkaan ja todennäköisempään paikkaan eliittialalla. Kyse on valikoitumisharhasta: eliittilukioihin todennäköisesti valikoituu opiskelijoita, jotka päätyvät opiskelemaan yliopistoon ja tietyille yliopistoaloille muita todennäköisemmin lukiosta riippumatta. Siten pelkkä kuvailu eliittilukioissa opiskelien opiskelupaikoista ei tarkoita, että menestys johtuisi kyseisissä lukioissa opiskelusta. Arviointia hankaloittaa se, että emme tiedä ns. kontrafaktuaalia eli sitä, mihin kyseiset henkilöt olisivat päätyneet, jos he olisivat opiskelleet ”tavallisessa” lukiossa.

Tässä tutkimuksessa valikoitumisharha pyritään poistamaan hyödyntämällä kvasikokeellista regressioepäjatkuvuusasetelmaa (*regression discontinuity design, RDD*). RDD:tä hyödynnettiin alun perin kasvatuspsykologian alalla (Thistlewaite & Campbell 1960), mutta viimeisen 20 vuoden aikana se on noussut hyvin suosituksi menetelmäksi yhteiskuntatieteissä, erityisesti koulutuksen ja politiikan tutkimuksessa (Cattaneo & al. 2018). RDD:n ideana on tässä tapauksessa verrata niukasti eliittilukion sisäänpääsyrajan ylittäneitä juuri sen alapuolelle jääneisiin jonkin tuleman – kuten yliopiston aloittamistodennäköisyyden – suhteen. Rajan läheisyydessä, sekä sen ala- että yläpuolella, olevat opiskelijat ovat keskimäärin samanlaisia, jolloin ainoa ero opiskelijoiden välillä on se, ovatko he tulleet valituksi eliittilukioon vai eivät. Siten rajan alapuolelle niukasti jääneitä voidaan käyttää kontrolliryhmänä rajan ylittäneille, mikä mahdollistaa eliittilukiotarjouksen kausaali vaikutuksen estimoinnin. RDD:n avulla vaikutusta voidaan kuitenkin tutkia vain rajan läheisyydessä, eikä tuloksia voi täten yleistää koskemaan kaikkia eliittilukioihin hyväksytyttä.

² Eliittilukioiden opiskelijoilla tarkoitetaan tässä eliittilukion aloittaneita opiskelijoita. Näistä kuitenkin vain muutama prosentti aloittaa yo-kirjoituksensa jossain muussa lukiossa kuin siinä, johon on päässyt.

On syytä uskoa, että eliittilukiot parantaisivat opiskelijoidensa oppimistuloksia ja täten myös pidemmän aikavälin tulemia. Voidaan esimerkiksi ajatella, että hyvät opiskelijat hyötyvät tilanteesta, jossa he ovat muiden samalla tasolla olevien joukossa (Lazear 2001). Lisäksi eliittilukion maineella voi olla positiivinen vaikutus oppilaaseen. On myös mahdollista, että eliittilukioihin valikoituu keskimäärin parempia opettajia, jolloin mahdolliset vaikutukset oppimistuloksiin selittyisivät ennemminkin paremmilla opettajilla eliittikouluissa kuin vertaisryhmän vaikutuksilla (Hoekstra & al. 2016). Useimmissa aikaisemmissa tutkimuksissa eliittilukioiden vaikutukset oppimistuloksiin ovat kuitenkin olleet vähäisiä ja toisinaan jopa negatiivisia (Abdulkadiroğlu & al. 2014; Dobbie & Fryer 2014; Hoekstra & al. 2016). Toisaalta pidemmän aikavälin kausaali vaikutuksia eliittilukiossa opiskelulle on tarkasteltu vain vähän. Kuitenkaan esimerkiksi New Yorkin eliittikouluissa opiskelulle ei löydetty positiivista vaikutusta yliopiston aloittamiseen tai sieltä valmistumiseen (Dobbie & Fryer 2014).

Eliittikouluja koskevan kirjallisuuden ohella tutkimuksen aihe liittyy koulujen eriytmistä koskevaan yhteiskuntatieteelliseen ja koulutuspoliittiseen kirjallisuuteen. Koulujen eriytmistä koskevan koulutussosiologisen kirjallisuuden keskeisiä tuloksia on se, että perhetausta ja sosiaalinen luokka selittävät koulu- ja opiskeluvaihtoja ja ovat yhteydessä myös koulujen eriytmiseen (Kuusela 2003; Nori 2011; Seppänen & al. 2015). Käytännössä tämä tarkoittaa sitä, että korkeasti koulutettujen ja hyvätuloisten vanhempien lapset päätyvät eliittikouluihin todennäköisemmin kuin heikommasta sosioekonomisesta taustasta tulevat. Toisaalta jos eliittikoulu parantavat oppimistuloksia tai vaikuttavat menestykseen myöhemmissä opinnoissa tai työelämässä, eliittikoulu voivat entisestään vahvistaa perhetaustan merkitystä ja sitä kautta eriarvoisuutta. Tämän vuoksi tutkimus eliittilukioiden pidemmän aikavälin vaikutuksista on tärkeää myös koulujen eriytmistä sekä eriarvoisuutta koskevan koulutuspoliittisen keskustelun kannalta.

Artikkeli etenee seuraavalla tavalla: Ensin esittelemme tutkimusasetelman. Tämän jälkeen ku-

vaimme tutkimuksessa käytettyä aineistoa ja käsittelemme tutkimuksen tuloksia. Tulosten jälkeinen osio käsittelee tutkimusasetelman validiteettia ja tulosten robustisuutta. Lopuksi esitämme tutkimuksen yhteenvetdon.

Tutkimusasetelma

Lukioiden yleislinjojen yhteishaun valintaperusteet ovat pysyneet pääasiallisesti samana tutkimiemme vuosien aikana. Niiden mukaan opiskelijat valitaan lukioon peruskoulun päättötodistuksen (tai lisäopetuksen) lukuaineiden keskiarvon mukaan, mikäli lukioon on enemmän hakijoita kuin sillä on vapaita opiskelupaikkoja. Keskiarvo lasketaan kahden desimaalin tarkkuudella. Lukuaineita ovat äidinkieli ja kirjallisuus, toinen kotimainen kieli, vieraat kielet, uskonto tai elämäntutkimustieto, historia, yhteiskuntaoppi, matematiikka, fysiikka, kemia, biologia, terveystieto ja maantieto (Lukiolaki 1998).³

Toisen asteen yhteishaussa opiskelijat voivat hakea viiteen eri koulutukseen (oppilaitoksen ja koulutuslinjan yhdistelmä), ja oppilaitoksilla on kullekin tarjoamalleen koulutuslinjalle valittavan opiskelijamäärän kiintiö. Nämä kiintiöt täytetään viivästetyn hyväksynnän algoritmin mukaan (ks. myös Virtanen 2016). Algoritmin ensimmäisellä kierroksella opiskelijat ovat ehdolla ensimmäisiin vaihtoehtoihinsa. Jokainen oppilaitos asettaa nämä hakijat järjestykseen ja hyväksyy alustavasti kiintiönsä mahtuvat opiskelijat. Muut hakijat hylätään. Myös alustavasti hyväksytyt voidaan hylätä myöhemmillä kierroksilla. Tämän jälkeen algoritmi jatkuu seuraavalla tavalla: olkoon $x > 1$. Kierroksella x seuraavaan vaihtoehtoonsa ovat ehdolla kierroksella $x - 1$ hylätyt hakijat. Jokainen koulu asettaa tämän kierroksen hakijat ja aikaisemmin alustavasti hyväksytyt järjestykseen, ja hyväksyy alustavasti kiintiönsä mahtuvat. Alustavasti hyväksytyt voivat tulla hylätyksi myöhemmillä kierroksilla. Algoritmi jatkuu, kunnes jokainen hakija on hyväksytty johonkin kouluun tai jokainen hakija on ollut ehdolla kaikkiin vaihtoehtoihinsa (Abdulkadiroğlu & al. 2014).

³ Lisäksi vuodesta 2007 alkaen hakijoiden hakutoivejärjestys on otettu huomioon siinä tapauksessa, että kahdella rajalla olevalla hakijalla on sama lukuaineiden keskiarvo. Mikäli hakutoivejärjestys on sama, on otettu huomioon kaikkien aineiden keskiarvo (Lukiolaki 1998).

Lukioiden sisäänpääsyrajat syntyvät tämän prosessin seurauksena, jolloin kunkin koulun sisäänpääsyrajaksi muodostuu viimeisen hyväksytyen hakijan lukuaineiden keskiarvo, ja joidenkin lukioiden rajat nousevat hyvin korkeiksi. Tällaisista korkean keskiarvorajan eliittilukioista päädytään myös todennäköisemmin yliopistoon ja yliopiston eliittialueille. Mutta onko kyse vain hyvien opiskelijoiden valikoitumisesta eliittilukioihin vai tuottavatko eliittilukiot jonkinlaista arvonlisää, joka näkyy esimerkiksi suuremmassa todennäköisyydessä päästä yliopistoon?

Koska emme voi järjestää koeasetelmaa, jossa sijoittaisimme opiskelijoita satunnaisesti opiskelijaineikseltaan eritasoisiin lukioihin, meidän täytyy käyttää jotain muuta tapaa eliittilukion vaikutuksen arvioimiseksi. Tässä tutkimuksessa hyödynnetään ajatusta kontrafaktuaalista eli siitä, mitä eliittilukion opiskelijalle olisi tapahtunut, jos hän ei olisi päässyt eliittilukioon. Kullekin eliittilukioon hakeneelle voidaan siis kuvitella kaksi potentiaalista tulemaa kunkin tulostuuttujan, esimerkiksi yliopistoon pääsemisen, suhteen. Voidaan esimerkiksi ajatella, että opiskelija pääsee yliopistoon, jos hän tulee valituksi eliittilukioon, mutta ei pääse yliopistoon, jos hän jää eliittilukion ulkopuolelle. Tällöin eliittilukion vaikutus olisi yliopistoon pääseminen.

Eliittilukion vaikutus voidaan siis nähdä potentiaalisten tulemien erotuksena. Ongelmana on, että emme havaitse molempia potentiaalisia tulemia. Hakija joko tulee tai ei tule valituksi, emmekä täten voi verrata potentiaalisia tulemia yksilötasolla. Ratkaisu tähän on verrata sellaisia ryhmiä, joista toiset tulevat valituksi ja toiset eivät. Tämän vuoksi tässä tutkimuksessa hyödynnetään lukioiden sisäänpääsyrajoja käyttämällä regressioepäjatkuvuusasetelmaa (*regression discontinuity design, RDD*), jossa juuri sisäänpääsyrajan ylittämisen kontrafaktuaali muodostetaan niistä, jotka ovat jääneet niukasti rajan alapuolelle (Lee & Lemieux 2010).

Tällaisen ns. kvasikokeellisen asetelman avulla voidaan imitoida satunnaistettua koeasetelmaa, vaikka asetelma ei olisi kokeellinen (Angrist & Pischke 2010). Lukion sisäänpääsyrajan määrätymisessä voidaan olettaa olevan satunnaisuutta, sillä lukioilla on vain tietty määrä paikkoja, jolloin rajaksi määrätty viimeisen paikan saavan hakijan lukuaineiden keskiarvo. Vaikka hakija voi aikaisempien vuosien sisäänpääsyrajojen perusteella arvioida, mihin raja sijoittuu, on niissä vuosikoh-

taista vaihtelua (Tervonen & al. 2017), eikä tarkkaa rajaa voi tietää ennakkoon. Lisäksi hakijat eivät pysty päättämään keskiarvojaan itse, vaan ne perustuvat peruskoulun opettajien antamiin arvosanoihin. Myös näissä arvosanoissa voidaan olettaa olevan esimerkiksi opettajan arvostelusta johtuvaa satunnaisuutta. Täten hakija ei voi tietää varmasti, tuleeko hän valituksi vai ei, ja eliittilukioon pääseminen on rajan ympäristössä ikään kuin satunnaistetun koeasetelman määräämää (Lee & Lemieux 2010).

RDD:n ideana onkin, että ”käsittely” (*treatment*) on ainakin osaksi sen määräämää, kummalle puolelle rajaa käsittelyn määräävän muuttujan (peruskoulun päättötodistuksen lukuaineiden keskiarvo) arvo sijoittuu. Jos muuttujan arvo on rajalla tai niukasti sen yläpuolella, kuuluu yksilö käsittelyryhmään, kun taas niukasti rajan alapuolelle jäävät kuuluvat kontrolliryhmään. Käsittelyn vaikutus on näiden ryhmien ero jonkin tulostuuttujan ehdollisessa odotusarvossa rajalla (Imbens & Lemieux 2008). Täten sisäänpääsyrajan ylittämisen vaikutus todennäköisyyteen päästä yliopistoon saadaan juuri rajan yläpuolella olevan joukon sisäänpääsytodennäköisyyden ja juuri rajan alapuolella olevan joukon sisäänpääsytodennäköisyyden erotuksesta.

Koska estimointi perustuu rajan ympäristöön, tulee määritellä, mikä alue on riittävän lähellä rajaa, jotta sitä voidaan pitää rajan ympäristönä. Asetelmassa tulee siis valita ympäristö eli ikkunan (*bandwidth*) koko. Tässä artikkelissa on käytetty erästä Sebastian Calonicon ja kollegoiden (2017) esittämää tapaa ja valittu ikkunan koko keskineliövirheoptimaalisesti. Valinta on tehty kunkin lukion kullekin tulostuuttujalle erikseen.

Aineisto

Tutkimuksen aineistona käytämme toisen asteen yhteisvalinta-aineistoa vuosilta 2000–2008, yhdistettyä työntekijä–työnantaja-aineistoa (FLEED) vuosilta 1988–2011 sekä opiskelijarekisteriä vuosilta 2000–2012. Yhteisvalinta-aineisto koostuu muun muassa hakijoiden hakukohteista, peruskoulun päättötodistuksen keskiarvosta ja valintatiedoista. FLEED sisältää tietoa muun muassa koulutuksesta ja tuloista, ja sen avulla olemme voineet määritellä hakijoiden odotetut tulot ja heidän vanhempiansa koulutustason. Opiskelijarekisteristä havaitsemme, missä kukin opiskelija on kuna-

kin vuonna opiskellut. Viimeisten yhteisvalintavuosien opiskelijarekisterin tiedot ovat kuitenkin puutteellisia, sillä opiskelijan on tullut päästä yliopistoon vuoteen 2012 mennessä, jotta hänet nooterataan tässä tutkimuksessa yliopisto-opiskelijana. Esimerkiksi vuonna 2008 yhteisvalintaan osallistuneet ovat pystyneet hakemaan korkeakouluhin vasta vuonna 2011 (olettaen, että he ovat suorittaneet lukion kolmessa vuodessa), ja he ovat saattaneet päästä yliopistoon vasta vuoden 2012 jälkeen. Aineistojen yhdistämisestä ja duplikaatihavaintojen poistamisesta on kerrottu liitteessä.⁴

Näiden aineistojen lisäksi olemme käyttäneet tietoja Helsingin ja Espoon lukioiden sisäänpääsyrajoista vuosilta 2000–2008.⁵ Kullekin lukiolle on laskettu näiden vuosien ajalta keskimääräinen sisäänpääsyraja, jonka mukaan lukiot on asetettu järjestykseen korkeimmasta rajasta matalimpaan. Helsingissä keskimääräisten sisäänpääsyrajojen erotus on suurin neljänneksi ja viidenneksi sijoittuneiden lukion välillä, joten neljän ensimmäisen lukion (Ressu, Märsky, Viikki ja Norssi) valitseminen on luontevaa. Espoossa ensimmäiseksi ja toiseksi sijoittuneiden lukioiden välillä erotus on toiseksi suurin, mutta suurin ero tulee vasta kahdeksanneksi ja yhdeksänneksi sijoittuneiden lukioiden kohdalla, joten pelkästään ensimmäiseksi sijoittunut lukio (Etelä-Tapiola) erottuu eliittilukiona. Parhaan tietomme mukaan Vantaalla ja Kauniaisissa ei ole vastaavia korkean keskiarvorajan lukiota, ja Etelä-Tapiolan keskimääräinen sisäänpääsyraja olisi Helsingissä viidenneksi korkein, joten käsittelemämme lukiot ovat olleet vuosina 2000–2008 yleislinjoiltaan viisi korkeimman sisäänpääsyrajan lukiota pääkaupunkiseudulla.

Kunkin lukion kohdalla teemme viivästetyn hyväksynnän algoritmin (Abdulkadiroğlu & al. 2014) mukaisen oletuksen, jonka mukaan tarjouksen on saanut jokainen kyseessä olevaan lukioon ehdolla ollut ja sisäänpääsyrajan ylittänyt. Tämän tutkimuksen tulokset perustuvat näihin oletettuihin tarjouksiin. Aineistosta havaitut tarjoukset eroavat kuitenkin hieman oletetuista tarjouksista, mikä voi johtua jonkinlaisesta mittavirheestä tai sääntöjen vastaisesta toiminnasta opiskelijavalinnassa. Nämä havaitut tarjoukset on esitetty liitekuviossa 1 (kuviossa ”Eliitti” viittaa yhdistet-

tyyn estimointiin).

Tutkimuksen analyysi keskittyy Atila Abdulkadiroğlun ja kollegoiden (2014) tapaan ”jyrkkiin otoksiin”. Jyrkät otokset koostuvat kunkin lukion hakijoista, jotka a) hakevat kyseiseen lukioon ensimmäisenä vaihtoehtonaan, b) eivät tule valituksi ensimmäiseen vaihtoehtoonsa ja hakevat kyseiseen lukioon toisena vaihtoehtonaan, c) eivät tule valituksi kahteen ensimmäiseen vaihtoehtoonsa ja hakevat kyseiseen lukioon kolmantena vaihtoehtonaan ja niin edelleen aina viidenteen vaihtoehtoon saakka. Yhdistetty estimointi sisältää kaikki jyrkät otokset, jolloin yksittäinen hakija voi esiintyä estimoinnissa useampana havaintona. Täten yhdistetyssä estimoinnissa keskivirheet on klusteroitu yksilötasolla.

Tulokset

Eliittilukion aloittaminen ja vertaisryhmä

Taulukossa 2 on esitetty, kuinka paljon kunkin eliittilukion sisäänpääsyrajan ylittäminen vaikuttaa todennäköisyyteen aloittaa kyseisessä lukiossa. Nämä ns. first stage -tulokset ovat samoja kuin aikaisemmassa, eliittilukioiden vaikutuksia yo-kirjoitusten tuloksiin käsittelevässä tutkimuksessa (Tervonen & al. 2017), koska sen tutkimusasetelma ja yhteisvalinta-aineisto ovat identtisiä tämän tutkimuksen kanssa. Havaitsemme, että jokaisessa lukiossa sisäänpääsyrajan ylittäminen lisää todennäköisyyttä aloittaa kyseisessä lukiossa. Yksittäisissä lukioissa todennäköisyys kasvaa 35,7–84,0 prosenttiyksikköä ja yhdistetyssä estimoinnissa 43,3 prosenttiyksikköä. Kuvio 1 havainnollistaa tilannetta yhdistetyn estimoinnin osalta.⁶ Tässä, kuten artikkelin muissakin kuvioissa, ”standardoitu lukuaineiden keskiarvo” kertoo etäisyyden sisäänpääsyrajasta ja esimerkiksi arvo 10 tarkoittaa 0,1 arvosanaa rajan yläpuolella. Kuviosta nähdään, että aloittamisen todennäköisyys on juuri rajan alapuolella hieman alle 0,4 ja rajan yläpuolella hieman yli 0,8. Todennäköisyys aloittaa eliittilukiossa hyppää siis sisäänpääsyrajalla, mutta aloittamistodennäköisyys ei muutu rajalla jyrkästi nolasta yhteen. Tämä viittaa siihen, että osa rajan ylittäneistä

4 Kaikki liitekuviot ja -taulukot löytyvät artikkelin yhteydestä osoitteesta www.julkari.fi.

5 Alkuperäisiä sisäänpääsyrajoja on korjattu, mikäli yhteisvalinta-aineistosta on havaittu, että raja on todennäköisesti virheellinen (Tervonen & al. 2017).

6 Tämän artikkelin RDD-kuvioissa havainnot on jaettu ”laatikoihin”, joiden otoskeskiarvoja pisteet edustavat. Laatikoiden määrä on pidetty mahdollisimman suurena niin, että kokonaiskuva pysyy informatiivisena. Lokaalit lineaariset estimaatit on kuvattu regressiosuorilla rajan lähellä olevia havaintoja painottaen.

Taulukko 2. Rajan ylittämisen vaikutus eliittilukion aloittamiseen ja vertaisryhmään

Tulosmuuttuja	E-T	Norssi	Viikki	Märsky	Ressu	Eliitti
Aloittaminen	0,761*** (0,056)	0,615*** (0,088)	0,840*** (0,051)	0,512*** (0,112)	0,357*** (0,101)	0,433*** (0,055)
N	2 001	1 064	1 233	1 045	2 392	7 735
Vertaisryhmän ominaisuudet						
Lähtötaso	31,61*** (6,344)	37,92*** (10,53)	46,23*** (7,193)	34,86*** (12,45)	20,81** (8,772)	28,81*** (4,478)
Lähtötason keskihajonta	-8,310*** (1,683)	-5,776* (3,018)	-14,52*** (2,276)	-1,125 (3,840)	-7,193*** (2,594)	-8,480*** (1,088)
Tyttöjen osuus	0,095*** (0,014)	-0,057** (0,028)	-0,005 (0,024)	0,051** (0,024)	0,035** (0,015)	0,034*** (0,009)
Korkeakoulutettujen isien osuus	0,007 (0,013)	0,063** (0,030)	0,050** (0,022)	0,016 (0,025)	0,015 (0,021)	0,021* (0,012)
Korkeakoulutettujen äitien osuus	0,039*** (0,010)	0,049 (0,030)	0,039** (0,019)	0,069*** (0,022)	0,016 (0,017)	0,039*** (0,009)
N	1 967	1 042	1 207	1 008	2 316	7 540

Keskivirheet on esitetty suluissa.

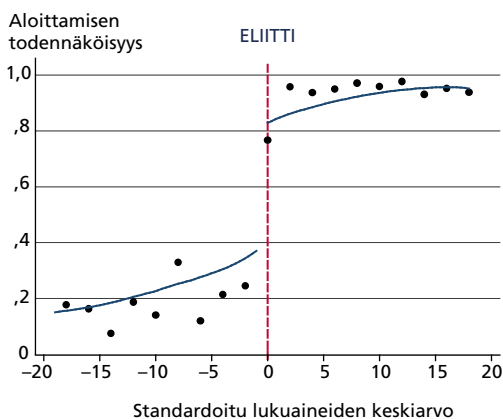
*** p<0,01, ** p<0,05, * p<0,1

hakijoista ei aloita ja osa rajan alapuolelle jääneistä hakijoista puolestaan aloittaa eliittilukiassa.

On huomioitava, että yhdistetyssä estimoinnissa yhden eliittilukion rajan alapuolelle jäänyt on voinut ylittää toisen eliittilukion sisäänkäysrajan ja myös aloittaa lukio-opinnot tässä toisessa lukiassa, jolloin hakija havaitaan sekä rajan ylä- että alapuolella. Tällaiset hakijat kattavat yhdiste-

tyn estimoinnin aineistosta kuitenkin alle 4 prosenttia. Suurempi syy asetelman epätäydellisyydelle voidaan havaita liitekuviosta 2, jossa havainnot on esitetty lukiokohtaisesti: kaikki rajan ylittäneet eivät aloita lukiassa, josta ovat saaneet tarjouksen, ja osa hakijoista onnistuu aloittamaan lukiassa, jonka rajan alapuolelle he ovat jääneet. Tällaista asetelmaa kutsutaan sumeaksi RDD:ksi: rajan ylittäminen määrää käsittelyn vain osaksi, mikä on verrattavissa koeasetelmaan, jossa osa koehenkilöistä ei noudata heille annettuja ohjeita (Lee & Lemieux 2010). Osa tarjouksen saaneista ei siis ota vastaan heille osoitettua ”eliittilukiokäsittelyä”, ja osa kontrolliryhmään kuuluvista onnistuu saamaan eliittilukiokäsittelyn tarjouksen puutteesta huolimatta. Asetelman epätäydellisyyden vuoksi joudumme tyytymään eliittilukiotarjouksen vaikutuksen estimoimiseen eliittilukiassa opiskelun vaikutuksen estimoimisen sijaan.

Taulukko 2 kuvaa myös sitä, kuinka vertaisryhmä muuttuu, kun hakija ylittää eliittilukion sisäänkäysrajan. Kunkin yksilön vertaisryhmällä tarkoitetaan tässä hänen kanssaan samalla koulutuslinjalla samaan aikaan aloittaneiden joukkoa, eli esimerkiksi vuonna 2000 Ressun yleislinjalla



Kuvio 1. Rajan ylittämisen vaikutus eliittilukion aloittamiseen.

aloittaneen vertaisryhmä koostuu muista samaan aikaan ko. linjalla aloittaneista.⁷ Rajan ylittäminen johtaa jokaisessa lukiossa vertaisryhmään, jossa lähtötaso (peruskoulun päättötodistuksen lukuaineiden keskiarvo) on keskimäärin korkeampi kuin rajan alle jäävien vertaisryhmässä. Muutos vertaisryhmän lähtötasossa on 0,21–0,46 arvosanaa. Tämä on seurausta korkeista sisäänpääsyrjoista, jolloin eliittilukioiden opiskelijoiden lähtötaso on keskimäärin korkeampi kuin niissä lukioissa, joihin rajan alapuolelle jäävät päätyvät.

Joissain lukioissa myös muut vertaisryhmän ominaisuudet muuttuvat. Tarkastelemme tässä vaikutusta vertaisryhmän lähtötason homogeenisuuteen, jonka mittarina käytämme lukuaineiden keskiarvon keskihajontaa. Lisäksi tutkimme vaikutusta sukupuolijakaumaan ja korkeakoulutettujen vanhempien osuuteen. Nämä ovat kiinnostavia muutoksia muun muassa siksi, että tutkimuksissa on havaittu vertaisryhmän homogeenisuuden olevan positiivisessa tilastollisessa yhteydessä äidinkielen yo-kirjoitusten tuloksiin (Kanninen 2013), tyttöjen osuuden kasvun vertaisryhmässä parantavan sekä tyttöjen että poikien oppimistuloksia (Lavy & Schlosser 2011) ja korkeakoulutettujen isien lapsien valikoituvan yliopistoihin keskimääräistä todennäköisemmin (Nori 2011).

Lähtötason keskihajonta pienenee jokaisessa lukiossa, joskin estimaatti on tilastollisesti merkitsevä vain Ressun, Viikin ja Etelä-Tapiolan lukioissa. Rajan ylittäminen siis johtaa lähtötasoltaan homogeenisempaan vertaisryhmään. Recessu, Märskysssä ja Etelä-Tapiolassa rajan ylittäminen johtaa suurempaan tyttöjen osuuteen vertaisryhmässä, kun taas Norssissa vaikutus tyttöjen osuuteen on negatiivinen. Kun tarkastellaan korkeakoulutettujen⁸ vanhempien osuutta vertaisryhmässä, käy ilmi, että estimaatit ovat positiivisia jokaisessa lukiossa. Isillä ne ovat tilastollisesti merkitseviä Norssissa ja Viikissä, äideillä taas Viikissä, Märskysssä ja Etelä-Tapiolassa. Vertaisryhmän ominaisuudet siis muuttuvat eliittilukioiden sisäänpääsyrjalla pääasiallisesti oppimistuloksia ja yliopistoon valikoitumista edistävasti.

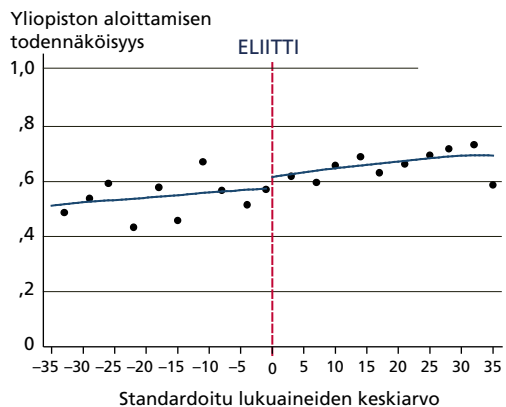
7 Aloittamisesta havaitsemme vain oppilaitoksen – emme aloittamislinjaa. Tarjouksesta havaitsemme kuitenkin myös linjan, joten oletamme sen olevan aloittamislinja, mikäli hakija on aloittanut ko. oppilaitoksessa. Jos linjan tunnistaminen ei ole ollut mahdollista, on vertaisryhmänä käytetty koko koulun vuosikurssia.

8 Korkeakoulutetulla viitataan tässä henkilöön, joka on suorittanut maisterin tai sitä korkeamman tutkinnon.

Yliopiston aloittaminen ja opintoalan valinta

Kuviossa 2 on esitetty, kuinka eliittilukion sisäänpääsyrjan ylittäminen vaikuttaa todennäköisyyteen aloittaa yliopistossa (tämä on esitetty lukio-kohtaisesti liitekuviossa 3). Rajalla on nähtävissä pieni positiivinen hyppäys, joka ei kuitenkaan ole tilastollisesti merkitsevä, kuten päätulokset esittävästä taulukosta 3 voidaan nähdä. Tämän perusteella ei siis voida sanoa, että eliittilukioon pääseminen johtaisi todennäköisempään opiskelupaikkaan yliopistossa. Tässä taulukossa on lisäksi esitetty myös yksittäisten lukioiden tulokset. Norssin ja Etelä-Tapiolan estimaatit ovat positiivisia ja melko suuria, ja Etelä-Tapiolan estimaatti on myös tilastollisesti merkitsevä. Koska Etelä-Tapiolan sisäänpääsyrjan ylittäminen ei aikaisemman tutkimuksen (Tervonen & al. 2017) mukaan johda parempiin yo-kirjoitusten arvosanoihin, voi olla, että Etelä-Tapiolan lukiossa opitaan jotain sellaista, josta ei ole hyötyä yo-kirjoituksissa, mutta joka auttaa pääsemään yliopistoon. Tuloksia tulkittaessa tulee kuitenkin muistaa, että yksittäiset merkitsevät tulokset voivat johtua myös satumasta.

Vaikka eliittilukioiden rajan ylittävät opiskelijat eivät päätyisi yliopistoon kontrolliryhmään kuuluvia useammin, voi eliittilukiolla olla vaikutusta heidän alavalintaansa. Taulukko 3 kertoo myös rajan ylityksen vaikutuksen eliittialoille – eli lääketieteelliselle, oikeustieteelliselle, kauppatieteelliselle tai teknillistieteelliselle alalle – päättämiseen. Tarkastelu on tehty vain yhdistetysti, sillä yksittäisissä lukioissa joidenkin alojen opiskelijamäärät ovat hyvin pieniä. Kaikki estimaatit ovat positiivi-



Kuvio 2. Rajan ylittämisen vaikutus yliopistossa aloittamiseen.

Taulukko 3. Rajan ylittämisen vaikutus yliopistossa ja eliittialoilla aloittamiseen

Tulosmuuttuja	E-T	Norssi	Viikki	Märsky	Ressu	Eliitti
Yliopisto	0,131** (0,064)	0,209 (0,152)	0,069 (0,092)	-0,073 (0,107)	0,027 (0,073)	0,040 (0,039)
Kauppatieteellinen	-	-	-	-	-	0,013 (0,029)
Lääketieteellinen	-	-	-	-	-	0,022 (0,016)
Oikeustieteellinen	-	-	-	-	-	0,006 (0,016)
Teknillistieteellinen	-	-	-	-	-	0,012 (0,024)
N	1 993	1 060	1 225	1 037	2 372	7 687

Keskivirheet on esitetty suluisa.

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$

sia, mutta yksikään niistä ei ole tilastollisesti merkitsevä. Kuitenkin lääketieteen estimaatti on melko suuri (0,022), sillä taulukosta 1 voidaan laskea, että kaikista eliittilukioiden opiskelijoista vain noin seitsemän prosenttia päätyy lääketieteelliselle alalle. Melko suuretkin vaikutukset ovat täten lääketieteessä mahdollisia. Emme kuitenkaan löydä selvää näyttöä siitä, että eliittilukion sisäänpääsyrajan ylittäminen johtaisi korkeampaan todennäköisyyteen päästä millekään eliittialalle. Tarkastelu tehtiin myös käyttäen Tilastokeskuksen koulutusluokitusta, joka jaottelee yliopistokoulutuksen kahdeksaan eri alaan (kasvatustieteellinen ja opettajankoulutus, humanistinen ja taideala, kaupallinen ja yhteiskuntatieteellinen, luonnontieteellinen, tekniikka, maa- ja metsätalousala, terveystieteellinen ja sosiaaliala ja palveluala), mutta eliittilukiotarjous ei tuottanut tilastollisesti merkitseviä vaikutuksia myöskään tällä jaottelulla. Nämä tulokset on esitetty liitetaulukossa 1 (pl. teknillistieteellinen ala, joka on esitetty jo taulukossa 3).

On joitain seikkoja, jotka voisivat selittää tilastollisesti merkitsevien vaikutusten puuttumisen. Ensimmäinen on se, että käytämme tulosmuuttujana yliopiston aloittamista, mutta emme huomioi välivuusia. Voimme kuitenkin tutkia, onko rajan ylittämisen vaikutusta välivuosiin. Välivuosien pituudeksi määrittelimme yliopisto-opintojen aloittamisvuoden ja oletetun abivuoden (kolme vuotta yhteishausta) erotuksen. Tulosten mukaan rajan ylittävät eivät keskimäärin myöskään aloita opintojaan rajan alapuolelle jääneitä nopeammin – tämä on esitetty taulukossa 4. Tämän vuok-

si myöskään aineisto-osiossa mainitun opiskelutietojen puuttumisen vuoden 2012 jälkeen ei siis pitäisi aiheuttaa ongelmaa, sillä seuraavien vuosien opiskelutietojen lisäys lisäisi todennäköisesti yliopisto-opiskelijoiden osuutta samassa suhteessa sekä rajan ylittävien että rajan alapuolelle jäävien joukossa. Yksittäisten lukioiden kohdalla on vaihtelua eivätkä estimaatit muodosta selkeää linjaa. Kuitenkin Viikin estimaatti on myös tilastollisesti merkitsevä, ja sen mukaan tarjouksen saaminen Viikistä johtaa n. 0,7 lisäviljuuteen.

Toiseksi emme pysty havaitsemaan ulkomaisiin yliopistoihin päätyviä aineistostamme. Jos siis eliittilukioon niukasti pääsevät opiskelijat päätyvät juuri rajan alle jääneitä useammin ulkomaisiin yliopistoihin, saattaa tuloksemme aliarvioida eliittilukion vaikutusta todennäköisyyteen aloittaa yliopisto-opinnot. Emme myöskään ole huomioineet kotimaisten yliopistojen tasoa. Jos esimerkiksi ajatellaan, että pääkaupunkiseudun yliopistot ovat tasokkaampia kuin muun Suomen yliopistot ja eliittilukion rajan ylittäminen lisää todennäköisyyttä päästä ensin mainittuihin, voidaan eliittilukioon pääseminen nähdä hyödyllisenä. Havaitsemme, että tällainen vaikutus on olemassa: rajan ylittäminen lisää todennäköisyyttä päästä pääkaupunkiseudulla sijaitsevaan yliopistoon 7,7 prosenttiyksikköä, kuten taulukosta 4 voidaan nähdä. Yksittäisten lukioiden estimaatit ovat positiivisia Märskyä lukuun ottamatta, ja Norssin estimaatti on sekä varsin suuri että tilastollisesti merkitsevä. Jos eliittilukioihin pääsevät opiskelijat myös hakeutuvat näihin yliopistoihin rajan alle jäänei-

Taulukko 4. Rajan ylittämisen vaikutukset muihin lukion jälkeisiin koulutustulemiin

Tulosmuuttuja	E-T	Norssi	Viikki	Märsky	Ressu	Eliitti
Välivuodet	-0,200 (0,262)	-0,260 (0,588)	0,707** (0,277)	-0,198 (0,320)	0,346 (0,304)	0,049 (0,126)
N	1 151	612	685	560	1 650	4 658
Pääkaupunkiseudun yliopisto	0,123* (0,069)	0,344** (0,159)	0,126 (0,083)	-0,087 (0,103)	0,108 (0,082)	0,077** (0,036)
N	1 993	1 060	1 225	1 037	2 372	7 687
Odotetut tulot	2 532* (1 295)	2 398 (2 687)	478,1 (1 510)	1 832 (2 432)	-986,6 (1 372)	460,4 (798,8)
N	1 995	1 061	1 225	1 038	2 381	7 700

Keskivirheet on esitetty suluissa.

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$

tä todennäköisemmin, voi vaikutuksen puute selittyä sillä, että pääkaupunkiseudun yliopistoihin on mahdollisesti vaikeampaa päästä kuin muun Suomen yliopistoihin.

Vaikka eliittilukiosta ei olisi hyötyä yliopistoon pääsemisessä, saattaa sillä olla vaikutuksia muihin asioihin, kuten tulevaisuudessa saataviin tuloihin. Tällainen on mahdollista, jos eliittilukioiden opiskelijat esimerkiksi hyötyvät lukioaikaisista verkostoista tai eliittilukio saa heidät hakeutumaan korkeapalkkaisille aloille. Emme kuitenkaan voi tutkia eliittilukion vaikutusta tuloihin siksi, että vuosina 2000–2008 yhteisvalintaan osallistuneet ovat pääasiassa liian nuoria, jotta heidän aikuisikänsä tuloja voisi tutkia luotettavasti. Tämän vuoksi tutkimme rajan ylittämisen vaikutusta odotettuihin tuloihin, joiden avulla voimme kuvata, johtaa-ko eliittilukioon pääseminen korkeatuloisemman opintoalan valintaan.

Hakijoiden odotetut tulot saamme vuoden 2011 FLEED-aineistosta laskemalla keskimääräiset ansiotulot 30–34-vuotiaille korkeimman suoritetun tutkinnon perusteella. Käytämme tätä lukua mittarina odotetuille tuloille kuhunkin koulutukseen päässeille. Jos hakija on siis päässyt opiskelemaan esimerkiksi lääketiedettä, ovat hänen odotetut tulonsa yhtä suuret kuin keskimääräiset lääketieteellisen tutkinnon suorittaneiden ansiotulot vuoden 2011 FLEED-aineistossa.⁹ Opintoaloissa olemme käyttäneet tässä samaa jakoa kuin Sari Pekkala Kerr ja kollegat (2018). Tässä luokit-

9 Mikäli havaitaan, että henkilö on opiskellut useammasa kuin yhdessä oppilaitoksesta, on näistä valittu korkeasteisin. Mikäli samantasoisia opiskelupaikkoja on useampi, on valittu se, josta saatavan tutkinnon odotettu tulotaso on korkein.

telussa yliopistoalat on jaettu 15 ja AMK-alat seitsemään eri luokkaan. Yliopistoaloilla on laajempien alojen (esim. yhteiskuntatieteet tai luonnontieteet) lisäksi tehty omia luokkia tarkemmille aloille (esim. lääketiede ja oikeustiede). Lisäksi omana luokkana on käsitelty niitä, jotka eivät ole päätyneet mihinkään kolmannen asteen oppilaitokseen. Tämä luokittelu on esitetty liitteessä.

Kuten taulukosta 4 voidaan nähdä, yhdistetystä estimoinnista odotettujen tulojen estimaatti on positiivinen, mutta se ei eroa tilastollisesti merkittävästi nolasta. Yksittäisten lukioiden kohdalla vaihtelua esiintyy, mutta estimaatit eivät ole tilastollisesti merkittäviä, joskin Etelä-Tapiolan estimaatti viittaa positiiviseen vaikutukseen. Yleisellä tasolla emme siis löydä perusteita sille, että eliittilukioon pääseminen johtaisi alavalintaan, jossa tulot olisivat keskimäärin korkeammat kuin niillä aloilla, joille eliittilukion ulkopuolelle jääneet ovat päätyneet. Estimoinnit on tehty myös käyttämällä kutakin tutkintokoodia omana alanaan, mutta tulokset eivät ole tilastollisesti merkittäviä tässäkään tapauksessa.¹⁰ Tämä lähestymistapa on myös hieman ongelmallinen, sillä joitain tutkintokoodeja on vuoden 2011 FLEED-aineistossa vain vähän – joissain tapauksissa ei lainkaan – jolloin odotettujen tulojen estimaatissa on epä-tarkkuutta.

Asetelman validiteetti ja tulosten robustisuus

Tässä osiossa osoitetaan, että RDD on soveltuva eli validi lähestymistapa tutkittavaan kysymykseen. Lisäksi tehdään erilaisia robustisuustarkaste-

10 Tuloksia ei ole esitetty tässä, mutta ne ovat saatavilla kirjoittajilta.

luja, joiden avulla esitetään, että tutkimuksen tulokset eivät ole riippuvaisia tekemistämme tutkimusasetelmaan liittyvistä valinnoista.

Tässä tutkimuksessa käytetyn tutkimusasetelman validiteetti voisi olla vaarassa, mikäli hakijoilla olisi mahdollisuus manipuloida käsittelyn määräävää muuttujaa (lukuaineiden keskiarvo) niin, että he tulisivat varmasti valituksi (Lee & Lemieux 2010). Jos esimerkiksi manipulaatiossa onnistuminen korreloisi negatiivisesti yliopisto-opintojen aloittamisen kanssa, voisi eliittilukion positiivisen vaikutuksen puuttuminen selittyä tällä valikoitumisella. Koska hakijan pitäisi tietää eliittilukion sisäänpääsyraja ennakkoon ja pystyä määräämään lukuaineiden keskiarvonsa tarkkaan onnistuakseen manipulaatiossa, emme pidä sitä todennäköisenä.

Tästä huolimatta manipulaation mahdollista esiintymistä voidaan tutkia. Esimerkiksi sisäänpääsyrajalla tapahtuva hakijamassan hyppäys kertoisi mahdollisesta manipulaatiosta (Imbens & Lemieux 2008). Tällöin juuri rajan yläpuolella havaittaisiin selvästi enemmän hakijoita kuin rajan alapuolella. Yleinen tapa tämän tutkimiseen on käyttää Justin McCraryn (2008) esittämää testiä. Tällaisesta testistä ei ole kuitenkaan hyötyä, jos epäjatkuvuuden tiedetään johtuvan jostain muusta asiasta kuin manipulaatiosta (Zimmerman 2014). Tilanne on juuri tällainen asetelmassamme, sillä jotkut keskiarvot ovat yleisempiä kuin toiset. Pelkkä hakijamassan hyppäys rajalla ei siis välttämättä kerro mitään manipulaatiosta (ks. tarkemmin Tervonen & al. 2017). Tämän vuoksi joudumme hyödyntämään muita tapoja asetelman validiteetin tutkimiseksi.

Koska RDD on kvasikokeellinen asetelma, sitä voidaan analysoida ja testata ikään kuin kyseessä olisi satunnaistettu koeasetelma. Täten käsittely- ja kontrolliryhmien tulee olla lähtökohdiltaan keskimäärin samanlaisia. RDD:ssä tämä tarkoittaa, että muuttujien, jotka ovat määrättyneet ennen valintatilannetta, tulisi olla jatkuvia sisäänpääsyrajalla (Lee & Lemieux 2010). Mahdollista manipulaatiota voidaan siis tutkia testaamalla, ”vaikuttaako” rajan ylitys johonkin muuttujaan, johon sen ei tiedetä vaikuttavan (Imbens & Lemieux 2008). Tässä tutkimuksessa tällaisina muuttujina on käytetty peruskoulun päättötodistuksen taito- ja taidaineiden arvosanoja, sukupuolta ja vanhempien koulutustasoa. Yksittäisten lukiodien kohdalla ainoastaan kaksi estimaattia neljästäkymmenestä on merkitseviä viiden prosentin merkitsevyystasolla, joten testi ei viittaa manipulaation olemas-

saaloon – viiden prosentin merkitsevyystasolla on odotettavissa, että viisi prosenttia regressioista on merkitseviä sattumalta. Nämä tulokset on esitetty liitetaulukossa 2.

Asetelmaa voidaan testata lisäksi tutkimalla sitä, havaitsemme ”vaikutusta” paikoissa, joissa sen pitäisi olla nolla. Tällaisten ”plaseborajojen” kohdalla tulomuuttujan ei tulisi hypätä, sillä niissä rajan ympäristössä olevat eivät eroa toisistaan käsittelyn suhteen kuten oikean sisäänpääsyrajan ympäristössä olevat eroavat. Kyseessä oleva testi voidaan tehdä esimerkiksi jakamalla aineisto kahteen osajoukkoon sisäänpääsyrajan mukaan ja käyttämällä osajoukkojen standardoidun lukuaineiden keskiarvon mediaaneja uusina rajoina (Imbens & Lemieux 2008). Plasebotarkastelu tehdään tässä vain yhdistetylle estimoinnille. Sisäänpääsyrajan vasemmanpuoleiselle osajoukolle plaseborajaksi määritetty $-0,42$ ja oikeanpuoleiselle $0,37$. Näissä pisteissä hyppyt eivät eroa tilastollisesti merkitsevästi nollassa, kuten voidaan nähdä liitetaulukosta 3.

Tekemiemme testien mukaan ei siis ole syytä epäillä, etteikö tutkimusasetelma olisi validi. Asetelman validiteetin tutkimisen lisäksi olemme tehneet analyysin tueksi funktiomuodon, kernel-funktion ja ikkunan koon robustisuustarkasteluja, sillä näihin liittyvät valinnat voivat vaikuttaa tuloksiin. Tutkimuksessa käytetyn lokaalin lineaarisen regression ja triangulaarisen kernel-funktion lisäksi pääestimoinnit on tehty myös käyttäen kvadraattista polynomifunktiota, jolloin tulokset ovat pääosin samansuuntaisia. Tätä korkeamman asteen polynomifunktioita ei ole käytetty niiden mahdollisen harhaanjohtavuuden vuoksi (Gelman & Imbens 2018). Kernel-funktion valinnalla ei yleensä ole suurta merkitystä (Lee & Lemieux 2010), ja tässäkin tapauksessa kaksi muuta vaihtoehtoa (tasainen ja Epanechnikov) tuottavat pääosin samansuuntaisia tuloksia kuin triangulaarinen kernel-funktio. Näiden lisätarkastelujen tuloksia ei ole esitetty tässä, mutta ne ovat saatavilla kirjoittajilta.

RDD:ssä käytettävä ikkuna vaikuttaa analysoitavien havaintojen määrään, joten estimointi on esitetty liitekuviassa 4 käyttäen erikokoisia ikkunoita. Piste-estimaatit on esitetty pisteillä ja 95 prosentin luottamusväli viivoilla. Pääestimoinnissa käytetyn ikkunan kohdalla luottamusväli on merkitty punaisella. Yhdistetyssä estimoinnissa piste-estimaatti pysyy hyvin samanlaisena eri ikkunakokoja käyttäen ja nolla sisältyy jokaiseen

luottamusväliin. Myös yksittäisten lukioiden estimaatit pysyvät ei-merkitsevinä lukuun ottamatta Etelä-Tapiolaa, jonka estimaatti oli päätuloksissakin positiivinen ja merkitsevä. Etelä-Tapiolan estimaatit ovat positiivisia ja merkitseviä jokaisella ikkunalla, jonka suuruus on 0,4 arvosanaa tai suurempi. Pienempiä ikkunoita käytettäessä estimaatti on pääosin positiivinen, mutta ei tilastollisesti merkitsevä.

Yhteenvedon voidaan siis todeta, että asetelmaan liittyvillä valinnoilla ei ole merkittävää vaikutusta tuloksiin. Robustisuustarkastelujen lisäksi olemme tehneet estimoinnin käyttäen rajan ylittämistä instrumenttimuuttujana havaitulle tarjoukselle, koska havaitut tarjoukset eivät ole täysin samoja kuin oletetut tarjoukset (ks. liitekuvio 1). Tämä kasvattaa yhdistetyn estimoinnin piste-estimaattia jonkin verran, mutta se pysyy ei-merkitseväenä, kuten liitetulokosta 3 voidaan nähdä. Oletus siitä, että kaikki rajan ylittäneet ovat saaneet tarjouksen, ei siis ole tulosten kannalta merkittävä.

Yhteenvedo

Tässä tutkimuksessa tarkasteltiin eliittilukion sisäänpääsyrajan ylittämisen vaikutusta yliopiston aloittamisen todennäköisyyteen ja opintoalan valintaan. Tutkimuksessa hyödynnettiin regressioepäjatkuvuusasetelmaa (RDD), jossa verrataan juuri sisäänpääsyrajan ylittäneitä niukasti sen alle jääneisiin. Tällaisen kvasikokeellisen asetelman tarkoituksena on estimoida eliittilukioon pääsemisen kausaalivaikutusta hyödyntäen sisäänpääsyrajan luomaa satunnaisuutta eliittilukiotarjouksissa.

Tutkimuksemme mukaan sisäänpääsyrajan ylittämällä ei ole tilastollisesti merkitsevää vaikutusta todennäköisyyteen aloittaa yliopistossa siitä huolimatta, että rajan ylittäminen johtaa lähtötasoltaan parempaan vertaisryhmään. Poikkeuksena tästä on Etelä-Tapiolan lukio, jonka rajan ylittäminen näyttää johtavan todennäköisempään yliopistopaikkaan. Tulokset ovat robusteja useille asetelmaan kuuluville valinnoille, kuten ikkunan koolle. Täten sillä, kuinka suurta sisäänpääsyrajan ympäristöä tutkitaan, ei ole merkittävää vaikutusta tuloksiin.

Tulostemme mukaan eliittilukioilla ei ole vaikutusta myöskään ns. eliittialojen – kauppatieteellisen, lääketieteellisen, oikeustieteellisen tai tek-

nillistieteellisen alan – aloittamiseen. Myöskään muilla aloilla vaikutusta ei havaita. Tätä tulosta tukee myös se, että emme havainneet eliittilukiotarjouksella olevan tilastollisesti merkitsevää vaikutusta odotettuihin tuloihin. Eliittilukiot eivät siis näytä lisäävän todennäköisyyttä päätyä korkeapalkkaiselle alalle. Sen sijaan todellisiin tuloihin eliittilukiolla voi olla vaikutusta esimerkiksi lukioaikana syntyneiden verkostojen kautta. Emme kuitenkaan pysty tutkimaan vaikutuksia todellisiin tuloihin, sillä aineistossamme olevat henkilöt ovat tähän tarkoitukseen liian nuoria.

Emme siis löydä näyttöä siitä, että eliittilukiot lisäävät koulutustasoon tai -alaan liittyvää eriytymistä. Vaikka eliittilukioiden opiskelijat ovat valikoinut joukko, tulostemme perusteella ei vaikuta siltä, että eliittilukiot itsessään vahvistaisivat koulutuspolkujen eriytymistä. Tämä on linjassa aikaisemman tutkimuksen kanssa, jonka mukaan eliittilukion sisäänpääsyrajan ylittäminen ei paranna ylioppilaskirjoitusten arvosanoja (Tervonen & al. 2017). Tuloksia tulkitessa tulee muistaa nolatuloksen lokaalisuus: tutkimuksen kohteena on vain rajan läheisyydessä olevat hakijat. Tuloksiin voi siis vaikuttaa myös esimerkiksi se, että tutkittavat eliittilukiokäsittelyn saavat opiskelijat ovat lähtötasoltaan vertaisryhmänsä heikoimpia. Tällä voi olla negatiivinen vaikutus oppimistuloksiin, tai kyseiset opiskelijat olisivat voineet hyötyä sellaisessa lukiossa olemisesta, jossa heidän suhteellinen ”sijoituksensa” olisi korkeampi (Cicala & al. 2018; Duflo & al. 2011; Murphy & Weinhardt 2018). Emme siis voi yleistää tulosta koskemaan kaikkia eliittilukioiden opiskelijoita. Esimerkiksi lähtötasoltaan vertaisryhmänsä parhaille eliittilukiosta voi olla hyötyä. Tulkinnassa tulee huomioda myös se, että emme pysty sulkemaan pois pieniä positiivisia tai pieniä negatiivisia vaikutuksia estimaattien epätarkkuuden vuoksi.

Tämän tutkimuksen tulokset viittaavat siihen, että eliittilukioiden opiskelijoiden muiden lukioiden opiskelijoita suuremmat todennäköisyydet päätyä yliopistoon ja eliittialoilta johtuvat hyvien opiskelijoiden valikoitumisesta eliittilukioihin. Kun valikoitumisharha poistetaan, katoavat myös erot. Eliittilukiot vaikuttaisivat olevan siis eliittia opiskelija-aineiksensa akateemisen ja mahdollisesti muun osaamisen perusteella – eivät tuottamansa arvonlisän suhteen.

KIRJALLISUUS

- Abdulkadiroğlu, Atila & Angrist, Joshua & Pathak, Parag: The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica* 82 (2014): 1, 137–196.
- Angrist, Joshua & Pischke, Jörn-Steffen: The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives* 24 (2010): 2, 3–30.
- Calonico, Sebastian & Cattaneo, Matias D. & Farrell, Max H. & Titiunik, Rocio: rdrobust: Software for regression-discontinuity designs. *The Stata Journal* 17 (2017): 2, 372–404.
- Cattaneo, Matias D. & Idrobo, Nicholas & Titiunik, Rocio: A Practical Introduction to Regression Discontinuity Designs: Volume I. Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge: Cambridge University Press, 2018 (tulossa).
- Cicala, Steve & Fryer, Roland G. Jr. & Spenkuch, Jörg L: Self-Selection and Comparative Advantage in Social Interactions. *Journal of the European Economic Association* 16 (2018): 4, 983–1020.
- Dobbie, Will & Fryer, Roland G. Jr: The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics* 6 (2014): 3, 58–75.
- Duflo, Esther & Dupas, Pascaline & Kremer, Michael: Peer effects, teacher incentives, and the impact of tracking. *American Economic Review* 101 (2011): 5, 1739–1774.
- Gelman, Andrew & Imbens, Guido: Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 2018 (tulossa).
- Hoekstra, Mark & Mouganie, Pierre & Wang, Yaojing: Peer Quality and the Academic Benefits to Attending Better School. Working Paper 22337. NBER Working Paper Series, 2016.
- Imbens, Guido W. & Lemieux, Thomas: Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142 (2008), 615–635.
- Kanninen, Ohto: Five Essays on Economics of Education. European University Institute, 2013.
- Kuusela, Jorma: Lukioiden tuloksiin vaikuttavista tekijöistä. *Moniste* 13/2003. Helsinki: Opetushallitus, 2003.
- Lavy, Victor & Schlosser, Analía: Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics* 3 (2011), 1–33.
- Lazear, Edward P: Educational Production. *The Quarterly Journal of Economics* 116 (2001): 3, 777–803.
- Lee, David S. & Lemieux, Thomas: Regression Discontinuity Designs in Economics. *Journal of Economic Literature* 48 (2010), 281–355.
- McCrary, Justin: Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142 (2008), 698–714.
- Murphy, Richard & Weinhardt, Felix: Top of the Class: The Importance of Ordinal Rank. NBER Working Paper No. 24958, 2018.
- Nori, Hanna: Keille yliopiston portit avautuvat? Tutkimus suomalaisiin yliopistoihin ja eri tieteenaloille valikoitumisesta 2000-luvun alussa. Turun yliopiston julkaisuja. Sarja C, osa 309, 2011.
- Pekkala Kerr, Sari & Pekkarinen, Tuomas & Sarvimäki, Matti & Uusitalo, Roope: Post-Secondary Education and Information on Labor Market Prospects: A Randomized Field Experiment. *Työpäperi*, 2018.
- Seppänen, Piia & Kalalahti, Mira & Rinne, Risto & Simola, Hannu (toim.): Lohkoutuva peruskoulu. Jyväskylä: Suomen Kasvatustieteellinen seura, 2015.
- Tervonen, Lassi & Kortelainen, Mika & Kanninen, Ohto: Eliittilukioiden vaikutukset ylioppilaskirjoitusten tuloksiin. VATT Tutkimukset 186. Helsinki, 2017.
- Thistlethwaite, Donald L. & Campbell, Donald T.: Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment. *The Journal of Educational Psychology* 51 (1960): 6, 309–317.
- Virtanen, Hanna: Essays on Post-Compulsory Education Attainment in Finland. Aalto University publication series. Doctoral Dissertations 87. Helsinki, 2016.
- Zimmerman, Seth D: The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics* 32 (2014): 4, 711–754.

ENGLISH SUMMARY

*Lassi Tervonen & Mika Kortelainen & Ohto Kanninen:
Effects of elite high schools on university enrolment
and field of study choice (Eliittilukioiden vaikutukset
yliopisto-opintojen aloittamiseen ja koulutusalan
valintaan)*

Finnish elite high school students enrol in university and so-called elite fields of study more often than Finnish high school students on average. However, those who attend elite high schools are also higher-achieving in terms of baseline grade point average (GPA) from comprehensive school. This selection bias must be taken into account in studying the causal effects of elite high schools.

This study focuses on five elite high schools in the Helsinki region and aims to solve the problem of selection bias by using a regression discontinuity design (RDD). In our case RDD exploits the entrance thresholds of elite high schools as a rule which assigns

applicants near the threshold into treatment and control groups. By comparing the outcomes (e.g. the probability of enrolment in a university) of these groups we can estimate the causal effects of an elite high school offer on various educational outcomes, such as university enrolment.

We find that crossing the threshold of an elite high school leads to a higher-achieving peer group in terms of baseline GPA. However, the elite high school offer does not have a statistically significant effect on the probability of enrolment in a university or on the probability of enrolment in an elite field of study. The only exception is Etelä-Tapiola high school, which has a positive effect on the probability of enrolment in a university.

Keywords: education, regression discontinuity design, peer effects, school choice.