

# OPINTOMENESTYKSEN ENNUSTAMINEN JA OPIKELIJAVALINNAN VAIKUTUKSEN ARVIOINTI AALTO-YLIOPISTON KAUPPAKORKEAKOULUSSA

Maisterin tutkielma  
Emma Mieskonen  
Aalto-yliopiston kauppakorkeakoulu  
Taloustiede  
Syksy 2017

---

**Tekijä** Emma Mieskonen

---

**Työn nimi** Opintomenestyksen ennustaminen ja opiskelijavalinnan vaikutuksen arviointi Aalto-yliopiston kauppakorkeakoulussa

---

**Tutkinto** Kauppatieteiden maisteri

---

**Koulutusohjelma** Taloustiede

---

**Työn ohjaaja** Matti Sarvimäki

---

**Hyväksymisvuosi** 2017**Sivumäärä** 78**Kieli** Suomi

---

### Tiivistelmä

Suomen korkeakoulujen opiskelijavalintaa on kritisoitu tehottomuudesta vuosikausia, ja valintakokeita painottavan nykyjärjestelmän ongelmat aiheuttavat ylimääräisiä kustannuksia hakijoille, korkeakouluille ja yhteiskunnalle. Ratkaisuksi on tarjottu ylioppilastutkinnon painottamista opiskelijavalinnoissa. Opetus- ja kulttuuriministeriön onkin esittänyt, että pitkää valmentautumista vaativista opiskelijavalinnoista luovutaan, ja vuoteen 2020 mennessä pääosa uusista opiskelijoista valita suoraan ylioppilastutkintoon perustuen. Opiskelijavalintaa ja opintomenestyksen ennustamista on Suomessa tutkittu runsaasti, mutta tutkimusta valintamenettelyjen vaikutuksesta valitun opiskelijajoukon opintomenestykseen ei ole juuri tehty.

Tutkielman tavoitteena on esitellä empiirinen menetelmä, jolla vaihtoehtoisia valintamenettelyitä on mahdollista arvioida opiskelijavalinnan järjestävän korkeakoulun näkökulmasta. Käytännöllisenä tavoitteena on selvittää, miten valintakokeesta luopuminen tai sen painoarvon pienentäminen voisi vaikuttaa opiskelijavalinnan tuloksiin Aalto-yliopiston kauppakorkeakoulussa. Tutkielma vastaa seuraaviin kysymyksiin: 1) Miten ylioppilastutkintoarvosanat ja valintakoe ennustavat opintomenestystä? 2) Millainen ylioppilastutkintoarvosanojen valintapisteytysmallin tulisi niiden ennustearvon mukaan olla? 3) Kuinka eri ihmiset tulisivat valituksi, jos käytössä olisi pelkästään ylioppilastutkintoon perustuva valinta? 4) Miten ylioppilastodistuksen perusteella valitut opiskelijat menestyisivät verrattuna niihin opiskelijoihin, jotka eivät olisi tulleet valituksi pelkkään ylioppilastutkintoon perustuvassa valinnassa? Aineistona käytetään Aalto-yliopiston kauppakorkeakouluun vuosina 2005 – 2012 valittujen opiskelijoiden opiskelija- ja kursssirekisterimerkintöjä.

Opiskelijavalinnan ongelmana on valita parhaat mahdolliset opiskelijat mahdollisimman pienin resurssi-investoinnein. Opintomenestys määritellään tutkielmassa Aalto-yliopiston kauppakorkeakoulun käyttämiin opintomenestysindekseihin perustuen arvosanakeskiarvoilla ja opintopistekertymillä. Kyvykkyyttä korkeakouluopintoihin mitataan ylioppilastutkintoarvosanoilla ja valintakoe pisteillä. Tutkielman tulokset ovat linjassa aiemman kirjallisuuden kanssa osoittaessaan, että ylioppilastutkintoarvosanat ennustavat OLS-regressiossa korkeakouluopintojen kurssiarvosanojen muodostumista varsin hyvin. Erityisesti matemaattinen osaaminen indikoi menestymistä kauppatieteellisissä opinnoissa. Opintopistekertymien ennustaminen valintakriteereillä ei kuitenkaan ole mielekästä. Valintakokeen merkitys opintomenestyksen ennustamiseen jää mitättömän pieneksi. Marginaalisten opiskelijoiden opintomenestystä verratessa todetaankin, että erillisen valintakokeen järjestäminen ei tuo opiskelijavalintaan lisäarvoa Aalto-yliopiston kauppakorkeakoulussa: puhtaalla todistusvalinnalla olisi kouluun valittu hyvin samankaltainen, jopa hieman paremmin menestynyt opiskelijajoukko. Tutkielman tulokset ovat hyvin riippuvaisia sekä metodologian ja aineiston rajoitteista, että menestyksen ja opiskelijavalinnan tavoitteiden määrittelystä.

---

**Avainsanat** opiskelijavalinta, opintomenestys, ylioppilastutkinto, valintakoe, ennustaminen

---

---

**Author** Emma Mieskonen

---

**Title of thesis** Predicting study success and evaluating student selection impact at the Aalto University School of Business

---

**Degree** Master of Science in Economics and Business Administration

---

**Degree programme** Economics

---

**Thesis advisor** Matti Sarvimäki

---

**Year of approval** 2017**Number of pages** 78**Language** Finnish

---

**Abstract**

The Finnish university student selection system has been criticized for its inefficiency for decades: the current system is costly for the applicant, the selecting university and the society. As a solution, the Finnish Ministry of Education and Culture has proposed a reform to replace the current admission examination heavy selections with a simpler selection based on the national matriculation examination by year 2020. Although a lot of research has been done on student selection and study success forecasting on a general basis, few analyses have been made on how student selection criteria can affect the average study success of a university's students.

The aim of this thesis is to provide a limited empirical method for analyzing the effects of alternative student selection criteria from the university's point of view. The practical goal is to find out how the average study success of the school's student population might change, if the admission examination was relinquished. The thesis answers the following questions: 1) How do matriculation examination grades predict study success? 2) What should the matriculation examination scoring be like based on the prediction results? 3) How would the chosen student population change, if the student selection was done purely based on the matriculation examination? 4) How would the average study success of the student population change, if the student selection was done purely based on the matriculation examination? The study is based on the student and course data of the entry cohorts 2005 – 2012 of the Aalto University School of Business.

The student selection problem can be set as choosing the best potential students with the least possible costs. Study success is defined in the thesis as grade point averages (GPA) and study credit accumulations as per two study success indices used by the Business School itself. Applicant ability is measured by their matriculation examination grades and admission examination points. The results of the thesis are in line with previous literature in finding that the matriculation examination is a good predictor of university GPA in a basic OLS regression model. Especially mathematical skills seem to indicate better study success. The models are not successful in predicting study credit accumulations. The prediction significance of the admission examination is found to be minor. Further, it seems that the admission examination has not brought added value in the student selection: a very similar, slightly more successful student population would have been chosen in a matriculation examination based student selection. The results of the study are highly dependent on the limitations of the methodology and given data, as well as the definitions of study success and student selection objectives.

---

**Keywords** student selection, study success, matriculation exam, admission exam, forecasting

---

# Sisällysluettelo

<b>1 JOHDANTO</b>	<b>1</b>
1.1 TUTKIMUSKYSYMYKSET JA TAVOITTEET	4
<b>2 OPISKELIJAVALINNAN TAVOITTEET JA TUTKIMINEN</b>	<b>7</b>
2.1 OPINTOMENESTYKSEN JA PARHAAN POTENTIAALISEN OPISKELIJAN MÄÄRITTELY	7
2.2 PARHAAN POTENTIAALISEN OPISKELIJAN VALITSEMINEN	11
2.3 OPISKELIJAVALINNAN JA OPINTOMENESTYKSEN EMPIIRISEN TUTKIMUKSEN ESITTELY	14
<b>3 AINEISTON ESITTELY</b>	<b>24</b>
3.1 DATA	24
3.2 OPINTOMENESTYKSEN MITTAAMINEN: SELITETTÄVÄT MUUTTUJAT	26
3.2.1 KOLMEN VUODEN OPINTOPISTEKERTYMÄ	28
3.2.2 KOLMEN VUODEN ARVOSANAKESKIAARVO	29
3.2.3 KAIKILLE YHTEISTEN OPINTOJEN ARVOSANAKESKIAARVO	29
3.3 VALINTAKRITEEREISSÄ MENESTYMINEN: SELITTÄVÄT MUUTTUJAT	30
3.3.1 YLIOPPILASTUTKINTOARVOSANAT	31
3.3.2 VALINTAKOEPISTEET	34
3.4 KONTROLLIMUUTTUJAT	36
<b>4 OPINTOMENESTYKSEN ENNUSTAMINEN</b>	<b>38</b>
4.1 METODOLOGIA	38
4.1.1 OLS-OLETUKSET	40
4.1.2 YLISOVITTAMINEN	41
4.2 ENNUSTETULOKSET	43
4.2.1 KOLMEN VUODEN OPINTOPISTEKERTYMÄ	44
4.2.2 KOLMEN VUODEN ARVOSANAKESKIAARVO	48
4.2.3 KAIKILLE YHTEISTEN OPINTOJEN ARVOSANAKESKIAARVO	51
4.3 ANALYYSI	55
4.3.1 VAIHTOEHTOINEN YLIOPPILASTUTKINTOARVOSANOJEN PISTEYTYYS	56

<b>5 OPISKELIJAVALINNAN VAIKUTUKSEN ARVIOINTI</b>	<b>59</b>
<b>5.1 METODOLOGIA</b>	<b>59</b>
<b>5.2 ARVIOINTITULOKSET</b>	<b>61</b>
5.2.1 VALINTAMENETTELYN EROTTELEVUUS	62
5.2.2 VALINTAMENETTELYN YHTEYS KESKIMÄÄRÄISEEN OPINTOMENESTYKSEEN	63
<b>6 KESKUSTELU JA RAJOITTEET</b>	<b>70</b>
<b>7 JOHTOPÄÄTÖKSET</b>	<b>77</b>
<b>LÄHTEET</b>	<b>VI</b>
<b>LIITTEET</b>	<b>XI</b>
LIITE 1: KAUPPATIETEELLISEN ALAN YHTEISVALINTA	XI
LIITE 2: JAKAUMAKUVIOT	XIV
LIITE 3: PERSENTILIN LASKEMINEN	XV
LIITE 4: HAJONTAKUVIOT	XVI
LIITE 5: YLIOPILASTUTKINTOARVOSANOJEN YHTEYS ARVOSANAKESKIAARVOON	XVII
LIITE 6: KEINOTEKOISEN OPISKELIJAVALINNAN VALINTAMENETTELYT	XVIII

## **Taulukot**

Taulukko 1. Valintakriteerien ennustemerkityksellisyys

Taulukko 2. OLS regressiotulokset: Kolmen vuoden opintopistekertymät ja arvosanakeskiarvot (Mukailtu: Pietilä 2012)

Taulukko 3. Ylioppilastutkintoarvosanojen merkityksellisyys (Mieskonen 2016)

Taulukko 4. Muuttujien yhteenveto

Taulukko 5. Selitettävien muuttujien parittainen korrelaatio

Taulukko 6. Ylioppilastutkinnon arvostelu (Ylioppilastutkintolautakunta 2014)

Taulukko 7. Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet 3 vuoden opintopistekertymän OLS regressiomallissa (1)

Taulukko 8. OLS regressiotulokset: 3 vuoden opintopistekertymä

Taulukko 9. Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet 3 vuoden arvosanakeskiarvon OLS regressiomallissa (1)

Taulukko 10. OLS regressiotulokset: 3 vuoden arvosanakeskiarvo

Taulukko 11. Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet kaikille yhteisten opintojen arvosanakeskiarvon OLS regressiomallissa (1)

Taulukko 12. OLS regressiotulokset: Kaikille yhteisten opintojen arvosanakeskiarvo

Taulukko 13. Ylioppilastutkintoarvosanojen vaihtoehtoinen pisteytys

Taulukko 14. Valintamenettelyjen marginaalisten opiskelijoiden yhteenveto

Taulukko 15. Muiden kuin marginaalisten opiskelijoiden yhteenveto

## **Kuviot**

Kuvio 1. Ylioppilastutkintoarvosanojen jakauma

Kuvio 2. Valintamenettelyn erottelevuus

Kuvio 3. Ylioppilastutkinnon painotuksen merkitys keskimääräiselle opintomenestykselle

# 1 Johdanto

Suomen korkeakoulujärjestelmän opiskelijavalintaa on kritisoitu tehottomuudesta vuosikausia. Opetus- ja kulttuuriministeriön tuore selvitys (2016b) kertoo karuja lukuja opiskelijavalinnan nykytilasta: vuoden 2015 uusista ylioppilaista 75 % haki korkeakouluun valmistumisvuonaan, mutta vain noin 16 % pääsi välittömästi jatkamaan opintojaan yliopistossa, 32 % jossain tutkintotavoitteisessa jatkokoulutuksessa. Jatko-opintojen ulkopuolelle jäi 68 % tuoreista ylioppilaista, kun kymmenen vuotta sitten vastaava luku oli 58 %. Suurin osa tarjolla olevista opiskelupaikoista menee nykyjärjestelmässä siis toista tai useampaa kertaa pyrkiville hakijoille ja tehoton opiskelija-allokaatio näkyy väli vuosien ja tutkintoaan vaihtavien opiskelijoiden suuressa määrässä. Kun keskimääräisen vastavalmistuneen ylioppilaan ikä on 19 vuotta, aloitetaan korkeakouluopinnot keskimäärin 24 vuoden iässä: OECD maiden vertailututkimukset osoittavat, että Suomessa sekä aloitetaan että päätetään opinnot keskimäärin huomattavasti eurooppalaisia vertailukohteita vanhempina (Opetus- ja kulttuuriministeriö 2016b). Opiskelijavalinnan ongelmat aiheuttavatkin ylimääräisiä kustannuksia niin hakijoille kuin valintaa tekeville korkeakouluille ja yhteiskunnalle. Ongelmista keskusteltaessa nousee poikkeuksetta esiin etenkin erilaisten koulu- tai alakohtaisten valintakokeiden suuri määrä: keväällä 2016 erillisiä valintakokeita järjestettiin yli tuhat kappaletta (Opetus- ja kulttuuriministeriö 2016b). Vaikka usealla alalle hakeminen on teoriassa mahdollista, pitkää valmistautumista ja huomattavia ajallisia resursseja vaativat valintakokeet pakottavat hakijan taktikoimaan hakukohtien priorisoinnissa ja keskittymään käytännössä vain yhteen hakuun täysillä (Pekkarinen & Sarvimäki 2016).

Myös opiskelijavalinnan ongelmien korjaamisesta on Suomessa puhuttu jo vuosikymmeniä<sup>1</sup>. Käytännössä ratkaisuksi opiskelijavalinnan tehostamiseen on usein tarjottu ylioppilastutkinnon painottamista valinnoissa: esimerkiksi Opetusministeriön selvitysmies Ahola (2004) esittää, että korkeakoulutuksen opiskelijavalinnan tulisi perustua

---

<sup>1</sup> Esimerkiksi selvitysmies Jussila ehdottaa jo vuoden 1996 raportissaan Opetusministeriölle opiskelijavalinnan muuttamista niin, että 50 – 70 % valittavista uusista korkeakouluopiskelijoista olisi tuoreita ylioppilaita.

pääosin lukio- ja ylioppilaskoemenestykseen. Ylioppilastutkinto on jo nyt yleisesti käytetty opiskelijavalinnan peruste, jonka arvosanojen ajatellaan olevan vertailukelpoisia läpi Suomen ja eri suoritusvuosien. Kevään 2016 valintakierroksella 89 % yliopistoon hakeneista ja 95 % paikan vastaanottaneista oli suorittanut suomalaisen tai kansainvälisen ylioppilastutkinnon. Ylioppilastutkintoon perustuvan, ensikertalaisia hakijoita suosivan opiskelijavalinnan uskotaan lyhentävän koulutukseen hakeutumiseen kuluvaan aikaa, ja lopulta jouduttavan nuorten siirtymistä työelämään. (Opetus- ja kulttuuriministeriö 2016b). Ikääntyvässä Suomessa työurien pidentäminen alkupäästä on ollut poliittisena tavoitteena pitkään. Pääministeri Juha Sipilän hallitusohjelma (Valtioneuvoston kanslia 2015a & 2015b) nimeääkin yhdeksi kuluvan hallituskauden kärkihankkeeksi korkeakoulujen valintamenettelyjen uudistamisen: tehokkaan opiskelijavalinnan tulisi mahdollistaa entistä joustavimmat opintopolut ja sujuvan siirtymän toisen asteen koulutuksesta yliopiston kautta työelämään. Uudistetun opiskelijavalinnan tavoitteena on vähentää hakuprosessien vuoksi vietettyjä väli vuosia ja aikaistaa ylemmän asteen opintojen aloitusta. Samankaltaisia tavoitteita on nimetty jo aiempien hallitusten tavoiteohjelmissa 2010-luvulla.

Opetus- ja kulttuuriministeri Sanni Grahn-Laasosen voimakkaasti ajama kärkihanke on edennyt hallituskauden alkuvuosien 2015 – 2017 aikana nopeasti strategisesta suunnitelmasta konkreettisiin tavoitteisiin ja sopimuksiin. Opetus- ja kulttuuriministeriön, korkeakoulujen ja ylioppilastutkintolautakunnan edustajista muodostettu työryhmä laati vuonna 2016 esityksen ”*Valmiina valintoihin*” toimenpideohjelmaksi ylioppilastutkintojen hyödyntämisestä korkeakoulujen opiskelijavalinnoissa (Opetus- ja kulttuuriministeriö 2016b). Työryhmä esittää hakevan yksilön, yhteiskunnan ja korkeakoulujen resurssien käytön optimointia erillisistä valintakokeista luopumalla ja toisen asteen koulutusnäyttöjä korostamalla. Periaatteena on, että samalla osaamisen näytöllä tulisi voida hakea useampaan koulutuskohteeseen, ilman tehottomuutta aiheuttavaa taktikointia valintakokeisiin osallistumisen ja valituksi tulemisen todennäköisyyden välillä. Lisäksi opiskelijavalinnan tulisi olla tasa-arvoinen, ennustettava ja muuttumaton, jotta nykyiset ja tulevat toisen asteen opiskelijat voivat alusta asti suunnitella opintonsa korkeakoulujen valintakriteerejä silmällä pitäen, ilman yllätyksiä tai muuttuvaa valintakoemateriaalia. Myöhään herääville hakijoille tulee kuitenkin edelleen tarjota mahdollisuus hakeutua koulutukseen vaihtoehtoisilla näytöillä.



Loppuvuonna 2016 valmistunut toimenpideohjelman esitys on tätä kirjoittaessa lausuntokierroksella korkeakouluissa ja suunnitelmapiirissä on tarkoitus viedä päätökseen joulukuussa 2017. Ensimmäiset askeleet täytäntöönpanoon on kuitenkin jo otettu. Opetus- ja kulttuuriministeriön ja kunkin korkeakoulun välisissä tulossopimuksissa vuosille 2017 – 2020 on sovittu, että pitkää valmentautumista vaativista opiskelijavalinnoista luovutaan vuoteen 2018 mennessä. Opiskelijavalintojen tulee siis perustua suoraan ylioppilastutkintoon tai toisen asteen koulutuksen aikana hankittuun osaamiseen uutta sisältöä painottavien valintakokeiden sijaan. Vuoteen 2020 mennessä pääosa uusista opiskelijoista tulee valita toisen asteen tutkintotodistusten perusteella, valtaosin suoraan ylioppilastutkinnon arvosanoihin perustuen. Tavoitteiden toteutumista ohjataan korkeakoulujen strategiarahoituksella (Opetus- ja kulttuuriministeriö 2016a). Käytännössä sopimus vaatii opiskelijavalinnan radikaalia uudistamista nopealla aikataululla, ja yliopistoja onkin pyydetty toimittamaan esityksensä ylioppilastutkintoarvosanojen uudesta lähtöpisteytyksestä vuoden 2017 joulukuuhun mennessä. Lisäksi Opetus- ja kulttuuriministeriö peräänkuuluttaa opiskelijavalintojen kehittämistä tukevan tutkimuksen tarvetta muun muassa valintakokeiden merkityksellisyyden arvioinnissa. Tämän tutkielman tavoitteena on osaltaan vastata näihin tarpeisiin.

Vaikka opiskelijavalintaa ja opintomenestyksen ennustamista on Suomessa tutkittu runsaasti, ei tutkimusta valintamenettelyjen vaikutuksesta valittavaan opiskelijajoukkoon tai sen opintomenestykseen ole juuri tehty. Yksi opiskelijavalinnan arvioinnin tärkeimmistä kysymyksistä kuitenkin on, miten vaihtoehtoiset valintamenettelyt toimisivat ja millainen vaikutus niillä olisi korkeakoulujen hakijoihin tai valituksi tuleviin opiskelijoihin ja näiden menestymiseen suhteessa nykyisiin valintamenettelyihin. Tämän tutkielman tavoitteena on esitellä yksinkertainen empiirinen menetelmä, jolla vaihtoehtoisia valintamenettelyitä ja muun muassa valintakokeen kannattavuutta opiskelijavalinnassa on mahdollista arvioida opiskelijavalinnan järjestävän korkeakoulun näkökulmasta. Käytännöllisenä tavoitteena on alustaa keskustelua hallituksen peräänkuuluttaman opiskelijavalinnan uudistamisen mahdollisista käytännön vaikutuksista kauppatieteellisellä alalla. Menetelmä on rajallinen ja puutteellinen, mutta pyrkii tarjoamaan perustan monipuolisempien arviointimenetelmien lähtökohdaksi.

Kauppatieteellinen ala houkuttelee vuodesta toiseen suuria hakijamääriä. Vuoden 2017 valintakierroksella kauppatieteitä pyrki opiskelemaan noin 10 000 hakijaa, joista noin 15 % tarjottiin opiskelupaikkaa. Aalto-yliopiston kauppakorkeakoulu on Suomen suurin ja valikoivin kauppatieteellinen koulutusyksikkö, johon hyväksytään vuosittain noin 8 % hakijoista. Vuonna 2017 yhtä aloituspaikkaa kohden Aaltoon pyrki 12 hakijaa. Kilpailu paikoista on siis kovaa, mikä korkeakoulujen puolelta tarkoittaa sitä, että opiskelijavalinnalla on suuri karsiva vaikutus: laajasta hakijajoukosta on varaa valita sellaiset opiskelijat kuin halutaan. Tämä tekee Aalto-yliopistosta mielenkiintoisen kohteen opintomenestyksen ennustamisen ja opiskelijavalinnan arvioinnin tutkimista ajatellen. (Kauppatieteellisen alan yhteisvalinta 2017)

## 1.1 Tutkimuskysymykset ja tavoitteet

Tutkielma pyrkii vastaamaan seuraaviin kysymyksiin:

- 1) Miten ylioppilastutkintoarvosanat ja valintakoe ennustavat opintomenestystä?
- 2) Millainen ylioppilastutkintoarvosanojen valintapisteytysmallin tulisi niiden ennustearvon mukaan olla?
- 3) Kuinka eri ihmiset tulisivat valituksi, jos käytössä olisi pelkästään ylioppilastutkintoon perustuva valinta?
- 4) Miten ylioppilastutkinnon perusteella valitut opiskelijat menestyisivät verrattuna niihin opiskelijoihin, jotka eivät olisi tulleet valituksi pelkkään ylioppilastutkintoon perustuvassa valinnassa?

Tutkielma koostuu kolmesta vaiheesta. Ensimmäinen vaihe esittelee opiskelijavalinnan teorian ja aiheen parissa tehdyn aiemman tutkimuksen (luku 2) ja seuraavissa vaiheissa käytetyn datan ja muuttujat (luku 3). Opiskelijavalinnan ongelma tiivistetään tutkielmassa pyrkimykseen valita parhaat mahdolliset potentiaaliset opiskelijat mahdollisimman pienin resurssi-investoinnein. Resurssitehokkaana pidetyn ylioppilastutkinnon lisäksi erillisiä valintakriteerejä kannattaa soveltaa ainoastaan, mikäli niiden tuoma lisäarvo opiskelijavalintaan on yhtä suuri tai suurempi kuin niiden kustannus. Paras potentiaalinen opiskelija ja tämän opintomenestys määritellään tutkielmassa Aalto-yliopiston kauppakorkeakoulun itsensä käyttämiin opintomenestysindekseihin perustuen

arvosanakeskiarvoilla ja opintopistekertymillä. Korkeakouluopintoihin potentiaalisimmin soveltuvimmat hakijat pyritään tunnistamaan opiskelijavalinnan käytössä olevien valintakriteerien – ylioppilastutkintoarvosanojen ja valintakoepisteiden – perusteella. Aiheen parissa aiemmin tehty tutkimus tukee tutkielman määrittelyjä ja tarjoaa työkaluja sekä opintomenestyksen ennustamiseen, että opiskelijavalinnan arviointiin. Aalto-yliopiston kauppakorkeakoulussa opintomenestykseen vaikuttavia tekijöitä on tutkittu runsaasti, mutta valintamenettelyjen vaikutuksesta opiskelijoiden keskimääräiseen opintomenestykseen ei ole ennakkotuloksia. Olemassa olevan kirjallisuuden perusteella vaikuttaa kuitenkin siltä, että arvosanakeskiarvojen ennustaminen ylioppilastutkintoarvosanoilla on varsin mielekästä. Opintopistekertymien selittäminen ja valintakoepisteiden käyttäminen ennustamiseen vaikuttaa sen sijaan haastavammalta.

Toisen vaiheen tarkoituksena on deskriptiivisesti tutkia ylioppilastutkintoarvosanojen ja valintakokeen yhteyttä opiskelijan opintomenestykseen Aalto-yliopiston kauppakorkeakoulussa (luku 4). Pyrkimyksenä on tunnistaa opintomenestyksen kannalta merkityksellisimmät ylioppilaskoeaineet ja näiden arvosanat, ja muodostaa näiden perusteella käsitys menestyneimpien opiskelijoiden aine-arvosana kombinaatioista. Käytännöllisenä tavoitteena on määritellä ylioppilastutkinnon arvosanoille empiriaan perustuva pisteytys, joka tukisi pyrkimystä valita Kauppakorkeakouluun ne hakijat, joiden potentiaali menestyä opinnoissaan on suurin. Aineistona käytetään Aalto-yliopiston kauppakorkeakoulun opiskelija- ja kurssirekisterimerkintöjä vuosina 2005 – 2012 opintonsa aloittaneiden opiskelijoiden osalta. Tutkielman tulokset ovat linjassa aiemman kirjallisuuden kanssa: vaikuttaa siltä, että ylioppilastutkintoarvosanat ennustavat korkeakouluopintojen kurssiarvosanojen muodostumista varsin mielekkäästi, mutta opintojen etenemisen ennustamiseen ei niistä ole. Erityisesti matemaattinen osaaminen vaikuttaa korostuvan kauppatieteellisissä opinnoissa menestymisessä. Valintakoe sen sijaan ei tuo opintomenestyksen ennustamiseen huomattavaa lisäarvoa.

Kolmannen vaiheen tarkoituksena on tutkia, kuinka eri ihmiset tulisivat valituksi ja miten opiskelijoiden keskimääräinen opintomenestys muuttuisi, jos Aalto-yliopiston kauppakorkeakoulun opiskelijavalinta perustuisi puhtaasti ylioppilastutkintoon (luku 5). Käytännöllisenä tavoitteena on selvittää, miten valintakokeesta luopuminen tai sen painoarvon pienentäminen vaikuttaisi opiskelijavalinnan tuloksiin. Tutkielmassa havaitaan, että noin 70 % Aalto-yliopiston kauppakorkeakoulun nykyisistä opiskelijoista olisi tullut

valituksi valintamenettelystä riippumatta. Loput 30 % uusista opiskelijoista valitaan eri tavalla valintamenettelystä riippuen. Näiden marginaalisten opiskelijoiden tutkiminen on tutkielman tavoitteiden kannalta kiinnostavaa. Analyysissa havaitaan, että ryhmät ovat opintomenestykseltään hyvin samankaltaisia, kuitenkin niin, että puhtaaseen todistusvalintaan perustuen valitut opiskelijat ovat hieman vertailuryhmäänsä menestyneempiä. Tulosten perusteella valintakoe ei ole tuottanut lisäarvoa tutkielman keinotekoisessa opiskelijavalinnassa: hyvin samankaltainen, jopa hieman paremmin menestynyt opiskelijajoukko olisi saatu valittua ilman valintakokeen tuottamaa informaatiota. Todellisessa opiskelijavalinnassa viimeisten, marginaalisten hakijoiden välisten erojen voidaan odottaa olevan suurempia kuin tutkielman tuloksista voi suoraan päätellä. Erillisen valintakokeen järjestäminen ei siis tässä rajatussa valinnassa ole tulosten valossa ollut kannattavaa Aalto-yliopiston kauppakorkeakoulussa. On kuitenkin huomattava, että tutkielman tulokset ovat hyvin riippuvaisia sekä käytetyn mallin ja aineiston rajoitteista, että opintomenestyksen ja opiskelijavalinnan tavoitteiden määrittelystä. Tuloksia ei tulisi sellaisenaan yleistää tai soveltaa opiskelijavalinnan suunnitteluun ilman rajoitteiden syvällistä ymmärtämistä. Tutkielman tuloksista ja rajoitteista käydään tarkempaa keskustelua luvussa 6; luku 7 tiivistää tutkimuksen tulokset johtopäätöksiin.

## 2 Opiskelijavalinnan tavoitteet ja tutkiminen

Opiskelijavalinnan onnistumista on arvioitava suhteessa sille asetetuille odotuksille, ja siksi sitä tutkiessa on kriittistä ymmärtää mitä opiskelijavalinnalla tavoitellaan. Seuraavat alaluvut tarjoavat pohjan tutkielman empiiriselle osalle määrittelemällä, mitä tutkielmassa tarkoitetaan opintomenestyksestä, parhaasta potentiaalisesta opiskelijasta tai opiskelijavalinnan resurssitehokkaasta järjestämisestä puhuttaessa. Viimeinen alaluku tarjoaa katsauksen siihen, miten ja millaisin tuloksin opiskelijavalintaa on aiemmassa kirjallisuudessa tutkittu.

### 2.1 Opintomenestyksen ja parhaan potentiaalisen opiskelijan määrittely

Onnistuneen opiskelijavalinnan suunnittelun kannalta on ensiarvoisen tärkeää ymmärtää, kenen koetaan olevan menestynyt, miten paras potentiaalinen opiskelija määritellään ja miten opintomenestystä mitataan. Ahola (2004) toteaaakin, että opintomenestyksen ymmärtämisen ongelmana on määritellä mitä arvioidaan: sopivuutta koulutukseen ja mahdolliseen ammattiin, koulutettavuutta yleensä vai menestystä eri tavoin mitattavissa suorituksissa<sup>2</sup>. Useimmiten opintomenestystä ei määritelläkään yhden mitattavan muuttujan kautta, vaan erilaisia kriteerejä yhdistellen. Yliopistojen opiskelijavalinnan arvioinnin ja opintomenestyksen tutkimisen kannalta onkin ratkaisevaa, miten menestyskriteerit määritellään, valitaan, havaitaan ja mitataan. Sakari Aholan ja Anu Kokon (2001) mukaan paras potentiaalinen opiskelija kiteytyy neljään ominaisuuteen: menestynyt opiskelija opiskelee tehokkaasti, ei keskeytä, valmistuu ajallaan ja työllistyy helposti.

---

<sup>2</sup> *Opiskelijavalintaa laajemmin arvioitaessa olisi tärkeää ymmärtää, mikä vaikutus valinnoilla on opintojen jälkeiseen menestymiseen: miten esimerkiksi hakijoiden tuleva työllisyystilanne, työrooli, asema tai ansiotaso olisivat ennustettavissa valintahetkellä. Tutkielman aineisto ei kuitenkaan sovellu yliopiston jälkeisen menestymisen arviointiin, sillä aineisto ei sellaisenaan ole yhdistettävissä valmistuneiden opiskelijoiden työtilanne- tai tulotietoihin. Tutkielma rajataankin nimenomaan opintojen aikaisen opintomenestyksen analyysiin. Ahola (2004) argumentoi kuitenkin, että yliopistoaikainen opintomenestys on tutkimusten mukaan yhteydessä menestymiseen myöhemmin työelämässä.*

Opiskelijavalintaa korkeakoulun näkökulmasta arvioitaessa opintomenestys voidaan määritellä valitsevan korkeakoulun omien tavoitteiden kautta: korkeakoulun kannattaa valita sellaiset opiskelijat, jotka parhaiten auttavat sitä saavuttamaan sen omia pyrkimyksiä. Esimerkiksi kovatasoisen tutkimusyliopiston voisi olla kannattavaa valita opiskelijoita, joita kiinnostaa jo opintojen alussa osallistua tutkimusryhmiin, tai jotka vaihtoehtoisesti vaativat mahdollisimman vähän opetusta, jolloin professorien aikaa vapautuisi arvokkaaseen tutkimustyöhön. Ainoastaan tutkintomääriä mittaavan yliopiston taas kannattaisi valita mahdollisimman nopeasti valmistuvia opiskelijoita välittämättä näiden laadusta tai muusta kontribuutiosta. (Pietilä 2012)

Aalto-yliopiston kauppakorkeakoululla on itsellään käytössä muutamia opintomenestysindekseiksi kutsuttuja laskentakaavoja, joiden perusteella on mahdollista tehdä päätelmiä Kauppakorkeakoulun opiskelijoilleen asettamista tavoitteista. Erikoistumisohjelman valinnassa opiskelijat arvotetaan seuraavalla kaavalla (Aalto-yliopiston kauppakorkeakoulu 2017a):

(1)

$$P_i = \sum_j^n kop_j * x_j$$

Tässä opiskelijan  $i$  opintomenestysindeksi  $P_i$  on laskettu  $n$  kappaleen liiketoimintaosaamisen perusteiden kurssin  $j$  opintopistemäärän  $kop_j$  ja arvosanan  $x_j$  tulojen summana. Kansainvälisen vaihto-opintokokonaisuuden paikkojen allokoinnissa tärkein valintaperuste on opintojen laatua ja etenemistä painottava opintomenestysindeksi (Aalto-yliopiston kauppakorkeakoulu 2017b):

(2)

$$P_i = \frac{op_i * GPA_i}{a}$$

Tässä  $P_i$  on koulun laskema opintomenestysindeksi opiskelijalle  $i$ ,  $op_i$  tämän suorittama opintopistemäärä,  $GPA_i$  opintopisteillä painotettu keskiarvo ja  $a$  läsnäololukukausien määrä. Kauppakorkeakoulun käyttämät opintomenestysindeksit painottavat käytännössä kahta asiaa: saatuja arvosanoja ja korkeaa opintopistekertymää suhteessa lyhyeen läsnäoloaikaan. Näitä mittareita tullaan hyödyntämään tutkielmassa opintomenestyksen arvioinnissa.

Opiskelijoiden saamia arvosanoja ja niistä johdettua arvosanakeskiarvoa käytetään yleisesti kyvykkyyden mittarina empiirisessä tutkimuksessa (mm. Ahola 2004, Rothstein 2004, Pietilä 2012). Arvosanat ovat mittarina helppokäyttöisiä: ne ovat suoraan vertailukelpoisia, yhtenäisin kriteerein myönnettyjä kvantitatiivisia arviointeja opiskelijan opintomenestyksestä. Suomen peruskouluissa, lukioissa, ylioppilastutkinnossa ja korkeakouluissa arvosanoja annetaan yhteisesti sovitun skaalan mukaisesti arviona kerrytetystä henkisestä pääomasta. Arvosanoja käytetään opiskelijan potentiaalisen menestystason arviointiin myös yliopiston ulkopuolella: työelämään siirryttäessä on kauppatieteellisillä aloilla verrattain tavallista, että ansioluettelon ja työhakemuksen lisäksi pyydetään hakemuksen yhteydessä myös yliopisto- ja lukioaikaiset opintosuoritustodistukset. Aholan (2004) ja Pietilän (2012) mukaan useat tutkimukset ovat myös osoittaneet positiivisen korrelaation yliopisto-opintojen arvosanakeskiarvon ja myöhemmän palkkatason välillä.

Opintopistekertymiä käytetään empiirisessä tutkimuksessa opintojen etenemisen mittarina (mm. Häkkinen 2004, Pietilä 2012, Kupiainen & Marjanen 2014<sup>3</sup>). Opintojen etenemistahdin on esitetty olevan muun muassa opiskelumotivaation tai alalle soveltumisen mittari: toiselle alalle haikailevat, toisen opiskelupaikan vastaanottaneet tai muuten alaan tympääntyneet opiskelijat etenevät opinnoissaan innostuneita nuoria hitaammin ja keskeyttävät opintonsa muita useammin (Saarenmaa, Saari & Virtanen 2010). Nopeasti

---

<sup>3</sup> Kupiainen ja Marjasen vuoden 2014 tutkimuspaperi ylioppilastutkinnon hyödyntämisestä opiskelijavalinnassa on toistaiseksi palautettu valmisteluun, ja julkaistaneen laajennettuna kirjana loppuvuodesta 2017. Tutkimuksen tulokset eivät ole siis käyneet vielä läpi lopullista vertaisarviointia, ja niihin on syytä suhtautua pienellä varauksella. Tutkimuksen tuloksista on käytössä myös luentomateriaaleja (Kupiainen 2014).

valmistuvat opiskelijat sen sijaan ovat muita useammin kiinnostuneita opiskelualastaan sekä motivoituneita ja kyvykkäitä etenemään opinnoissaan suunnitelmallisesti tutkintoa ja työelämää kohti. On kuitenkin huomattava, että Saarenmaan et al. (2010) mukaan heikko opiskelumotivaatio tai kokemus väärällä alalla opiskelemisesta (34 %) on vasta kolmanneksi yleisimmin raportoitu hitaan etenemisen syy työssäkäynnin (52 %) ja elämäntilanteen tai henkilökohtaisten syiden (43 %) jälkeen. Muita yleisiä opintoja hidastavia tekijöitä ovat raportin mukaan muun muassa opintojen työläys, puutteellinen opintojen ohjaus, toimeentulovaikeudet, puutteelliset opiskelutaidot sekä aktiivisuus opiskelijajärjestössä tai muussa harrastustoiminnassa. Opintojen eteneminen ei siis ole yksiselitteinen opintomenestyksen mittari.

Opinnoissa eteneminen ja niiden loppuunsaattaminen ovat joka tapauksessa Suomen hallituksen ja yliopistojen erityisen mielenkiinnon kohteina: nopeammat tutkinnot mahdollistavat hallituksen tahtotilan mukaisesti nopeamman työelämään siirtymisen, ja siksi huomattava osa yliopistojen julkisesta rahoituksesta on sidottu lyhyempien tutkintoaikojen edistämiseen (Opetus- ja kulttuuriministeriö 2016a). Käytännön mittareina yliopistojen rahoitusmallissa hyödynnetään vuosittaisia opintopistekertymiä (10 %) ja suoritettujen tutkintojen määrää (20 %). Opintopistekertymien ohella soveltuvia opinnoissa etenemisen mittareita ovatkin esimerkiksi keskeyttämis- ja valmistumistodennäköisyys (Öckert 2001, Häkkinen 2004) sekä valmistumisaika (Rothstein 2004). Käytössä olevan aineiston rajallisuuden vuoksi tutkielmassa opintojen etenemisen mittarina käytetään kuitenkin ensisijaisesti ajallisesti rajoitettuja opintopistekertymiä.

Opiskelijan potentiaalisen menestystason tai kyvykkyyden määrittelyyn voi siis käyttää monipuolista mittaristoa optimoitavien tavoitteiden näkökulmasta riippuen. Aalto-yliopiston kauppakorkeakoulun omien opintomenestysindeksien perusteella voidaan korkeakoulun opintomenestyksen mittareiksi tutkielmassa määritellä opiskelijoiden arvosanakeskiarvot ja opintopistekertymät. Yliopiston päämäärän voidaan ajatella olevan mahdollisimman pienin resurssi-investoinnein valita ne hakijat, jotka todennäköisimmän menestyvät näillä kriteereillä mitattuna eli ovat hakijajoukon parhaita potentiaalisia opiskelijoita.



## 2.2 Parhaan potentiaalisen opiskelijan valitseminen

Opiskelijavalintaa voidaan teoreettisesti tarkastella kolmesta näkökulmasta: yhteiskunnan, korkeakoulun tai hakijan (Opetus- ja kulttuuriministeriö 2016b: [Talman 2014]). Yhteiskunnan optimointiongelmana on saada opiskelijat sulavasti ja kustannustehokkaasti toisen asteen koulutuksesta korkeakouluopintoihin ja edelleen työelämään. Korkeakoulun intressinä on tunnistaa hakijajoukosta aiemmassa luvussa määritellyt parhaimmat potentiaaliset opiskelijat mahdollisimman pienin resurssi-investoinnein. Hakijan kannalta valinnan on tärkeää olla sujuva, tehokas ja oikeudenmukainen. Pekkarinen ja Sarvimäki (2016) tiivistävät valintamenettelyn tavoitteet neljään päämäärään: 1) hakijoiden ja korkeakoulujen toiveet toteutuvat mahdollisimman hyvin, 2) valintajärjestelmä on reilu, 3) hakijoille annetaan riittävästi uusia mahdollisuuksia ja 4) resursseja hukataan mahdollisimman vähän. Tutkielman analyysin kannalta tärkeintä on ymmärtää valitsevan korkeakoulun näkökulma opiskelijavalintaan.

Korkeakoulun näkökulmasta opiskelijavalintaa ohjaavat niin rajalliset resurssit kuin kiinnostus valita hakijoiden joukosta opintoihin soveltuvimmat opiskelijat. Opiskelijavalinnan ongelman voi tiivistää pyrkimykseen valita hakijoiden joukosta parhaimmat potentiaaliset opiskelijat mahdollisimman vähäisillä investoinneilla itse valintaprosessiin. Hakijan potentiaalista menestystasoa voi arvioida erilaisina kyvykkyyksinä. Kyvykkyyden voi määritellä kontekstista riippuen usein eri tavoin, esimerkiksi älykkyytenä, osaamistasona, taitavuutena, tehokkuutena tai vaikka motivaationa. Sakari Ahola ja Anu Kokko (2001) määrittelevät kauppatieteellisen alan määrääviksi kyvykkyyksiksi motivaation, sopivat taidot ja tiedot, lahjakkuuden ja alalle soveltuvat henkilökohtaiset ominaisuudet, jotka Parikka [1999] lainaten voi tiivistää hyviin viestintä- ja sosiaalisiin taitoihin, valmiuteen ottaa vastuuta sekä kykyyn oppia jatkuvasti ja soveltaa teoreettista tietoa. Potentiaalisesti kyvykkäimpien hakijoiden valitseminen on korkeakoululle tärkeää, jotta opiskelijoille asetetut tavoitteet saavutetaan mahdollisimman tehokkaasti.

Empiirisesti on kuitenkin oletettava, että opiskelijavalinnan vastuhenkilöt eivät ole kykeneviä suoraan havainnoimaan hakijoiden kyvykkyyttä tai potentiaalia menestyä opinnoissaan. Sen sijaan koulutukseen soveltuvimmat hakijat on tunnistettava käyttäen havaittavissa tai luotavissa olevia kvantitatiivisia signaaleja opiskelijoiden

menestyspotentiaalista. Opetus- ja kulttuuriministeriön tekemän selvityksen mukaan korkeakoulut arvioivat valintaperusteillaan muun muassa hakijoiden opiskelukykyä ja tiedon omaksumistaitoja, lukion oppimäärään kuuluvia ja alakohtaisia tietoja ja taitoja, kiinnostusta ja motivaatiota alan opiskeluun sekä soveltuvuutta alan töihin (Opetus- ja kulttuuriministeriö 2016b). Käytännössä Kauppakorkeakoulun nykyvalinnassa (liite 1) hyödynnettäviä signaaleja on kaksi: ylioppilastutkintoarvosanat ja valintakoepisteet.

Ylioppilastutkinto on suomalaisen lukiojärjestelmän päättökoe, jonka suorittaminen antaa opiskelijalle kelpoisuuden hakea kolmannen asteen koulutukseen. Hakijan, korkeakoulujen ja yhteiskunnan näkökulmasta ylioppilastutkinto on kustannustehokas signaali hakijan kyvykkyydestä. Opetus- ja kulttuuriministeriön opiskelijavalinnan uudistamista suunnitellut työryhmä katsookin, että ylioppilastutkinnon tulisi olla korkeakoulujen opiskelijavalinnan ensisijainen valintaperuste. Työryhmän mielestä ylioppilastutkinnon sivuuttaminen opiskelijavalinnoissa on vaikeaa, sillä ylioppilastutkinto on tavoitavuudellaan ja riippumattomana, standardoituna ja vertailukelpoisena<sup>4</sup> osaamista arvioivana kokeena Suomessa ainutlaatuinen. Yli puolet kustakin suomalaisesta ikäluokasta suorittaa nykyisen ylioppilastutkinnon; kevään 2016 valintakierroksella 89 % yliopistoon hakeneista ja 95 % paikan vastaanottaneista oli suorittanut suomalaisen tai kansainvälisen ylioppilastutkinnon. Ylioppilastutkinto mittaa sitä, miten hyvin kokelas on omaksunut lukion opetussuunnitelman mukaisen ainekohtaisen tieto- ja taitotason sekä saavuttanut ylioppilaalta edellytettävän yleisen kypsyystason (Ylioppilastutkintolautakunta 2016). Kypsyydellä viitataan tässä yhteydessä esimerkiksi kriittisen ajattelun, oppimiskyvyn,

---

<sup>4</sup> Kupiainen et al. (2016) ovat esittäneet huolen siitä, että ylioppilastutkintoarvosanat eivät standardoidusta arviointimenetelmästä huolimatta ole olleet keskenään vertailukelpoisia eri aineiden ja tutkintokertojen välillä ylioppilaskokeiden vapaavalintaisuuden lisääntyttyä. Myös Ylioppilastutkintolautakunta (2014) on tiedostanut ongelman: pisteiden laskumenetelmä on johtanut tilanteisiin, joissa kokelaat saavat liian alhaisen arvosanan osaamiseensa nähden, ja joissa eri aineiden ja jopa saman aineen eri tutkintokertojen arvosanat eivät ole olleet vertailukelpoisia keskenään. Ongelma on korjattu vuonna 2014 käyttöön otetulla uudella pisteidenlaskumenetelmällä. Tutkielman kannalta havainto tarkoittaa käytännössä kahta asiaa: 1) aineiston opiskelijoiden ylioppilastutkintoarvosanat voivat kärsiä vertailukelvottomuudesta, jota pyritään korjaamaan ylioppilastutkinnon suoritusvuosien välistä vaihtelua kontrolloimalla ja 2) aineiston opiskelijoiden ylioppilastutkintoarvosanat eivät ole suoraan yhteismitallisia Kauppakorkeakoulun tulevien hakijoiden kanssa, mikä vaikuttaa saatujen tulosten yleistettävyyteen. Pitkällä aikavälillä ylioppilastutkintoarvosanojen laskennan uudistus tulee kuitenkin tarjoamaan entistä luotettavamman keinon hakijoiden vertailuun.

tiedon hankinnan ja opitun soveltamisen kaltaisiin taitoihin. Ylioppilastutkintoa hyödynnetäänkin opiskelijavalinnassa laajasti, usein jonkinlaisen opintomenestystä tai alalle soveltuvuutta parhaalla tavalla ennustavan pisteytysmallin kautta. (Opetus- ja kulttuuriministeriö 2016b)

Ylioppilastutkinto on kuitenkin koettu myös ongelmalliseksi opiskelijavalinnan kriteeriksi. Korkeakouluissa on koettu, etteivät ylioppilastutkinnon arvosanat tarjoa riittävän tarkkaa mittaria hakijoiden paremmuusjärjestyksen laskemiseen: kapean arvosana-asteikon ei koeta erottelevan hakijoita riittävästi. Ylioppilastutkintoon onkin ehdotettu tarkennuksia pisteidenlaskuun niin, että se erottelisi opiskelijat ja näiden osaamistasot entistä paremmin. Ylioppilastutkintoa on kritisoitu myös siitä, ettei se tarjoa mahdollisuutta alakohtaisen soveltuvuuden tai motivaation mittaamiseen. Opiskelijavalinnan toisena tärkeänä elementtinä onkin toimineet lukuisat alakohtaiset valintakokeet. Korkeakouluissa valintakokeiden uskotaan mittaavan erityisesti kyseisellä alalla tarvittavia tietoja ja taitoja, mutta myös erilaisten geneeristen valmiuksien, kuten esimerkiksi motivaation, tiedonhankintataitojen ja kriittisen ajattelun, hallintaa. Valintakokeen tavoitteen voidaan myös uskoa toteutuneen, mikäli siihen valmistautuminen onnistuneesti kasvattaa opiskelijan osaamista tai ymmärrystä alasta tai mittaa hakijan motivaatiota tai soveltuvuutta opiskella kyseistä alaa ja siten sitoutua opintojen loppuun saattamiseen. Valintakokeiden puolesta argumentoidaan myös, kun puhutaan lukiossa huonosti menestyneiden tai ammatillisen koulutuksen suorittaneiden väylistä yliopistokoulutukseen<sup>5</sup>. Nyt vallitseva näkemys kuitenkin on, että laajaa valmistautumista vaativat valintakokeet tuottavat sekä hakijoille että yhteiskunnalle huomattavan kustannuksen hidastaessaan opiskelijoiden siirtymää koulutusasteelta toiselle ilman pitäviä perusteita kokeiden tuomista hyödyistä. (Opetus- ja kulttuuriministeriö 2016b)

---

<sup>5</sup> Opetus- ja kulttuuriministeriön työryhmä (2016b) on ehdottanut, että ylioppilastutkintoa uudella tavalla hyödyntämällä esimerkiksi erilliset ylioppilaskokeet voisivat soveltua opiskelijavalinnassa käytettäväksi osaamisen arvioinniksi myös muille hakijoille kuin lukionsa päättävillä nuorilla. Tämän mahdollisuuden tarkastelu on yksi monista avauksista, jota ministeriön on tarkoitus työstää eteenpäin opiskelijavalinnan uudistuksen edetessä.

Opiskelijavalinnan arvioinnissa onkin ajankohtaista selvittää, tuovatko valintakokeet hakijoiden vertailuun perusteltavaa lisäarvoa. Taloustieteen näkökulmasta erillisen valintakokeen järjestäminen on opiskelijavalinnan kannalta kannattavaa ainoastaan, mikäli se antaa hakijoista sellaista parhaan valinnan kannalta arvokasta informaatiota, jota ei saada muutoin olemassa olevan tiedon perusteella pääteltyä. Valintakokeen järjestämiseen käytettävien resurssien tulee siis olla kustannuksiltaan yhtä suuret tai pienemmät kuin valintakokeen perusteella valittujen opiskelijoiden tuoma arvo suhteessa opiskelijoihin, jotka olisi valittu ilman sen järjestämistä. Resurssien optimoinnin kannalta korkeakoulun kannattaa investoida opiskelijavalintaan siihen pisteeseen asti, jossa hakijasta saadun lisäinformaation arvo on yhtä suuri kuin lisätyn valintakriteerin rajakustannus. Mahdolliset negatiiviset ulkoisvaikutukset huomioon otettaessa tulisi valintakokeen lisäarvon olla huomattava, jotta valintakokeen järjestäminen olisi kokonaisresurssien optimoinnin kannalta järkevää. Tutkielman empiirisen osan tavoitteena on esitellä ehdotus menetelmästä, jolla valintakokeen kannattavuutta opiskelijavalinnassa on mahdollista arvioida opiskelijavalinnan järjestävän korkeakoulun näkökulmasta. (Opetus- ja kulttuuriministeriö 2016b)

### 2.3 Opiskelijavalinnan ja opintomenestyksen empiirisen tutkimuksen esittely

Yliopistotason opiskelijavalintaa, opintomenestystä ja sen ennustettavuuteen vaikuttavia tekijöitä on tutkittu runsaasti niin Suomessa kuin ulkomailla. Opetusministeriön selvitysmies Sakari Ahola (2004) käsittelee tutkimuksia kolmessa menetelmäryhmässä Hypénin [1996] luokittelun mukaan: ensimmäinen keskittyy sosioekonomisten tekijöiden vaikutukseen koulutukseen hakeutumisessa ja valinnoissa pärjäämisessä, toinen tutkii aiemman koulumenestyksen ja yliopistoaikaisen opintomenestyksen yhteyksiä ja kolmas pyrkii ymmärtämään valintamenettelyn vaikutusta tulevaan opintomenestykseen. Tulevat kappaleet tarjoavat lyhyen katsauksen etenkin Hypénin tunnistamien toisen ja kolmannen menetelmän mukaisiin relevantteihin, suomalaista opiskelijavalintaa koskeviin lähivuosisikymmenten tutkimuksiin.

Suomalaiset opiskelijavalinnan tutkimukset koostuvat lähinnä yliopistojen sisäisistä, alakohtaisista selvityksistä, joilla on pyritty ymmärtämään ylioppilastutkintoarvosanojen ja

mahdollisen valintakokeen yhteyttä opintomenestykseen etenkin opintojen alkuvaiheessa. Opintomenestyksen ymmärtämisessä on ulkomaisissa tutkimuksissa useimmiten käytetyn kurssiarvosanakeskiarvon ohella keskitytty erityisesti opintojen etenemisen ja valmistumisaikojen ennustamiseen, sillä pitkät valmistumisajat on jo vuosikymmeniä koettu Suomessa ongelmalliseksi (Opetus- ja kulttuuriministeriö 2016b). Kotimaiset selvitykset ovat päätyneet eri aloilla poikkeaviin päätelmiin ylioppilastutkintoarvosanojen ja valintakokeen merkityksellisyydestä opiskelijavalinnassa. Yleinen löydös vaikuttaa olevan, että arvosanakeskiarvot korreloivat vahvasti aiemman ja tulevan opintomenestyksen välillä, mutta opintojen etenemistä on vaikea selittää opiskelijan valintahetkellä havaittavissa olevien ominaisuuksien perusteella. Yleisimmin erilaisia korrelaatio- ja regressioanalyyssejä hyödyntävissä tutkimuksissa on havaittu, että noin 25 % opiskelijan arvosanojen vaihtelusta on selitettävissä valintahetkellä havaittavissa olevilla tekijöillä (Häkkinen 2004). Opintojen etenemistahti vaikuttaa sen sijaan olevan voimakkaasti yhteydessä opintojen aikaisiin tapahtumiin. Esimerkiksi yhteiskuntatieteiden, liikuntatieteiden, kasvatustieteiden ja teknisen alan poikkitieteellisen tutkimuksen tehnyt Iida Häkkinen (2004) toteaa, että alasta riippuen suurimmillaan 10 - 30 % opintopistekertymien variaatiosta on ennustettavissa valintahetkellä. Myös nuorten taustatekijöillä on suuri vaikutus menestykseen opintoihin hakeutumisessa, opinnoissa suoriutumisessa ja myöhemmin työelämässä, vaikka kotitaustan vaikutus opiskeluun vaikuttaa pienentyneen lähivuosina (Opetus- ja kulttuuriministeriö 2016b).

Opiskelijavalinnan menettelyiden vaikutusta valittuun opiskelijajoukkoon tai näiden menestymiseen on Suomessa tutkittu hyvin vähän. Eri alojen opiskelijavalinnan tutkimuksia koostanut Ahola (2004) toteaa, että opiskelijavalinnan käytännöt ovat alakohtaisesti hyvin vaihtelevia, eikä yleistä päätelmää valintamenettelyjen toimivuudesta voida tehdä. Esimerkiksi valintakokeen oikea painotus ja ylioppilastutkinnon perusteella annettavien pisteiden kannattava laskentatapa vaihtelevat reilusti aloittain: Häkkinen (2004) havaitsee, että kasvatustieteissä valinnan suorittaminen puhtaasti ylioppilastutkintoarvosanojen perusteella kasvattaisi opiskelijoiden keskimääräistä opintomenestystä, mutta yhteiskuntatieteissä ja teknisellä alalla muutos olisi päinvastainen. Häkkisen identtisillä metodeilla usealla alalle toistettu tutkimus on hyvä osoitus siitä, että valintakriteerien ja opintomenestyksen yhteys on alakohtaista ja voimakkaasti vaihtelevaa – valintamenettelyt lienee siis jatkossakin syytä määritellä aloittain eriävien vaatimusten mukaisesti.

Toisaalta on tärkeää huomata, että Häkkisen tutkimusta lukuun ottamatta alakohtaisia tuloksia ei useinkaan ole saavutettu täysin identtisillä tutkimusasetelmilla, ja käytetyillä tutkimusmetodeilla voi olla tuloksiin suuriakin vaikutuksia. Janne Hakoniemi (1995) esimerkiksi toteaa Turun yliopiston historian opiskelijoita tutkittuaan, että valintakokeen painoarvoa historian opiskelijavalinnassa tulisi nostaa, sillä menestyneiden opiskelijoiden joukossa on opiskelijoita, joiden lähtöpisteet ovat matalat ja pääsykoepisteet korkeat. Samaa aineistoa analyysissään käyttänyt Ahola (2004) tulee kuitenkin ristiriitaisesti siihen tulokseen, että sekä lähtöpisteiden että valintakokeen ennustearvo on historian opiskelijoilla samankaltainen, ja Hakoniemen havainnoimaa heikompi. Ahola (2004) toteaaakin, että tutkimusten menetelmien luotettavuus ja käytettävyys vaihtelevat suuresti ja huomauttaa jo Tuula Lotin [1988] todenneen, että ”valinta saattaa toimia, mutta tutkimus ei”. Vaihtelevat aineistorajaukset ja tarkasteluvälit, datan laatu, mukaan otettavat muuttujat ja regressiomallien erot tuovat haasteita tulosten vertailulle. Tutkimukset kärsivät myös yleisestä selektio-ongelmasta: opintomenestystä mittaavissa tutkimuksissa mukana ovat luonnollisesti vain koulutukseen valituksi tulleet opiskelijat, jolloin luotettavaa vertailua valitsematta jääneiden hakijoiden potentiaalisesta menestyksestä on hyvin vaikeaa saavuttaa. Ilman poikkeuksellisia koeasetelmia parempiin ennustemalleihin on vaikea yltää.

Myös Aalto-yliopiston kauppakorkeakoulussa opintomenestystä ja opiskelijavalintaa on tutkittu runsaasti. Ensimmäisen laajan opiskelijavalintaa arvioivan tutkimuksen toteutti Arja Gillberg (1987), jonka opintomenestysanalyysi perustui opiskelijoiden taustatekijöiden, valintakoepisteiden ja ylioppilastutkintoarvosanojen lisäksi laajaan laadulliseen kyselyaineistoon. Gillberg havaitsi, että opiskelijoiden ylioppilastutkinnon yleisarvosana oli vahva arvosanakeskiarvon ennustaja, mutta aiemman koulumenestyksen yhteys opinnoissa etenemiseen oli pieni<sup>6</sup>. Valintakokeen ennustearvo sekä arvosanoihin että opintotahtiin oli

---

<sup>6</sup> Gillberg (1987) toteaa tutkimuksessaan, että ylioppilastutkinnon ja valintakokeen sijaan opintojen etenemisen parhaita ennustajia olivat tehokas opiskelutyyli, opintoihin sitoutuminen ja tutkielmanprosessiin liittyvät tekijät (mm. motivaatio ja palkitsevyys). Tutkijan mukaan opintomenestykseen vaikuttavat lisäksi valtavasti opiskeluympäristö, opiskelutyyli, opintojen organisointi sekä esimerkiksi opiskelijoiden sosiaaliset suhteet ja muu sopeutuminen opiskeluun.

huomattavasti ylioppilastutkintoa pienempi, mutta hyvät valintakoepisteet olivat yhteydessä matalaan keskeyttämisprosenttiin. Seuraavissa kappaleissa esiteltävät Aalto-yliopiston kauppakorkeakoulussa tai sen edeltäjissä useiden vuosikymmenten aikana tehdyt tutkimukset ovat päätyneet varsin vastaaviin päätelmiin. Tämä on havaintona merkittävä: voidaan todeta, että Kauppakorkeakoulun opiskelijoiden opintomenestyksen yhteys käytettyihin valintamenettelyihin on pysynyt samankaltaisena vuosikymmenestä toiseen.

Myös Erkki Komulainen ja Anne-Marja Nevgi (1993) perustivat tutkimuksensa valintakriteerien lisäksi laadullisiin kyselytuloksiin, jotka kartoittivat muun muassa opiskelijoiden minäkuva, ammatillista imagoa ja motivaatiota. Opintomenestyksen mittarina tutkijat käyttivät arvosanakeskiarvolla kevyesti painotettuja opintopistekertymiä, joiden ennustamisessa lukioaikainen kurssikeskiarvo ja ylioppilastutkintoarvosanat osoittautuvat tehokkaammaksi kuin opiskelijoiden valintakoepisteet. Laaja regressiomalli onnistui selittämään 27 % opintomenestysindeksin varianssista – suurin osa selitysvoimasta tuli kuitenkin tässäkin tutkimuksessa kyselytulosten laadullisista mittareista valintakriteerien sijaan. Komulainen ja Nevgi totesivat lisäksi, että valintakriteerien selitysvoima pienenee selkeästi opintojen edetessä – onkin selvää, että yksin asuva ja omaa talouttaan pyörittävä loppuvaiheen opiskelija on usein hyvin eri elämäntilanteessa kuin vanhempiansa luona asuva, valintakokeeseen valmistautuva nuori.

Erkki Komulainen (mm. 1987, 1990a, 1990b, 1992, 1996) on tutkinut laajasti myös valintakokeen merkitystä opiskelijavalinnassa ja yli kymmenenä peräkkäisenä vuonna tehdyt selvitykset tarjoavat kattavan kuvan valintakokeen ominaisuuksista ja kehityksestä. Useana vuonna toistettu tutkimus pyrki selvittämään, mikä yhteys hakijoiden taustatekijöillä ja lukiomenestyksellä on saatuihin valintakoepisteisiin, sekä miten valintakokeen eri osiot onnistuivat erottelemaan kokeen tehneet hakijat toisistaan. Komulainen toteaa, että Kauppakorkeakouluun pyrkivät hakijat ovat keskimäärin lukio-opinnoissaan hyvin menestyneitä opiskelijoita, ja että valituiksi tulleet ovat keskimäärin osoittaneet korkeampaa matemaattista lahjakkuutta aiemmissä opinnoissaan kuin muut hakijat. Tutkimustulokset

---

*Opintojen eteneminen vaikuttaakin liittyvän enemmän opintojen aikaisiin tekijöihin kuin valintahetkellä havaittavissa oleviin opiskelijoiden ominaisuuksiin.*

osoittavat myös, että korrelaatio eri valintakoeosioiden ja aiemman opintomenestyksen kanssa on usein voimakasta – valintakoe mittaa siis ainakin osittain samoja ominaisuuksia kuin muin keinoin havaittavissa oleva aiempi koulumenestys. Riitta-Liisa Peltonen, Arja Gillberg ja Erkki Komulainen (1990) ovat myös pyrkineet ennustamaan valintakokeessa menestymistä hakijoiden valmistautumista kartoittavan laadullisen kyselyaineiston perusteella.

Viimeisimmät tutkimukset opintomenestyksen ennustamisesta Aalto-yliopiston kauppakorkeakoululla ovat Lotta Pietilän (2012) ja allekirjoittaneen (Mieskonen 2016) tutkimukset ylioppilastutkintoarvosanojen ja valintakokeen selitysvuimasta. Pietilän tutkimus on taloustieteen laitokselle tehty maisterin tutkielma; Mieskosen tutkimus on tehty toimeksiantona Aalto-yliopiston kauppakorkeakoululle. Molempien tutkimusten pyrkimyksenä on ollut löytää valintakriteerien yhteys opintomenestykseen, jotta valintakriteereillä tunnistettaisiin jatkossa paremmin ne hakijat, joiden potentiaali menestyä opinnoissaan on suurin. Pietilä ja Mieskonen hyödynsivät molemmat aineistonaan Kauppakorkeakoulun omia opiskelijarekistereitä. Pietilän aineistona ovat vuosina 2005 – 2009 opintonsa aloittaneet opiskelijat sekä näinä vuosina kouluun ensisijaisesti pyrkineet hakijat, Mieskosen data rajautuu vuosina 2005 – 2013 kouluun hyväksytyihin ja paikan vastaanottaneisiin opiskelijoihin. Molemmat ennustavat opintomenestystä kolmen vuoden opintopistekertymien ja arvosanakeskiarvojen kautta Rothsteinin (2004) OLS-regressiomallilla; Mieskonen analysoi lisäksi opiskelijoiden tutkielma-arvosanoja, valmistumisaikoja sekä valmistumis- ja keskeyttämistodennäköisyyksiä kuvailevan tilastoanalyysin keinoin. Siinä missä Pietilän ensisijainen fokus on valintakokeen merkityksellisyyden mittaamisessa, Mieskosen tutkimus antaa analyysissaan suurimman painoarvon yksittäisten ylioppilastutkintoarvosanojen ennustearvon ymmärtämiseen. Tutkimusten lähtökohta on siis eri: Pietilä pyrkii selvittämään kannattaako erillisen valintakokeen järjestäminen, Mieskosen tavoitteena on ylioppilastutkinnon tehokkaan lähtöpisteytysmallin luominen. Tämän tutkielman pyrkimyksenä on yhdistää nämä tavoitteet opiskelijavalinnan arvioinnin empiiriseksi malliksi.

Sekä Pietilän että Mieskosen tutkimusten tulokset ovat linjassa Gillbergin (1987) havaintojen kanssa: valintakriteerien yhteys yliopistoaikaisiin arvosanoihin on suhteellisen



selkeä, mutta opintojen etenemistähdin ennustaminen tuntuu olevan kovin hankalaa Aalto-yliopiston kauppakorkeakoulussa. Pietilän mallilla käytetyt valintakriteerit ennustavat arvosanakeskiarvojen vaihtelusta hieman alle 25 %, kun opintopistekertymien vaihtelusta malli selittää vain noin 3 % (taulukko 1). Laajojen kontrollimuuttujien lisääminen nostaa opintopistekertymämallin selitysvoiman vajaaseen kymmeneen prosenttiin. Tulosten valossa vaikuttaa siltä, että opiskelijoiden etenemistähtiin vaikuttavat huomattavasti enemmän kontrollimuuttujien mittaamat ominaisuudet kuin ylioppilastutkintoarvosanat. Mieskosen selitysasteet ovat samankaltaiset, mutta kauttaaltaan hieman matalammat. Mieskosen analyysi opiskelijoiden valmistumisajan ennustamisesta vahvistaa kuitenkin opintopistekertymämallien tuloksia: ylioppilastutkinnon ( $R^2 = 6\%$ ) tai valintakokeen ( $R^2 = 0,01\%$ ) sijaan valmistumisaikaa selittävät parhaiten mallin kontrollimuuttujat ( $R^2 = 23\%$ ). Opintojen keskeyttämisen ennustaminen on valintakriteerien perusteella suoraviivaisempaa: Mieskosen mukaan korkeat ylioppilastutkintoarvosanat saanut opiskelija on Kauppakorkeakoulun todennäköisin keskeyttäjä. Näiden opiskelijoiden itseluottamus ja kyky saada toinen opiskelupaikka lienee muita opiskelijoita korkeampi. Ylioppilastutkintoarvosanat tarjoavat kuitenkin selkeästi paremman indikaattorin hakijoiden muusta opintomenestyksestä kuin erillinen valintakoe, jonka tuoma lisäarvo ylioppilastutkinnon rinnalla jää hyvin pieneksi.

Taulukko 1. Valintakriteerien ennustamerkityksellisyys

### Valintakriteerien ennustamerkityksellisyys

(Pietilä 2012)

	Opintopistekertymät	Arvosanakeskiarvot
Ylioppilastutkintoarvosanat	2,3 %	20,6 %
Valintakoe	0,5 %	4,0 %
$R^2$	2,8 %	24,6 %

Taulukko raportoi Pietilän (2012) kontrolloimattomien regressiomallien selitysasteet ja niiden jakautumisen valintakriteerien välille.

Taulukko 2. OLS regressiotulokset: Kolmen vuoden opintopistekertymät ja arvosanakeskiarvot (Mukaiutu: Pietilä 2012)

### OLS regressiotulokset: Kolmen vuoden opintopistekertymät ja arvosanakeskiarvot

(Pietilä 2012)

	Opintopistekertymät	Arvosanakeskiarvot
Valintakoepersentiili	.428*** (.112)	.011*** (.001)
Lähtöpisteytyspersentiili	.261*** (.047)	.010*** (.001)
Vakio	106.312*** (11.813)	1.801*** (.147)
Havaintoja	1336	1336
R <sup>2</sup>	.032	.205
Korjattu R <sup>2</sup>	.028	.201

Taulukko raportoi jatkuvan muuttujan  $\beta$ -kertoimen. Suluissa kertoimien robustit keskivirheet. Kertoimet ovat merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa t-testissä. Malli sisältää kontrollin opintojen aloitusvuodelle.

Vaikka valintakokeen tuoma ennustearvo jää käytetyissä malleissa vähäiseksi, on yhden valintakoesijan muutoksen merkitys opintomenestysmuuttujien vaihtelua selittäessä kokoluokaltaan ylioppilastutkinnon lähtöpisteitä suurempi (taulukko 2). Opintopistekertymämallissa Pietilä tulkitsee valintakoepersentiilin kertoimen 0,43 seuraavasti: opiskelija, jonka pisteet ovat yhtä suuret tai korkeammat kuin 80 % valintakokeen tehneistä hakijoista, suorittaa kolmen opiskeluvuoden aikana keskimäärin 9 opintopistettä enemmän kuin opiskelija, jonka pisteet ylittävät 60 % hakijoista. Vastaava kerroin ylioppilastutkinnon perusteella lasketuille lähtöpisteille on noin 40 % pienempi. Mieskosen tutkimuksessa valintakokeen kerroin ei ole täysin vertailukelpoinen, mutta valintakokeen minimi- ja maksimipisteiden saaneiden opiskelijoiden välillä ero on Pietilän tuloksiin nähden kaksinkertainen. Ero Pietilän ja Mieskosen tulosten välillä selittynee poissaolojen poikkeavilla käsittelyillä: Pietilän mallissa opintojen aikaiset poissaolot kontrolloidaan, Mieskosen mallissa ei. Arvosanakeskiarvojen ennustemallissa muuttujien kertoimet ovat sadasosan tarkkuudella samat: kerroin 0,01 ennustaa 80. persentiiliin

opiskelijan keskiarvon olevan keskimäärin 0,22 yksikköä 60. persentiilin opiskelijaa korkeampi.

Taulukko 3. Ylioppilastutkintoarvosanojen merkityksellisyys (Mieskonen 2016)

Ylioppilastutkintoarvosanojen merkityksellisyys

Äidinkieli	Pitkä matematiikka	Lyhyt matematiikka	Reaaliaineet	Pitkä kieli	Lyhyt kieli	
<b>L</b>						Menestyneimmät
<b>L</b>	<b>E</b>		<b>L</b>		<b>L</b>	
<b>E</b>				<b>L</b>		
<b>M</b>	<b>M</b>	<b>L</b>	<b>E</b>	<b>EM</b>	<b>E</b>	Keskiarvoiset
<b>CB</b>	<b>C</b>	<b>E</b>	<b>M</b>	<b>C</b>	<b>MC</b>	
<b>AI</b>	<b>BA</b>	<b>MC</b>	<b>CBAI</b>	<b>BAI</b>	<b>BA</b>	Heikoimmat
	<b>I</b>	<b>BAI</b>			<b>I</b>	

Mieskonen analyysin mukaan yksittäisten ylioppilastutkintoarvosanojen indikaattorimuuttujat tarjoavat kuitenkin opiskelijoiden arvosanakeskiarvoille paremman ennusteen kuin näiden perusteella lasketut lähtöpisteet. Indikaattorimuuttujat tarjoavat myös valmiiksi painotettua lähtöpisteytystä objektiivisemmän keinon tarkastella ylioppilasaineiden ja näiden arvosanojen yhteyttä opiskelijoiden opintomenestykseen. Tutkimuksen havainnot on koottu taulukkoon 3, joka esittelee arvosanaryhmien yhteyden opiskelijoiden keskimääräiseen opintomenestykseen. Koonti huomioi arvosanakeskiarvojen, opintopistekertymien ja valmistumisajan korrelaatio- ja regressioanalyysissä saadut tulokset painottaen suurimpia kertoimia ja selitysvoimaa niin että havaitut merkityksellisyserot ja arvosanojen keskinäiset suhteet on kuvattu mahdollisimman objektiivisesti. Tulokset tukevat desiilijakaumiin perustuneen visualisoinnin antamaa kuvaa ylioppilastutkintoarvojen merkityksellisyydestä opintomenestykselle. Mieskonen huomauttaa, että vahvat matemaattiset osaajat muodostavat opiskelijoiden selkeän huippuluokan: lähes 30 % laudaturin pitkässä matematiikassa kirjoittaneista opiskelijoista kuuluu arvosanakeskiarvoiltaan parhaiten menestyneeseen desiiliin. Vastaavasti lyhyen matematiikan kirjoittaneita ei desiilijaon

yläpäästä juuri löydy. Yleinen trendi on, että huippuarvosanoihin ylioppilaskokeissa yltäneet menestyvät myös yliopistossa.

Vaikka Pietilän ja Mieskosen tutkimukset perustuvat pitkälti samaan aineistoon ja lähestymistapaan, ei tutkimusten metodologia ole identtinen. Pietilä käsittelee ylioppilastutkintoarvosanoja jatkuvina muuttujina tai niiden perusteella lasketun lähtöpisteityksen kautta, Mieskonen selvittää yksittäisten aine-arvosana-indikaattorimuuttujien suoraa yhteyttä opintomenestykseen. Tutkimusten tulokset eivät ole siis suoraan vertailukelpoisessa muodossa. Tämän tutkielman tulevissa luvuissa esiteltävät menetelmät ovat luontevammin verrattavissa Pietilän analyysiin, mutta indikaattorimuuttujien käsittelyn kannalta vertailukohtana toimii paremmin Mieskosen tutkimus. Pietilän ja Mieskosen tutkimukset antavat hyvän pohjan ja vertailukohdan tutkielman ensimmäisen empiirisen vaiheen analyysille (luku 4), mutta eivät varsinaisesti tarjoa mallia opiskelijavalinnan arviointiin. Tutkimusten pohjalta voidaan todeta, että ylioppilastutkintoarvosanat vaikuttavat ennustavan arvosanakeskiarvoja kohtuullisesti, mutta opintojen etenemistahdin ennustajiksi niistä ei ole. Valintakoe tuo ennustamiseen vain hieman lisäarvoa opintomenestyksen määrittelystä riippumatta. Tutkimukset eivät kuitenkaan vastaa kysymykseen siitä, miten opiskelijavalinnan muuttaminen vaikuttaisi opiskelijoiden keskimääräiseen menestymiseen. Tutkimusten pohjalta ei voida myöskään luoda perusteltua kantaa siitä, tuoko valintakoe opiskelijavalintaan tarpeeksi lisäarvoa, jotta sen järjestäminen on kannattavaa. Tähän vastaamiseksi tulisi ymmärtää, pärjäävätkö valintakokeen ansiosta valituksi tulleet opiskelijat tarpeeksi paljon paremmin kuin niiden vuoksi valitsematta jääneet opiskelijat olisivat pärjänneet.

Valintamenettelyjen vaikutusta valittaviin opiskelijoihin ei ole Aalto-yliopiston kauppakorkeakoulussa aiemmin tutkittu. Malli valintamenettelyjen arviointiin on kuitenkin olemassa. Iida Häkkinen (2004) on tutkinut valintamenettelyjen vaikutusta valittaviin opiskelijoihin ja näiden opintomenestykseen yhteiskuntatieteiden, liikuntatieteiden, kasvatustieteiden ja teknisen alan opiskelijavalinnassa. Häkkisen kiinnostuksen kohteena on ollut ymmärtää, miten valintakriteerien muuttaminen – esimerkiksi valintakokeesta lähtöpisteisiin siirtyminen – muuttaisi valituksi tulevien opiskelijoiden kokoonpanoa ja keskimääräistä opintomenestystä. Vertailemalla marginaalisia opiskelijoita – hakijoita jotka

olisivat tulleet valituksi yhdellä valintamenettelyllä, mutta eivät toisella – Häkkinen toteaa, että valintamenettelyllä on selkeä vaikutus opiskelijajoukon keskimääräiseen opintomenestykseen. Esimerkiksi neljän vuoden opintopistekertymillä mitattuna joukot poikkeavat toisistaan 22 opintopisteen verran; suurimmillaan marginaalisten ryhmien välillä keskimääräinen seitsemän vuoden valmistumisprosentti yhdellä ryhmällä on 59 % ja toisella 31 %. Tulokset vaihtelevat voimakkaasti aloittain, eikä yhtenäistä päätelmää alojen välillä ole mahdollista tehdä – kaikilla aloilla suurin osa opiskelijoista tulisi kuitenkin valituksi valintamenettelystä riippumatta. Häkkisen menetelmät marginaalisten ryhmien tunnistamisessa ja valintamenettelyjen vertailussa ovat erittäin hyödyllisiä opiskelijavalinnan arvioinnin mallin rakentamisessa, ja niitä tullaan hyödyntämään tutkielman toisessa empiirisessä vaiheessa.

## 3 Aineiston esittely

Seuraavat alaluvut esittelevät tutkimuksessa käytetyn datan ja muuttujat sekä niiden ominaisuudet ja rajoitteet.

### 3.1 Data

Aineistona analyysissä on käytetty vuosina 2005 – 2012 Aalto-yliopiston kauppakorkeakoulun Helsingin kampuksen yhteisvalinnassa hyväksytyjen ja paikan vastaanottaneiden hakijoiden valintadataa, sekä kesään 2015 mennessä näille opiskelijoille kertynyttä kurssimerkintädataa sekä muita opiskelijarekisterimerkintöjä. Aineisto on haettu Aalto-yliopiston kauppakorkeakoulun tietokannasta kesäkuussa 2015. Aineistosta on rajattu pois opiskelijat, joilta puuttuu informaatio ylioppilastutkinnossa tai valintakokeessa suoriutumisesta<sup>7</sup>, sekä opiskelijat, jotka ensimmäisen vuoden poissaolon vuoksi eivät ole ehtineet suorittamaan kolmea täyttä opintovuotta aineiston tarkastelujakson puitteissa<sup>8</sup>. Rajausten jälkeen analyysissä on mukana 2285 opiskelijaa, jotka on hyväksytty opiskelemaan Aalto-yliopistoon kauppatieteellisen alan yhteisvalinnassa valintakoepisteidensä tai valintakoepisteisiin ja ylioppilastutkintoarvosanoihin perustuvien yhteispisteidensä perusteella.

---

<sup>7</sup> Ylioppilastutkintoarvosanat puuttuvat henkilöiltä, jotka ovat tulleet valituksi jonkin muun opiskelukelpoisuuden osoittavan tutkinnon perusteella. Näitä ovat *International Baccalaureate (IB)*, *European Baccalaureate (EB)* ja *Reifeprüfung (RP)*. Valintakoepisteet puuttuvat esimerkiksi *Scholastic Aptitude Test (SAT)* –koepisteiden, *VIKSU*-kilpailun tai avoimen yliopiston opintojen perusteella valituilta opiskelijoilta. Lisäksi vuonna 2007 valittiin 30 opiskelijaa suorassa todistusvalinnassa ilman valintakoetta. Näiden opiskelijoiden opintomenestystä suhteessa muihin opiskelijoihin on kuitenkin tutkittu aiemmin, katso Mieskonen (2015).

<sup>8</sup> Tutkimuksen vertailuajanjaksona käytetään opiskelijoiden kolmen ensimmäisen opintovuoden tunnuslukuja. Vuosikurssilla 2012 on 41 opiskelijaa, jotka ovat ehtineet suorittaa vasta kaksi kokonaista lukuvuotta ensimmäisen vuoden poissaolon vuoksi. Nämä opiskelijat on rajattu analyysistä pois selitettävien muuttujien vertailukelpoisuuden parantamiseksi. Tärkeimmät analyysit on toistettu nämä opiskelijat sisällyttäen, eikä opiskelijoiden poisrajaaminen vaikuta tutkimuksen tuloksiin laadullisesti.

Taulukko 4. Muuttujien yhteenveto

## Muuttujien yhteenveto

Muuttuja	Kaikki opiskelijat						Yhteispisteillä valitut			Koe pisteillä valitut		
	N	Ka	Med	SD	Min	Max	N	Ka	Med	N	Ka	Med
Opiskelijoiden lukumäärä	2285						1315			970		
<b>Selittävät muuttujat</b>												
3 vuoden opintopistekertymä	2270	147,94	156	49,1	9	354	1303	149,98	162	967	145,20	150
3 vuoden arvosanakeskiarvo	2263	3,60	3,62	0,57	1	5	1300	3,78	3,83	963	3,37	3,37
Yhteisten opintojen arvosanakeskiarvo	2209	3,43	3,44	0,77	1	5	1265	3,65	3,72	944	3,15	3,11
<b>Selittävät muuttujat</b>												
Valintakoe pisteet	2285	26,88	27	4,23	13,75	39	1315	26,56	26,50	970	27,31	27,13
Ylioppilastutkinnon lähtöpisteet	2285	24,60	26	8,62	0	40	1315	30,42	30	970	16,71	17
Ylioppilaskokeiden lukumäärä	2285	5,85	6	0,89	4	12	1315	6,15	6	970	5,45	5
Ylioppilaskoe arvosanojen keskiarvo	2285	5,42	5,50	0,88	2,33	7	1315	6	6	970	4,65	4,75
Äidinkieli	2285	5,24	5	1,16	0	7	1315	5,85	6	970	4,41	4
Pitkä matematiikka	1630	5,02	5	1,35	0	7	1045	5,63	6	585	3,92	4
Lyhyt matematiikka	605	5,79	6	1,16	2	7	263	6,38	7	342	5,33	5
Pitkä kieli	2275	5,59	6	1,12	2	7	1310	6,09	6	965	4,92	5
A-Englanti	2275	5,52	6	1,12	2	7	1310	5,99	6	965	4,87	5
Muu pitkä kieli kuin englanti	174	5,10	5	1,38	2	7	130	5,34	6	44	4,39	4
Lyhyt kieli	1877	5,44	6	1,29	0	7	1120	6	6	757	4,62	5
Reaaliaine	2193	5,69	6	1,20	0	7	1290	6,29	6	903	4,84	5
Historia	459	5,73	6	1,05	2	7	299	6,19	6	160	4,87	5
Yhteiskuntaoppi	658	5,81	6	1,06	2	7	443	6,20	6	215	4,99	5
Fysiikka	259	4,87	5	1,32	2	7	164	5,48	6	95	3,82	4
Kemia	171	4,82	5	1,40	0	7	117	5,38	5	54	3,61	4
Uskonto *	67	5,78	6	0,98	3	7	45	6,13	6	22	5,04	5
Filosofia	36	5,50	6	1,16	3	7	30	5,83	6	6	3,83	3,50
Psykologia	200	5,75	6	1,19	2	7	129	6,22	6	71	4,90	5
Biologia	131	5,29	5	1,16	2	7	77	5,86	6	54	4,48	5
Maantiede	194	5,85	6	1,12	3	7	118	6,30	6	76	5,14	5
Terveystieto	74	5,49	6	1,31	2	7	37	6,27	6	37	4,70	5
<b>Kontrollimuuttujat</b>												
Ikä valintahetkellä	2285	20,79	20	3,31	16	51	1315	20,51	19	970	21,18	21
Naisten osuus	2285	0,44		0,50	0	1	1315	0,48		970	0,38	
<b>Muut</b>												
Poissaoloprosentti	2285	0,14		0,35	0	1	1315	0,17		970	0,1	
Poissaolot vuosina poissaolleilla	315	1,01	1	0,26	0,5	5	221	1,00	1	94	1,03	1
Keskeyttämisprosentti	117	0,05			0	1	82	0,06		35	0,04	

\* Evankelisluterilainen uskonto, ortodoksinen uskonto tai elämäntätömustieto.

Opiskelijoiden valintadata sisältää tiedon opiskelijaan sovelletun valintamenettelyn lisäksi opiskelijan valintakoepisteistä sekä suoritetuista ylioppilastutkintokokeista, näiden arvosanoista, suoritusajankohdasta sekä ylioppilaaksi valmistumisvuodesta. Kauppakorkeakoulun opiskelijarekisteristä on saatu opiskelijoiden perustiedot (syntymäaika, sukupuoli, äidinkieli, yms.) sekä opinto-oikeus- ja läsnäolomerkinnät, eli muun muassa tiedot opinto-oikeuden alkamisesta ja päättymisestä, opinto-oikeuden päättymisen syystä (mm. valmistunut tai keskeyttänyt), opiskelun aloittamisesta, mahdollisista poissaoloista sekä suoritetuista tutkinnoista päivämäärineen. Kurssimerkintädata on haettu rekisteristä kokonaisuudessaan<sup>9</sup>, ja sisältää täyden informaation opiskelijan opintosuorituksista: kurssien nimet, koodit, tasot, opintopistemäärät, arvosanat, mahdolliset uusinnat, korotukset ja hyväksi lukemiset, suorituspäivämäärät sekä suorituksen kirjanneen laitoksen tiedot. Haetut tiedot on anonymisoitu ja yhdistetty yhdeksi aineistoksi opiskelijanumeron pohjalta luotuun tunnistenumeroon perustuen. Taulukko 4 vetää yhteen käytettyjen muuttujien ominaisuudet. Muuttujien jakaumakuviot löytyvät liitteestä 2.

### 3.2 Opintomenestyksen mittaaminen: selitettävät muuttujat

Opintomenestystä mitataan useimmiten opintojen laatua tai niissä edistymistä mittaavilla muuttujilla. Luvussa 2.1 tehtyyn opintomenestyksen määrittelyyn perustuen opiskelijoiden opintosuoritusta arvioidaan tutkielmassa näiden arvosanojen keskiarvoihin ja opintopistekertymiin perustuen. Opinnoista saadut arvosanat ovat vertailukelpoinen ja laajasti käytetty tapa mitata opiskelijoiden laadullista menestymistä opinnoissaan. Opintopistekertymät sen sijaan tarjoavat helposti ymmärrettävän ja tulkittavan tavan vertailla opiskelijoiden etenemistä opinnoissaan kohti suoritettavaa tutkintoa. Opintopistekertymiä ja arvosanakeskiarvoja tarkastellaan tutkielmassa opiskelijoiden kolmen ensimmäisen opintovuoden osalta. Lisäksi tutkitaan kaikille yhteisten opintojen arvosanakeskiarvoa. Taulukko 5 raportoi selitettävien muuttujien parittaiset

---

<sup>9</sup>Kurssimerkintädata on haettu Aalto-yliopiston kauppakorkeakoulun tietokannasta 2. – 5.6.2015. On siis huomattava, että lukuvuoden 2014 – 2015 data on aineistossa muutaman kuukauden vajaa, sillä lukuvuosi päättyy virallisesti heinäkuun loppuun. Kesällä suoritetaan kuitenkin keskimäärin hyvin vähän opintoja.



korrelaatiokertoimet. Opintopistekertymät korreloivat tilastollisesti merkitsevästi sekä kolmen vuoden ( $r = 0,36$ ;  $n = 2263$ ;  $p = 0,0000$ ) että kaikille yhteisten opintojen arvosanakeskiarvojen kanssa ( $r = 0,33$ ;  $n = 2209$ ;  $p = 0,0000$ ). Kolmen vuoden ja kaikille yhteisten opintojen arvosanakeskiarvojen välillä on voimakas positiivinen korrelaatio ( $r = 0,84$ ;  $n = 2209$ ;  $p = 0,0000$ ), mutta myös eroavaisuutta. Molempien muuttujien sisällyttäminen analyysiin lienee siis kannattavaa.

Kaikille yhteisten opintojen arvosanakeskiarvoja analysoidaan, jotta opiskelijoiden välille saadaan vertailukelpoinen menestysmittari ilman erikoistumisohjelmien aiheuttamaa varianssia. Myös kolmen vuoden muuttujien käytön perusteena on vertailtavuuden ja mielekkään analysoinnin varmistaminen: useamman vuoden yli ulottuvien muuttujien käyttö tasoittaa vuosittaisten muuttujien vahvaa vaihtelua ja korkeaa varianssia, silottaa opintojen alkuun liittyvien elämänmuutostekijöiden aiheuttamia väliaikaisia häiriötekijöitä ja antaa pitkäjänteisemmän kuvan opintojen sujumisesta (Häkkinen 2004). Kolmen vuoden vertailujakso vastaa myös Kauppakorkeakoulussa suoritettavan kolmivuotisen, 180 opintopisteen kauppatieteiden kandidaatin tutkinnon täysipäiväistä ohjesuoritusaikaa, joka tarjoaa perusteltavan vertailukohtaan analyysin opiskelijoiden etenemiseen opinnoissaan (Valtioneuvoston kanslia 2004).

Taulukko 5. Selitettävien muuttujien parittainen korrelaatio

### Selitettävien muuttujien parittainen korrelaatio

	(A)	(B)	(C)
(A) Kolmen vuoden opintopistekertymät	-		
(B) Kolmen vuoden arvosanakeskiarvot	.3556***	-	
(C) Yhteisten opintojen arvosanakeskiarvot	.3285***	.8425***	-

Taulukko raportoi parittaisen Pearsonin  $r$  korrelaatiokertoimen. Kertoimet ovat merkitseviä \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa  $t$ -testissä.

Monilla opiskelijoilla on poissaololukukausia opintojensa aikana (315 opiskelijaa, 14 % opiskelijoista), ja kaikkein yleisimmin lukukauden tai –vuoden pituinen poissaolo pidetään välittömästi opiskelupaikan vastaanottamisen jälkeen (aineistossa 300 opiskelijaa, eli 95 % kaikista poissaoloista). Opintojensa aloittamista lykkääviä opiskelijoita on vuosittain siis vajaat 40 kappaletta. Tästä syystä käytettävät kolmen vuoden menestysmittarit on analyysissa laskettu opintojen aloituksesta, ei opinto-oikeuden alkamisesta, lähtien. Opintojen aloituksen jälkeen kaikki lukuvuodet huomioidaan tasa-arvoisina, eikä mahdollisia ylimääräisiä poissaololukukausia oteta analyysissa huomioon. Näin toimitaan siksi, että opintojen etenemisen mittarin taustalla on motivaatio tutkintoon johtavan opiskeluajan lyhentämiseen, ja ylimääräiset poissaolot lykkäävät opiskelijan valmistumista vääjäämättä sen taustalla olevista syistä riippumatta tai jopa niistä johtuen. On huomattava, että poissaololla viitataan tässä lukukausiin, jolloin opiskelija on virallisesti ilmoittautunut poissaolevaksi yliopiston kirjoihin. Pietilä (2012) argumentoi, että useimmat opiskelijat eivät ilmoittaudu poissaoleviksi yliopistolle, vaikka eivät tosiasiasa suorittaisi lukukauden aikana yhtään opintopistettä, ja suorittaa laajan regression tällaisten poissaolojen syiden ymmärtämiseksi. Lukukausi voi kuitenkin näyttäytyä tilastoissa opintopisteettömänä, vaikka opiskelija olisi todellisuudessa edistänyt aktiivisesti opintojaan esimerkiksi tutkielmatyötä kirjoittaen tai yli lukukauden ulottuvia kursseja suorittaen. Tästä syystä tässä tutkimuksessa poissaoloiksi tulkitaan ainoastaan virallisesti tilastoidut ilmoittautumiset.

Seuraavat alaluvut esittelevät tarkemmin analyysissa käytettävät selitettävät muuttujat.

### 3.2.1 Kolmen vuoden opintopistekertymä

Opiskelijoiden opintojen etenemistä mitataan tutkielmassa kolmen vuoden kumulatiivisen opintopistekertymän muuttujalla. Useamman vuoden opintopistekertymiä<sup>10</sup>

---

<sup>10</sup> *Opintopistekertymien sijaan opinnoissa etenemistä voisi mitata myös muilla mittareilla, ja opiskelijoiden valmistumisaika olisi vaihtoehtoisista mittareista mielenkiintoisin. Tämä ei kuitenkaan valitettavasti ole datan puitteissa mielekäs: vuosikurskien 2005 – 2012 opiskeluvuosia 2005 – 2015 tarkastelevassa datassa ainoastaan neljä ensimmäistä vuosikurssia ovat ehtineet opiskella virallisen tutkinnon maksimisuoritusajan seitsemän vuotta. Valmistumisaikoja hyödynnetään kuitenkin kevyesti opiskelijavalinnan vaikutuksen arvioinnissa luvussa 5.*

opintomenestyksen tarkasteluun käyttävät myös muun muassa Häkkinen (2004), Pietilä (2012) sekä Kupiainen ja Marjanen (2014). Kolmen vuoden opintopistekertymä on laskettu yhdistämällä opiskelijoiden kurssidata läsnäoloinformaatioon. Kurssirekisteristä on ensin karsittu hylätyt ja korotetut kurssisuoritukset, ja suoritettut kurssit on suorituspäivämääränsä perusteella allokoitu lukukausille (syksy tai kevät). Läsnäoloinformaatiosta saatu opintojen aloituspäivä määrittelee aloituspisteen opiskelijan kolmen vuoden vertailujaksolle, josta lähtien opintopistekertymät on laskettu ensin vuositasolla, ja lopulta summattu yhteen kolmen ensimmäisen opiskeluvuoden osalta. Aineiston opiskelijoiden kolmen vuoden opintopistekertymät vaihtelevat välillä 9 – 354, keskimääräinen opintopistekertymä on 148. Muuttujan arvo puuttuu 15 opiskelijalta, jotka ovat jättäneet opintonsa paikan vastaanottamisen jälkeen aloittamatta.

### 3.2.2 Kolmen vuoden arvosanakeskiarvo

Opintopistekertymän tapaan myös opintosuoritusten arvosanoja käsitellään kolmen vuoden vertailujakson muuttujana. Pidempi ajanjakso mahdollistaa keskiarvon tasoittumisen opiskelijan todelliseen keskiarvoiseen suoritukseen, kun kaikille yhteisten perusopintojen lisäksi vertailuluku sisältää informaatiota myös vuodesta toiseen vaikeutuviissa erikoistumisohjelman opinnoissa suoriutumisen. Useamman vuoden arvosanakeskiarvoa käyttävät tutkimuksissaan muun muassa Rothstein (2004) ja Pietilä (2012). Opiskelijoiden arvosanakeskiarvot on laskettu perinteiseen tapaan opintopisteillä painotettuna aritmeettisena keskiarvona. Arvosanakeskiarvon laskennassa on huomioitu kaikki kolmen ensimmäisen opintovuoden aikana suoritettut opinnot, jotka on arvioitu numeroarvosanalla. Aineiston opiskelijoiden kolmen vuoden arvosanakeskiarvot vaihtelevat koko käytetyn arvosanaskaalan mukaisesti välillä 1 – 5, keskimääräinen arvosanakeskiarvo on 3,60. Muuttujan arvo puuttuu 22 opiskelijalta, jotka eivät ole suorittaneet hyväksytysti ainoatakaan numeroarvosanoin arvosteltavaa kurssia opintojensa aikana.

### 3.2.3 Kaikille yhteisten opintojen arvosanakeskiarvo

Kolmen vuoden arvosanakeskiarvojen lisäksi analyysissä huomioidaan kaikille yhteisten opintojen arvosanakeskiarvot. Kauppatieteiden kandidaatin tutkinto koostuu erikoistumisohjelman opintojen lisäksi kaikille yhteisistä perusopinnoista, niin kutsutuista

liiketoimintaosaamisen perusteista, jotka suoritetaan yleensä pääasiassa ensimmäisen opintovuoden aikana. Kaikille yhteisissä opinnoissa suoriutuminen on erittäin vertailukelpoinen menestysmittari niin vuosikurssien sisällä kuin rajoitetusti niiden välillä, sillä perusopintojen sisältö on pysynyt vuodesta toisiin hyvin samankaltaisena, eikä niiden arvosanat kärsi mahdollisista erikoistumisohjelmien välisistä arvosteluperiaatteiden tai vaikeustasojen eroavaisuuksista. Vuosina 2005 – 2012 opintonsa aloittaneet opiskelijat ovat suorittaneet perusopintonsa lähes poikkeuksetta vuoden 2005 tutkintosäännön perusteella, jolloin Liiketoimintaosaamisen perusteet koostuvat aineiston opiskelijoilla samoista kursseista vuosikurssista riippumatta (Helsingin kauppakorkeakoulun opintoasiaintoimisto 2005, 2006, 2007, 2008, 2009; Aalto-yliopiston kauppakorkeakoulun opintoasiaintoimisto 2010, 2011, 2012). Ensimmäisen vuoden perusopintojen arvosanakeskiarvoja tutkivat myös esimerkiksi Betts ja Morell (1999) sekä Rothstein (2004).

Kaikille yhteisten opintojen arvosanakeskiarvot on laskettu samoin kuin kolmen vuoden arvosanakeskiarvot yllä, huomioiden kaikki yhteisiin opintoihin sisältyvät, numeroarvosanoin arvioidut kurssit. Vuoden 2005 tutkintosäännön perusteella kaikille opiskelijoille yhteiset perusopinnot koostuvat seuraavista kursseista: Johtamisen perusteet, Matematiikkaa ja tilastotiedettä liikkeenjohtajille, Laskentatoimen perusteet, Rahoituksen perusteet, Taloustieteen perusteet, Yritysjuridiikan perusteet ja liike-elämän etiikka (vaihtoehtoisesti suoritettu kahtena kurssina: Yritysjuridiikan perusteet ja Liike-elämän etiikka), Markkinoinnin perusteet, Tuotantotalous ja Kvantitatiivinen analyysi taloustieteissä (Helsingin kauppakorkeakoulun opintoasiaintoimisto 2005).

Aineiston opiskelijoiden kaikille yhteisten opintojen arvosanakeskiarvot vaihtelevat välillä 1 – 5, keskimääräinen arvosanakeskiarvo on 3,43. Muuttujan arvo puuttuu 76 opiskelijalta. Puuttuvien arvojen määrä on hieman suurempi kuin kolmen vuoden arvosanakeskiarvon muuttujalla siksi, että pieni osa aineiston opiskelijoista on suorittanut kaikki perusopintonsa englanniksi, eikä näitä englanninkielisiä kursseja ole huomioitu kaikille yhteisten opintojen määrittelyssä.

### 3.3 Valintakriteereissä menestyminen: selittävät muuttujat

Analyysin selittävinä muuttujina ovat opiskelijoiden valintahetkellä mitattavissa olevat kyvykkyyden mittarit: opiskelijoiden ylioppilastutkintoarvosanat ja valintakoepisteet.

### 3.3.1 Ylioppilastutkintoarvosanat

Ylioppilastutkinto koostuu vähintään neljästä eri aineen ylioppilaskokeesta, jotka arvioidaan etukäteen määritellyn jakauman mukaisesti taulukon 6 osoittamalla tavalla. Kaikille pakollisen äidinkielen kokeen (suomi tai ruotsi) lisäksi tutkintoon on sisällytettävä vapaasti valiten toisen kotimaisen kielen, vieraan kielen, matematiikan tai jonkin reaaliaineen koe<sup>11</sup>. Tutkintoon voi sisällyttää myös vapaan määrän vapaavalintaisia kokeita. Aineiston opiskelijat ovat suorittaneet keskimäärin kuusi ylioppilaskoetta. (Ylioppilastutkintolautakunta 2014)

Taulukko 6. Ylioppilastutkinnon arvostelu (Ylioppilastutkintolautakunta 2014)

Kirjainlyhenne	Arvosana	Kompensaatiopisteet	Osuus arvosanoista
L	Laudatur	7	5 %
E	Eximia cum laude approbatur	6	15 %
M	Magna cum laude approbatur	5	20 %
C	Cum laude approbatur	4	24 %
B	Lubenter approbatur	3	20 %
A	Approbatur	2	11 %
I	Improbatur	0	5 %

Data sisältää tiedon opiskelijoiden ylioppilaskirjoitustuloksista kunkin aineen osalta saadun arvosanan tarkkuudella. Analyysissa ylioppilastutkintoarvosanat huomioitiin kuudessa ryhmässä: äidinkieli, pitkä matematiikka, lyhyt matematiikka, reaaliaineet, pitkä kieli ja lyhyt kieli<sup>12</sup>. Usean saman ryhmän aineen kirjoittaneilla (esimerkiksi monta eri reaaliainetta

---

<sup>11</sup> Pietilä (2012) huomauttaa, että syksyyn 2005 asti reaalikoe oli yksi yksittäinen koe, jossa oli kysymyksiä kaikkien reaaliaineiden alueelta. Kokeen suorittajan ei tarvinnut etukäteen ilmoittaa suorittamaansa aihealuetta, vaan sai vapaasti valita itselleen sopivimmat kysymykset koetilaisuuden aikana. Reaalikoe jaettiin vuonna 2005 erillisiin ainekokeisiin kannusteena yliopistoille hyödyntää relevantteja ainekokeita entistä enemmän opiskelijavalinnassa. Koska käsillä olevan tutkimuksen aineistona on Kauppakorkeakoulun vuosikurssit 2005 – 2012, on huomioitava, että vuosikurssien välillä on vaihtelua reaalikokeen luonteessa.

<sup>12</sup> Kupiainen et al. (2016) ovat esittäneet huolen ylioppilastutkinnon koeaineiden ja tutkintovuosien välisestä vertailukelvottomuudesta (alaviite 2). Tutkijat ehdottavatkin, että ylioppilastutkintoarvosanoja

kirjoittanut) on aina huomioitu aineryhmän paras arvosana. Äidinkielen pakollisen aineen lisäksi lähes kaikki aineiston opiskelijat ovat kirjoittaneet matematiikan (selkeästi yleisemmin pitkän (72 %) kuin lyhyen (26 %) laajuuden), pitkän kielen (99 %, useimmiten englannin) sekä jonkin reaaliaineen (96 %). Lyhyt kieli on ylioppilaskokeena hieman harvinaisempi (82 %). Aineiston opiskelijoiden ylioppilasarvosanat ovat maan huippuluokkaa, kuten kuvio 1 havainnollistaa: kaikkien aineiden arvosanat ovat painottuneet selkeästi asteikon yläpäähän, ja jopa yli 40 % opiskelijoista on saanut laudaturin tai eximian kaikissa aineluokissa. Matalampia arvosanoja sen sijaan aineiston opiskelijoilla on hyvin vähän; tilastollisista syistä arvosanalokat I ja A onkin jouduttu jättämään analyysistä pois<sup>13</sup>. Jakauma on selkeästi vinoutunut suhteessa ylioppilastutkintolautakunnan (kuviossa YTL) jakamiin arvosanoihin.

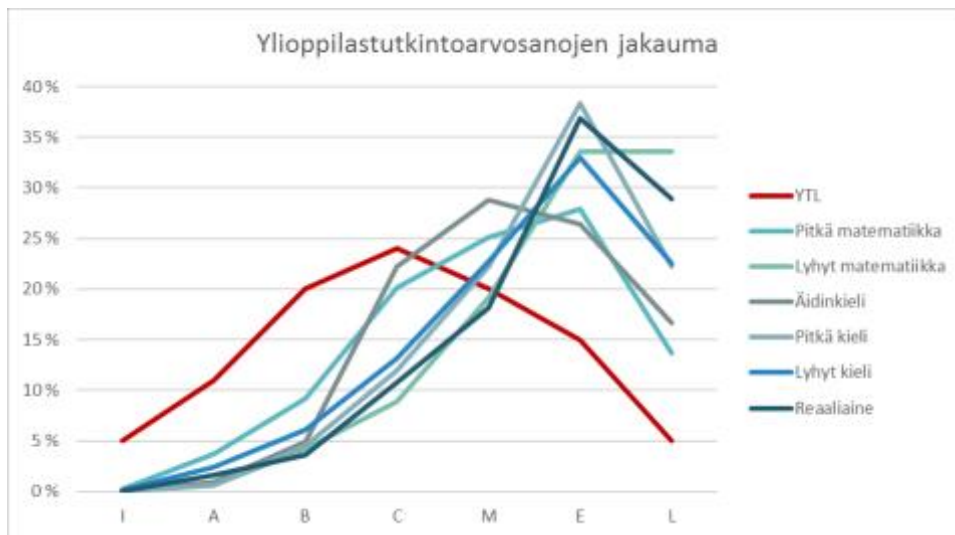
Ylioppilastutkintoarvosanoja käsiteltiin tutkimuksessa pääosin indikaattorimuuttujina, eli usean dikotomisen muuttujan ryhminä, sillä ylioppilastutkintoarvosanat ovat luonteeltaan kategorisia ja epätasaisesti jakautuneita. Indikaattorimuuttujaa käytettäessä jokaiselle arvosanalle luodaan oma dummy-muuttuja, ja tuloksia tulkitessa kutakin dummya verrataan yhteen poisjätettyyn kategoriaan. Esimerkiksi laudatur-dummyn ollessa vertailukategoriana tulkitaan eximian kerroin eximian ja laudaturin kirjoittaneiden opiskelijoiden välisenä erona, ja magnan kerroin magnan ja laudaturin kirjoittaneiden erona. Myös kyseisen aineryhmän ylioppilaskokeen kirjoittamatta jättäneet koodattiin omaksi luokakseen, jolloin kaikilla analyysin opiskelijoilla on jokin merkintä jokaisesta aineryhmästä – näin analyysi voidaan puuttuvista koesuorituksista huolimatta suorittaa kerralla koko aineistolle.

---

*hyödynnettäisiin koekohtaisesti: sen sijaan että opiskelijavalinnassa annetaan lähtöpisteitä ”kolmen parhaat pisteet antavan kokeen perusteella” kannattaisi määritellä, että esimerkiksi äidinkielen, matematiikan ja historian kokeista ansaitsee tietyn määrän pisteitä. Kokeet huomioidaan tässä kuitenkin koeryhmissä siksi, että opiskelijat ovat koevalinnoissaan hyvin hajautuneita eikä pienten aineiden erillinen huomioiminen ole analyysin kannalta mielekästä. Kevyt reaaliaineiden analyysi on kuitenkin tehty samalla aineistolla aiemmin, katso Mieskonen (2016).*

<sup>13</sup> *Arvosanalokat I ja A ovat aineistossa niin pienet, ettei ryhmien opintomenestyksen ennustaminen ole käytetyillä malleilla tilastollisesti mielekästä. Näitä arvosanoja käsitellään myöhemmässä analyysissä yhtenevästi puuttuvan arvosanan luokan kanssa, eli niin kuin kyseinen opiskelija ei olisi ainetta lainkaan kirjoittanut. Arvosanojen käsittely ei aiheuta ongelmaa tutkimuksen tulosten kannalta, sillä arvosanoista I tai A ei ole annettu tai ole jatkossakaan tarkoituksenmukaista antaa pisteitä opiskelijavalinnassa. Tärkeimmät analyysit on ajettu sekä ilman että sisällyttäen kyseiset luokat, eikä laadullisia eroja muiden arvosanaloukkien tuloksissa ole havaittu.*

Indikaattorimuuttujia käytettäessä on huomioitava, että jokainen regressiomuuttuja syö analyysistä aina yhden vapausasteen, ja arvosana-dummit vaativat usean muuttujan lisäämistä regressioon. Käytettävissä oleva aineisto on kuitenkin suhteellisen suuri, jolloin vapausasteiden menettäminen ei ole suuri ongelma. Mahdollisesta ylisovittamisen ongelmasta käydään keskustelua luvussa 4.1.2.



Kuvio 1. Ylioppilastutkintoarvosanojen jakauma

Lisäksi ylioppilastutkinnon analyysissä hyödynnettiin nykyisen opiskelijavalinnan mukaisia lähtöpisteitä (liite 1). Lähtöpisteitä käsiteltiin todellisten pisteiden sijaan persentiilinä Häkkisen (2004) ja Pietilän (2012) mallin mukaan. Persentiilit muodostettiin asettamalla opiskelijat vuosikursseittain lähtöpisteidensä mukaiseen paremmuusjärjestykseen suurimmasta pistemäärästä pienimpään. Yhtä suurille pistemäärille asetettiin sama sijaluku. Lopuksi sijaluvut skaalattiin välille 0 – 100 niin, että parhaat lähtöpisteet saaneet opiskelijat ovat vuosikursseittain persentiilisijalla 100 ja heikoimmat pisteet saaneet sijalla 0 (liite 3). Näin lasketut lähtöpistepersentiilit ovat standardoituja ja suoraan vertailukelpoisia sekä vuosikurssien välillä, että samoin tavoin laskettujen muiden persentiilimuuttujien kanssa.

Persentiilin sijaluku kertoo, kuinka suuri osa vuosikurssin opiskelijoista on saanut yhtä hyvät tai matalammat valintapisteet: persentiilisija 60 tarkoittaa, että opiskelijalla on korkeammat lähtöpisteet kuin 60 % samana vuonna valituiksi tulleista opiskelijoista. Muuttuja on jatkuva

ja sen regressiokerrointa tulkitaan yhden sijan muutoksena: jos opiskelijan persentiilisija kasvaa yhdellä (eli on yhden korkeampi), kasvaa tai laskee  $y$  saadun kertoimen verran. Varsinaisten lähtöpisteiden muutoksia ei muuttujalla voi tulkita, eikä ole tarvekaan: lähtöpisteiden (tai myöhemmin valintakoepisteiden) tarkoitus on vertailla saman valintakierroksen hakijoita toisiinsa, eivätkä ne sellaisenaan sovellu suoraan vuosikurssien välisen paremmuusjärjestyksen luomiseen. Persentiiliä käyttämällä mahdollinen muutos tai systemaattinen vuosivaihtelu annetuissa pisteissä ei pääse vaikuttamaan analyysiin, vaan opiskelijoita verrataan oman hakuvuotensa muihin hakijoihin ja näiden valintapisteisiin.

Lisäanalyysina ylioppilastutkintoarvosanoja käsiteltiin niiden arvosanakeskiarvon persentiilinä. Kupiainen ja Marjanen (2014) tutkivat Helsingin yliopiston seitsemän koulutusohjelman opiskelijoiden opintomenestystä kurssiarvosanojen ja opinnoissa etenemisen kautta, ja totesivat ylioppilastutkinnon koearvosanojen keskiarvon selittävän opintomenestystä erilaisin kaavoin laskettavia lähtöpisteitä paremmin usealla alla. Tämä on kiinnostava selvitys myös Kauppakorkeakoulussa. Keskiarvo laskettiin yksinkertaisena aritmeettisena keskiarvona opiskelijan kaikki ylioppilaskokeet niiden kompensatiopisteiden (taulukko 6) mukaan huomioiden. Keskiarvon voisi laskea myös painotettuna niin, että ainekohtaiset ylioppilastutkintoarvosanat kerrottaisiin aineen pakollisten tai esimerkiksi useimmin suoritettavien lukiokurssien määrällä.

### 3.3.2 Valintakoepisteet

Valintakoepisteet mittaavat opiskelijan suoriutumista kauppatieteellisen alan yhteisvalinnan valintakokeessa, jonka pisteenlasku perustuu positiivisia ja negatiivisia koepisteitä antaviin monivalintatehtäviin (liite 1). Valintakoepisteiden mahdollinen maksimiarvo on 40, ja valituksi tullakseen koepisteitä on saatava vähintään 10. Aineiston opiskelijoilla valintakoepisteet vaihtelevat välillä 13,75 ja 39 (keskiarvo 27). Myös valintakoepisteitä käsiteltiin todellisten pisteiden sijaan persentiilimuuttujana, joka on suoraan vertailukelpoinen muiden esiteltyjen persentiilimuuttujien kanssa. Valintakoepersentiili on myös samanmuotoinen Pietilän (2012) käyttämän valintakoemuuttujan kanssa, jolloin tutkimusten tulokset ovat näiltä osin vertailukelpoisia.



Valintakoepisteitä tai niiden persentiiliä tulkitessa on kuitenkin tärkeää muistaa, että valintakoepisteiden käsittely kyvykkyyden mittarina ei ole ongelmatonta. Yhteispistevalinnassa korkeat ylioppilastutkintoon perustuvat lähtöpisteet saanut hakija (60 % valittavista) tarvitsee vähemmän valintakoepisteitä tullakseen valituksi kuin pelkkien valintakoepisteiden perusteella arvioitavat hakijat (40 % valittavista). Opiskelijoiden keskuudessa tämän tiedetään aiheuttavan pistetaktikointia valintakokeessa sekä kokeeseen valmistautuessa, jolloin hyvä ylioppilastutkintomenestys näkyy huonompina valintakoepisteinä. Aineiston opiskelijoilla tämänkaltainen yhteys on empiirisesti todennettavissa korrelaatiokertoimilla: opiskelijoiden valintakoepisteet korreloivat negatiivisesti kaikkien aineiden ylioppilasarvosanojen kanssa. Vahvin negatiivinen korrelaatio on äidinkielen arvosanojen kanssa ( $r = -0,2055$ ,  $p = 0,0000$ ), heikoin lyhyen matematiikan kanssa (epämerkittävä yhteys:  $r = -0,0562$ ,  $p = 0,1481$ ).

Ilmiö on todennettavissa myös muuttujien keskiarvoissa (taulukko 4): ainoastaan valintakoepisteillä valittujen opiskelijoiden koepisteiden keskiarvo (mediaani) on 27,31 (27,13), kun yhteispisteillä valituilla vastaavat luvut ovat 26,56 (26,50). Vaikka keskiarvojen ero ei ole suuri, on ilmiö huomattava. Ylioppilastutkintoarvosanoja vertaamalla eri valintamenettelyillä valittujen opiskelijaryhmien välinen ero on keskimäärin kaksi arvosanaa yhteispisteillä valittujen opiskelijoiden hyväksi. Mikäli ylioppilastutkinto mielletään tarkoituksensa mukaiseksi vertailukelpoiseksi yleisen oppineisuuden (kyvykkyyden) mittariksi, voidaan todeta, että valintakoepisteet eivät yhtenevästi anna arviota opiskelijan kyvykkyydestä koko hakijamassalle. Opintomenestyksen ennustamisessa vaikuttaa kuitenkin siltä, että ryhmien välille ei valintakriteereistä johtuen ole tarvetta rakentaa erillisiä ennustemalleja, sillä valintamenettelyryhmille erillisinä toistetut regressiot päätyvät hyvin samankaltaisiin tuloksiin kuin koko opiskelija-aineistolle kokonaisuudessaan ajatut regressiot. Koko opiskelijamassalle tehdyt regressiot lienevät siis valintakokeen poikkeavasta valintamerkityksellisyydestä huolimatta sovellettavissa yhtä lailla molempien valintamenettelyjen perusteella valituksi tulleiden opiskelijoiden opintomenestyksen ennustamiseen.

### 3.4 Kontrollimuuttujat

Kontrollimuuttujat kuvastavat havaittavissa olevia tekijöitä, joiden voidaan kuvitella selittävän osansa analyysin selitettävien muuttujien variaatiosta, mutta joiden perusteella opiskelijavalintaa ei ole mahdollista suorittaa. Kontrollimuuttuja lisätään siis malliin vakioimaan ominaisuutta, jonka vaikutuksen tiedetään tai uskotaan vaikuttavan saatuihin ennusteisiin, mutta jonka vaikutuksesta tutkitulle ilmiölle ei varsinaisesti olla kiinnostuneita. Opintomenestyksen ennustamisessa tällaisia muuttujia ovat opiskelijoiden vuosikurssi ja ylioppilastutkinnon suoritusvuosi. Vuosikurssimuuttujan tavoitteena on vakioida mahdollinen kurssiarvostelujen tai muiden opintovuosien aikana tapahtuneiden muutosten vaihtelu; tutkintovuoden lisääminen pyrkii vähentämään mahdollista ylioppilaskirjoitusvuosien välistä varianssia ja lisäämään kokeiden välistä vertailukelpoisuutta<sup>14</sup>. Tutkintovuosi indikoi lisäksi välillisesti opiskelijan lukion päättämisen ja yliopisto-opintojen aloittamisen välillä kuluneen ajan eroja.

Opiskelijoiden vuosikurssilla viitataan tutkimuksessa kahteen asiaan. Ensinnäkin, valintakoepisteiden persentiili muodostettiin opiskelijan hakuvuoden mukaisesti (opinto-oikeuden alkamisvuosi), jolloin valintakoepisteitä verrataan toisiinsa saman koekerran sisällä. Toiseksi, opintomenestystä käsiteltiin opiskelijan opiskeluvuosikurssin mukaisesti (opiskelun aloittamisvuosi), jolloin kaikille yhteisiin opintoihin sisältyvät kurssit on todennäköisimmin suoritettu samanaikaisesti. Huomattavaa on, että suurimmalle osalle aineiston opiskelijoista tässä määritellyt vuosikurssit ovat samat. Regressiossa kontrollimuuttujaksi lisättiin näistä jälkimmäinen, joka otettiin mukaan usean dummy-muuttujan ryhmänä, indikaattorimuuttujana. Myös ylioppilastutkinnon suoritusvuosi lisättiin tutkimukseen indikaattorimuuttujana.

Aiemmissa opintomenestyksen ennustamiseen pyrkivissä tutkimuksissa kontrolleina on käytetty muun muassa opiskelijoiden sukupuolen, aloitusiän tai erikoistumisohjelman kontrolleja (mm. Häkkinen 2004, Pietilä 2012, Mieskonen 2016). Tämä ei kuitenkaan ole mielekäästä, kun opintomenestyksen ennustaminen liitetään opiskelijavalinnan arviointiin.

---

<sup>14</sup> *Lyhyt perustelu tutkintovuosien välisestä mahdollisesta vertailukelvottomuudesta alaviitteessä 2.*

Aiempien tutkimusten tuloksista tiedetään, että esimerkiksi opintomenestysmittarit tai hakijoiden ylioppilastutkintoarvosanat ovat keskimäärin erilaiset naisilla ja miehillä. Mikäli hakijoiden sukupuoli lisättäisiin malliin kontrollimuuttujana, ennustaisi regressiomalli tällöin esimerkiksi ylioppilastutkintoarvosanojen vaikutuksen opiskelijoiden arvosanakeskiarvoihin erikseen molemmille sukupuolille. Opiskelijavalinnan tavoitteen oletetaan tässä kuitenkin olevan parhaiden potentiaalisten opiskelijoiden valitseminen riippumatta heidän iästään, sukupuolestaan tai muista henkilökohtaisista ominaisuuksistaan<sup>15</sup>. Valintamenettelyjä ei ole mahdollista suunnitella naisille ja miehille erikseen, jolloin opintomenestystä ei ole kannattavaa ennustaa sukupuolten välillä eri tavoin. Tällöin opintomenestyksen ennustamisessa on luontevaa hyödyntää ainoastaan muuttujia, joiden perusteella opiskelijavalintaa voidaan tosiasiaassa suorittaa. Muiden opintomenestykseen vaikuttavien tekijöiden heijastuminen näihin muuttujiin on tällöin välttämätöntä, mutta tavoitteen kannalta jopa toivottavaa.

---

<sup>15</sup> Opetus- ja kulttuuriministeriö (2016b) on ilmaissut kiinnostuksena niin kutsuttujen ensikertalaisten hakijoiden valintamahdollisuuksien kasvattamiseen. Ensikertalaisia suosivan valinnan tavoitteena on yhtä lailla valita parhaat potentiaaliset opiskelijat annetusta hakijajoukosta, hakijajoukko vaan jaetaan erillisiin ”valintakoreihin”, joista valinta järjestyksessä suoritetaan. Tämä on poliittinen, suomalaisten nuorten nopeampaan valmistumiseen tähtäävä aloite, jolla ei ole merkitystä opintomenestyksen ennustamisen menetelmiin tutkielmassa.

## 4 Opintomenestyksen ennustaminen

Ensimmäisen empiirisen vaiheen tarkoituksena on tutkia ylioppilastutkintoarvosanojen ja valintakokeen ennustearvoa opiskelijan opintomenestykseen Aalto-yliopiston kauppakorkeakoulussa. Pyrkimyksenä on tunnistaa opintomenestyksen kannalta merkityksellisimmät ylioppilasaineet ja näiden arvosanat, ja muodostaa näiden perusteella käsitys menestyneimpien opiskelijoiden aine-arvosana kombinaatioista. Lisäksi tavoitteena on selvittää tuoko valintakoe ylioppilastutkintoarvosanoihin nähden lisäarvoa opintomenestyksen ennustamiseen. Tulokset koostetaan ylioppilastutkintoarvosanojen vaihtoehoiseksi pisteytysmalliksi, joka tarjoaa pohjan tutkielman toisen empiirisen vaiheen keinotekoiselle opiskelijavalinnalle.

Luku pyrkii vastaamaan seuraaviin tutkimuskysymyksiin:

- 1) Miten ylioppilastutkintoarvosanat ja valintakoe ennustavat opintomenestystä?
- 2) Millainen ylioppilastutkintoarvosanojen valintapisteytysmallin tulisi niiden ennustearvon mukaan olla?

### 4.1 Metodologia

Valintakriteerien ja opintomenestyksen välinen yhteys eli ylioppilastutkintoarvosanojen ja valintakokeen ennustearvo opintopistekertymiin ja arvosanakeskiarvoihin pyrittiin löytämään lineaarisella OLS regressiomallilla (pienimmän neliösumman menetelmä). Käytetty malli perustuu Rothsteinin (2004) mallinnukseen korkeakouluarvosanojen ennustamisesta. Opiskelijan  $i$  opintomenestys  $y_i$  mallinnettiin analyysissa seuraavasti:

(3)

$$y_i = \alpha + \beta_1 A_i + \beta_2 V_i + \gamma X_i + \varepsilon$$

Yhtälössä  $A_i$  viittaa ylioppilastutkintoarvosanoihin,  $V_i$  valintakoe pisteisiin ja  $X_i$  kontrollimuuttujiin,  $\varepsilon$  on virhetermi. Malli ei ole kausaalinen, pikemminkin tutkimuksen on tarkoitus deskriptiivisesti löytää ja selittää opintomenestystä parhaiten kuvaava

ennustemalli. Ylioppilastutkintoarvosanojen ja valintakoepisteiden tai näiden muunnelmien ei odoteta olevan opintomenestyksen syy tai lähde, vaan pikemminkin opiskelijan kyvykkyyttä tai menestymispotentiaalia heijastava, todettavissa oleva mittari. Valintakriteerien voidaankin toimiessaan ajatella olevan opiskelijan kyvykkyyden surrogaatti. Tutkimuksen keskeisenä kiinnostuksenkohteena on selvittää, kuinka paljon ennustevoimaa ylioppilastutkintoarvosanoilla on, ja kuinka paljon selitysvoimaa menetettäisiin, mikäli valintakoepisteitä ei olisi käytettävissä opiskelijavalinnassa. Tästä syystä analyysiin sisällytetään usean regressiomallin vertailu.

Malleja vertailemalla tunnistetaan kertoimissa ja selitysvoimassa tapahtuvien muutosten syyt yksittäistä regressiomallia tarkemmin. Selitysvoima onkin ennustemallia etsivässä tutkimuksessa avainasemassa, sillä ilman tarpeeksi korkeaa selitystasetta malli ja muuttujien kertoimet ovat arvottomia. Selitysvoimaa mitattiin tutkimuksessa taloustieteissä tuttuun tapaan  $R^2$  mitalla, joka osoittaa minkä osuuden selitettävän muuttujan vaihtelusta regressiomalli onnistuu selittämään.  $R^2$  lasketaan selitettävän muuttujan arvojen ja ennustearvojen korrelaation neliönä ja vaihtelee nollan ja yhden välillä. Lisäksi estimaatin keskimääräinen neliövirhe eli regressiomallin virhetermien keskihajonnan tunnusluku mittaa mallin selitysvoimaa: mitä pienempi keskivirhe, sitä pienempi virhetermien hajonta ja suurempi selitysvoima. Regressiomallien yhteydessä raportoidaan myös korjattu  $R^2$  (adjusted R-squared), joka ottaa huomioon malliin lisättyjen selittävien muuttujien lukumäärän, ja jota käytetään regressiomallien keskinäiseen vertailuun. Tällöin voidaan Rothsteinin mukaan laskennallisesti ajatella valintakriteerien kontribuution selitysvoimaan saatavan yhtälöstä  $\Delta R^2 \equiv R_A^2 - R_B^2$ , jossa valintakriteerien vaikutusta mallin selitysvoimaan tulkitaan suhteessa muihin malliin lisättäviin muuttujiin.

Opintomenestysmittareiden muodostumista pyritään tutkielmassa selittämään yhteensä kuudella regressiomallilla. Seuraavissa luvuissa käsitellyt mallit (1) ja (3) käsittelevät ylioppilastutkintoarvosanoja yksittäisinä indikaattorimuuttujina ja mallit (4) – (6) esittelevät vaihtoehtoisen tavan käsitellä ylioppilastutkinnon tarjoamaa informaatiota sen arvosanojen keskiarvon (4), nykyisen opiskelijavalinnan lähtöpisteytyksen (5) tai indikaattorimuuttujien ennustearvon (6) persentiileinä. Ylioppilastutkinnon arvosanojen keskiarvo ja lähtöpisteytys on laskettu luvussa 3.3.1 esitellyn mukaisesti; indikaattorimuuttujien ennustearvon persentiili on luotu niin, että opiskelijat on laitettu paremmuusjärjestykseen mallin (1)

perusteella ennustettujen opintomenestysmittareiden arvojen perusteella ja paremmuusjärjestys on aiemmin esitellyn mukaisesti skaalattu välille 1 – 100. Valintakoepisteiden persentiili on mukana malleissa (2) – (6).

#### 4.1.1 OLS-oletukset

Harhattomia ja tehokkaita ennusteita tuottaakseen OLS-regressio vaatii kuuden oletuksen toteutuvan: 1) selittävät muuttujat ovat eksogeenisiä, eli eivät korreloi virhetermin kanssa, 2) malli on lineaarinen, eli selitettävän ja selittävien muuttujien välillä vallitsee lineaarinen yhteys, 3) selittävät muuttujat eivät korreloi liian vahvasti keskenään (multikollineaarisuus), 4) virhetermi on hajautunut selittävästä muuttujista riippumatta eli homoskedastinen, 5) havainnot eivät ole aikariippuvia ja 6) suuria poikkeavia havaintoja ei ole. (Angrist & Pischke 2008)

Selittävien muuttujien eksogeenisuusrajoitteita on käsitelty luvussa 3.4. Lineaarisuusoletus on tarkistettu muuttujien graafisella mallinnuksella (liite 4), josta lineaarinen yhteys muuttujien välillä on visuaalisesti havaittavissa. Kuvioiden perusteella vaikuttaa siltä, että valintakoepisteiden tai niiden persentiilin ja opintomenestysmittareiden välinen lineaarinen yhteys on olemassa, mutta hyvin kevyt. Muiden selittävien muuttujien yhteys opintomenestysmittareihin on selkeämpi. Ylioppilastutkintoarvosanojen voidaan olettaa olevan hyvin korreloituneita keskenään, jolloin multikollineaarisuus on mallissa mahdollinen ongelma. Multikollineaarisuutta on mitattu regressioissa VIF-testillä (Variation Inflation Factor) ja toleranssimittareilla. Mallin korkeimmat multikollineaarisuusarvot  $VIF = 2,65$  ja toleranssi  $= 0,3778$  ovat kuitenkin selkeästi suositusrajoissa ( $VIF$  suositus  $< 10$  ja toleranssi  $> 0,1$ ), eikä multikollineaarisuuden tulisi näiden perusteella aiheuttaa regressiolle haasteita.

Homoskedastisuuden varmistamiseksi regressiomallit on ajettu robusteilla keskivirheillä. Lisäksi virhetermit mallinnettiin visuaalisesti virhetermin hajonnan satunnaisuuden varmistamiseksi: opintopistekertymien malleissa virhetermit ovat paikoin keskittyneitä, mutta lineaarisia yhteyksiä selittäviin muuttujiin ei havaita; arvosanakeskiarvojen malleissa virheet ovat toivotusti satunnaisesti hajaantuneet. Havaintojen aikariippuvuus ei ole mallissa ongelma, sillä mallin muuttujat eivät ulotu useiden aikaperiodien yli: esimerkiksi

opintopistekertymiä hyödynnetään yhtenä kolmen vuoden summalukuna kolmen erillisen vuosimuuttujan sijaan. Vuosikurssien ja tutkintovuosien välinen vaihtelu on vakioitu kontrollimuuttujilla. Viimeiseksi, poikkeavat havainnot eivät tuo mallissa haasteita, sillä persentiilimuuttujien arvot on standardoinnilla rajattu tietylle vaihteluvälille ja muiden muuttujien kanssa haastavia arvoja ei havaita. Regressiomallin voidaan siis todeta täyttävän OLS-oletukset tarvittavalla tasolla.

#### 4.1.2 Ylisovittaminen

On mahdollista, että tutkielman regressiomallit kärsivät ylisovittamisen ongelmasta etenkin useista muuttujaryhmistä koostuvia ylioppilastutkintoarvosanojen indikaattorimuuttujia käytettäessä. Ylisovittamista tapahtuu, kun regressiomallissa on liikaa selittäviä muuttujia suhteessa ennustettavan joukon kokoon. Mitä enemmän parametreja spesifioidaan, sitä enemmän mallin ennusteet mukailevat aineistoa ja sitä suuremmaksi mallin selitysarvo nousee. Esimerkiksi 100 opiskelijan aineistoa 100 muuttujalla ennustettaessa tuloksena on aineistoa täydellisesti mukailevat sovitearvot tai ”ennusteet”.

Ylisovitetun mallin tulkinta on intuitiivisesti ymmärrettävissä alisovitetun mallin kautta. Parhaaseen mahdolliseen ennustemalliin verrattuna alisovitettu malli jättää huomioimatta datassa olevia, ennusteen kannalta tärkeitä yhteyksiä. Ilmiö on verrattavissa kausaalisen tutkimuksen puuttuvan muuttujan harhaan: mallista on jäänyt pois muuttujia, jotka olisivat onnistuneesti selittäneet  $y$ :n vaihtelua. Näiden puuttuvien muuttujien vaikutus on nyt huomioitu mallin muiden muuttujien ja virhetermin kertoimissa ja malli on harhainen. Ylisovitetussa mallissa on päinvastoin mukana liikaa muuttujia, jolloin kertoimien ennusteet ovat suhteessa harhattomia, mutta niiden varianssi on oikeaa ennustemallia suurempi, sillä ennustetut arvot mukailevat aineistoa turhankin tarkasti. Tällöin ennusteen tarkkuus jää huonommaksi kuin optimaalisen ennustemallin perusteella olisi odotettavissa, ja malli estimoii yhteyksiä, jotka eivät ole relevantteja  $y$ :n ennustamisen kannalta. Malli on optimaalinen, kun ennustevirhe ja otosvarianssi ovat tasapainossa ja ennusteet ovat juuri hyväksyttävän luottamusvälin sisällä. (Burnham & Anderson 2002)

Ylisovittamisen haittoja pohtivassa kirjallisuudessa ongelma liitetään usein tiedonlouhinnan (engl. data mining) keinoin rakennettuihin ennustemalleihin: ylisovittamisen riski kasvaa

sitä suuremmaksi, mitä enemmän muuttujia malliin lisätään tehtyjen ennusteiden tai korrelaatioanalyysien pohjalta, kunnes saadut estimaatit ovat ”tarpeeksi hyviä” (Burnham & Anderson 2002)<sup>16</sup>. Ylisovittamisen riski on pienempi, mikäli ennustemalli spesifoidaan etukäteen määriteltyjen oletusten pohjalta ja arvioidaan sellaisenaan ilman lisäyksiä. Tästä näkökulmasta arvioiden ylisovittaminen ei olisi tutkielman asetelmalle suuri ongelma: esitellyt regressiomallit perustuvat suoraan etukäteen määriteltyyn, intuitiivisesti perusteltavaan ennustemalliin, joka sisältää vain opiskelijavalinnan hetkellä aidosti käytetyt muuttujat ja näiden välttämättömät kontrollit. Viimeisessä regressiomallissa (6) tilanne on kuitenkin eri, sillä mallissa hyödynnetään aiemman ennustemallin (1) tuottamaa lisäinformaatiota uuden ennusteen tekemiseen. Ylisovittamisen ehkäisemiseksi malli validoidaan out-of-sample menetelmällä.

Out-of-sample mallintamisessa regression ennustearvot validoidaan testaamalla niitä toiseen otokseen: yksinkertaisimmillaan käytettävissä oleva aineisto jaetaan kahteen osajoukkoon, joista ensimmäistä käytetään mallin rakentamiseen ja opintomenestyksen ennustamiseen (in-sample), ja toista ennustettujen parametrien testaamiseen ja mallin validointiin (out-of-sample). Tällainen yksinkertainen ristiin validointi tarjoaa analyysihetkellä käytettävän työkalun mallin ennustevoiman tarkistamiselle. Aineiston opiskelijoilla ristiin validointi on tehty niin, että kunkin vuosikurssin opintomenestys on ennustettu aineiston muiden vuosikurssien perusteella sovitettujen mallin parametreilla: esimerkiksi niin, että regressiomallin (1) parametrit on ennustettu vuosikurssien 2005 – 2011 opiskelijoilla, ja vuoden 2012 opintomenestys on ennustettu näihin parametreihin perustuen. Ristiin validoitu, regressiomallissa (6) käytetty indikaattorimuuttujien ennustearvojen persentiili perustuu out-of-sample mallinnuksen perusteella ennustettujen sovitearvojen skaalattuun paremmuusjärjestykseen, eikä näin ole yhtä altis ylisovittamisen riskille. Näin sovitettujen parametrien ja niiden ennustevoiman tilastollisten testien ajatellaan antavan luotettavamman

---

<sup>16</sup> Tällaisen mallinrakentamisen yhteydessä ylisovittamisen riskiä voidaan pienentää niin kutsutun Akaiken informaatiokriteerin (AIC, Akaike information criterion) avulla. AIC estimoi niin kutsuttuun informaatioteoriaan perustuen tilastollisen mallin suhteellista laatua verrattuna aineiston perusteella rakennettavissa oleviin muihin ennustemalleihin. AIC arvioi mallin ja aineiston yhteensopivuutta suhteessa mallin monimutkaisuuteen ja toimii siten työkaluna mallin valintaan. Tutkielman analyysiin AIC ei kuitenkaan tuo välttämätöntä lisäarvoa, sillä se ei tarjoa arviota ennustemallin absoluuttisesta laadusta.



kuvan mallin selitysvoimasta kuin koko aineiston perusteella sovitetut mallit antaisivat. Paras validointi mallille olisi ennustaa sen perusteella Kauppakorkeakouluun nyt valittavien uusien opiskelijoiden opintomenestystä ja muutaman vuoden kuluttua tutkia, kuinka hyvin ennusteet pitävät paikkaansa.

Ylisovittamista voi usean muuttujan mallissa tapahtua, vaikka malli olisi teoriassa oikein rakennettu. Jo käsitellyn lisäksi tutkielmassa ylisovittamisen riski on olemassa etenkin indikaattorimuuttujia hyödyntävissä malleissa. Käytännössä ylisovittamiseen viittaavia merkkejä näkyy esimerkiksi opintopistekertymien ennustemallin indikaattorimuuttujien kertoimia esittelevässä taulukossa 7, jonka kertoimia suoraviivaisesti tulkitsemalla voisi kuvitella, että lyhyestä matematiikasta C:n kirjoittaneet opiskelijat ( $\beta = 16,56$ ) ovat edenneet opinnoissaan poikkeuksellisen nopeasti. Olisi kuitenkin virheellistä tämän perusteella todeta, että lyhyen matematiikan C ennustaisi korkeita opintopistekertymiä. Huomattavasti todennäköisempää on, että kyseisen arvosanaryhmän kerroin kärsii ylisovittamisesta: pienten arvosanaryhmien kohdalla ennusteessa käytettävien muuttujien suhteellisen suuri määrä on johtanut tilanteeseen, jossa sovitearvot kuvaavat liiankin hyvin kyseisen otoksen jäseniä (lyhyestä matematiikasta C:n kirjoittaneet tämän aineiston opiskelijat), mutta eivät tarjoa hyviä ennusteita kuvitellulle populaatiolle (kaikki lyhyestä matematiikasta C:n kirjoittaneet hakijat). Ylisovitetussa mallissa saadut ennusteet selittävät hyvin kyseisen joukon jäsenten opintomenestystä, mutta tulokset eivät ole samalla tarkkuudella yleistettävissä laajemmalle joukolle. Mikäli ylisovittamisessa esiintyy, ovat mallin tarkkuutta kuvaavat estimaatit, esimerkiksi selitysarvo  $R^2$ , suurempia kuin muuttujien todellisten yhteyksien perusteella tulisi olla. Ylisovittamisen riski on huomioitava seuraavissa luvuissa esiteltyjä tuloksia yleistettäessä.

## 4.2 Ennustetulokset

Seuraavat alaluvut esittelevät opintomenestyksen ennustamisen regressiotulokset opintomenestysmittari kerrallaan.

#### 4.2.1 Kolmen vuoden opintopistekertymä

Kolmen vuoden opintopistekertymien regressiotulokset on esitelty indikaattorimuuttujien kerrointen osalta taulukossa 7 ja mallien (1) – (6) osalta taulukossa 8. Opintopistekertymien ennustaminen ei vaikuta mallin muuttujilla olevan kovinkaan mielekäästä. Ylioppilastutkintoindikaattorit selittävät vain vajaat 8 % opintopistekertymien vaihtelusta (malli (1)), ja indikaattorimuuttujien kokonaismerkitsevyys jää mallissa alle vaaditun merkitsevyystason. Korkeimpaan selitysasteeseen yltävä, sekä ylioppilastutkinnon arvosanaindikaattorit, että valintakoepercentiin sisältävä malli (3) sopii aineistoon vain hieman paremmin. Mallin (1) mukaisella out-of-sample regressiolla luotua mallia (6) voidaan pitää ylioppilastutkintoarvosanojen indikaattorimuuttujien ennustearvon testinä: mikäli malli (1) onnistuneesti ennustaa opiskelijoiden opintomenestystä, tulisi mallin (6) olla yhteensopiva opiskelijoiden todelliseen opintomenestykseen. Tällaista yhteyttä ei kuitenkaan havaita, päinvastoin, percentiilimuuttujan kerroin on muihin ylioppilastutkintomuuttujiin verrattuna hyvin pieni. Mallin keskimääräinen neliövirhe sen sijaan on muita malleja suurempi ja  $R^2$  pienempi, mikä osoittaa, että opintopistekertymien ennustaminen ei mallilla ole mielekäästä. Opintopistekertymillä ei näiden tulosten pohjalta ole ennustettavaa yhteyttä yksittäisiin ylioppilastutkintoarvosanoihin tai opiskelijoiden valintakoetuloksiin: voidaan todeta, että opintopisteiden määrää ei ole mahdollista luotettavasti ennakoita käytettyjen valintakriteerien perusteella.

Yksittäisten ylioppilastutkinnon arvosanaryhmien tarkastelu vahvistaa päätelmää: alemman arvosanan kirjoittaneet eivät opintopistekertymien osalta eroa tilastollisesti merkitsevästi laudaturin kirjoittaneiden joukosta kuin pitkän kielen aineryhmässä<sup>17</sup>. Esimerkiksi äidinkielen arvosanaryhmien välillä kerrointen merkki ja koko vaihtelevat ennustamattomasti arvosanan laskiessa laudaturista alaspäin niin, että suurimmat

---

<sup>17</sup> *Indikaattorimuuttujaa käytettäessä jokaiselle arvosanalle luodaan oma dummy-muuttuja, ja tuloksia tulkitessa kutakin dummya verrataan yhteen poisjätettyyn kategoriaan, tässä laudaturiin. Näin esimerkiksi esimian kerrointa tulkitaan esimian ja laudaturin kirjoittaneiden opiskelijaryhmien välisenä erona. Sekä pitkän että lyhyen matematiikan arvosanoilla vertailuryhmänä toimii pitkän matematiikan laudatur. Myös kyseisen aineryhmän ylioppilaskokeen kirjoittamatta jättäneet tai alle B:n kirjoittaneet on koodattu omaksi luokakseen, joka on taulukossa ryhmä "muut". Indikaattorimuuttujien tulkintaa on käsitelty tarkemmin luvussa 3.3.1.*

keskimääräiset kertymäärät ovat hienolla erolla esimian kirjoittaneilla opiskelijoilla<sup>18</sup>. Pitkän matematiikan kerrointen kokoluokka sen sijaan on kaikilla ryhmillä sama, eikä poikkeama lyhyen matematiikan kirjoittaneisiin ole sekään suuri kuin yhden poikkeavan arvosanaryhmän osalta. Tämän arvosanaryhmän C sekä reaaliaineiden ja pitkän kielen yksittäisten, merkitsevien aineryhmien osalta on syytä epäillä ylisovittamista.

Taulukko 7. Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet 3 vuoden opintopistekertymän OLS regressiomallissa (1)

### Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet

3 vuoden opintopistekertymän OLS regressiomallissa (1)

	Äidinkieli	Pitkä matemat.	Lyhyt matemat.	Reaaliaineet	Pitkä kieli	Lyhyt kieli
B	-6.9629 (5.8882)	6.7487 (6.0331)	-4.4237 (10.647)	-7.8957 (6.1894)	7.2011 (5.5373)	-16.7438*** (6.1338)
C	-2.9030 (4.0399)	6.4601 (5.0970)	16.5603** (7.3858)	-9.6671** (4.2504)	10.2896** (4.3240)	-14.9362*** (4.3845)
M	.0439 (3.5970)	6.0379 (4.9328)	5.4223 (6.1930)	-1.1611 (3.3399)	9.9215*** (3.2930)	-8.8292** (3.7206)
E	.9343 (3.4692)	5.8743 (4.8927)	4.1964 (5.7651)	-2.6407 (2.8017)	4.3673 (2.8795)	-7.6168** (3.1865)
L	x	x	4.2364 (5.4550)	x	x	x
Muut	11.4022 (12.715)	6.0979 (6.4715)		-4.6358 (4.9864)	-3.095 (10.369)	-15.870*** (3.6190)

Kategoria x on indikaattorimuuttujan vertailuryhmä, jonka suhteen muita kertoimia tulkitaan. Sekä pitkän että lyhyen matematiikan arvosanoja verrataan pitkän matematiikan L-kategoriaan. Taulukko raportoi indikaattorimuuttujan ryhmäkohtaisen  $\beta$ -kertoimen. Suluissa kertoimien robustit keskivirheet. Kertoimet ovat epämerkitseviä (x) tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa t-testissä.

Ainoastaan lyhyellä kielellä on opiskelijoiden opintopistekertymiin regressiomallissa kokonaisuudessaan merkitsevä vaikutus. Lyhyen kielen kertoimet kertovat, että laudaturin lyhyestä kielestä kirjoittaneet ovat edenneet opinnoissaan keskimäärin muita opiskelijoita

<sup>18</sup> Ryhmä ”muut” jätetään analyysissä huomiotta, sillä se koostuu äidinkielen aineryhmässä muutaman I:n tai A:n kirjoittaneen opiskelijan joukosta, joka jää tilastolliseen analyysiin turhan pieneksi.

nopeammin<sup>19</sup>. Eximian kirjoittaneisiin verrattuna laudaturin opiskelijat ovat suorittaneet keskimäärin 8 opintopistettä enemmän kolmen opintovuoden aikana. Laskennallisesti tämä vastaa noin 210 työtuntia, kun 1 opintopisteen määritelmällinen pituus on 26,7 tuntia (Valtioneuvoston kanslia 2004); kolmen vuoden tavoitellusta 180 opintopisteestä osuus on noin 4 %. Laudaturin ja B:n kirjoittaneiden välillä ero on suurempi: näiden opiskelijoiden keskimääräinen ero kolmen vuoden opiskelun aikana on 17 opintopistettä tai 9 % tutkintovaatimuksista.

Kokonaisuudessaan ylioppilastutkintoarvosanojen yhteys opintopistekertymiin on pieni, mutta valintakoetta voimakkaampi: esimerkiksi taulukon 8 mallissa (5) lähtöpisteytyksen perusteella lasketun persentiilin kerroin on lähes 1,4-kertainen valintakoepersentiilin kertoimeen suhteutettuna. Opiskelijoiden skaalattuun paremmuusjärjestykseen perustuvien muuttujien kertoimet ovat suoraan vertailukelpoiset keskenään. Lähtöpisteytyspersentiilin kerroin 0,15 mallissa (5) vertautuu siihen, että 20 sijaa korkeammat lähtöpisteet saanut opiskelija olisi suorittanut kolmen vuoden aikana noin 3 opintopistettä enemmän kuin matalammalla sijalla oleva opiskelija. Suurimmillaan opiskelijoiden välille saatava ero – liikuttaessa persentiilin sijalta 0 sijalle 100 – on siis noin 15 opintopistettä. Käytännössä kyse on kahden kuuden ja yhden kolmen opintopisteen kurssin laajuisesta erosta, jonka keskimääräinen opiskelija suorittaa suositusajan mukaisesti yhden opetusperiodin aikana<sup>20</sup>. Ero ei siis ole valtava, mutta efektiivinen: mikäli opintopisteiden kertymätahdin ero pysyy

---

<sup>19</sup> *Lyhyen kielen merkitsevä ero selittyy intuitiivisesti aineen ylioppilaskokeen ominaisuuksilla: lyhyt kieli on usein oman osaamisen vuoksi ylimääräisenä kirjoitettava aine, esimerkiksi toinen kotimainen kieli tai muiden kieliopinointojen näyttökoe, jonka kirjoittaja todennäköisesti on keskimääräistä, ainoastaan pakolliset kokeet suorittavaa lukio-opiskelijaa kyvykkäämpi tai ainakin halukkaampi osoittamaan kyvykkyytensä (Kupiainen et al. 2016). Rajatun aineiston analyysissä jää epäselväksi, liittyykö lyhyen kielen merkitsevyys opintomenestyksen ennustamisessa nimenomaan kielen kokeen mittaamiin ominaisuuksiin, vai pikemminkin kokeen statukseen ylimääräisenä, vapaaehtoisena kokeena. Tulosten pohjalta ei siis voida suoraan päätellä, että lyhyen kielen laudaturia tulevissa opiskelijavalinoissa painottamalla Kauppakorkeakouluun saataisiin opinnoissaan nopeammin eteneviä opiskelijoita. Sen sijaan voidaan sanoa, että lyhyen kielen kokeen arvosanoilla vaikuttaa olevan yhteys opintopistekertymiin tässä aineistossa. Muiden ylioppilastutkintoaineiden osalta vastaavaa päätelmää ei voida tehdä.*

<sup>20</sup> *Aalto-yliopiston kauppakorkeakoulussa lukuvuosi on nykyisin jaettu viiteen opetusperiodiin, jolloin tasaisesti jaettuna 60 opintopisteen vuosittainen suositus merkitsee noin 12 opintopisteen suorittamista per opetusperiodi. Aineiston opiskelijoiden opiskeluhetkellä opetusperiodeja on kuitenkin ollut vain neljä, jolloin opintopisteitä on ohjeellisesti tullut suorittaa 15 per periodi (Helsingin kauppakorkeakoulun opintoasiaintoimisto 2009).*

samankaltaisena myös kolmannen opintovuoden jälkeen, voi nopeammin etenevien opiskelijoiden odottaa valmistuvan muutamia kuukausia (kaksi opetusperiodia) vertailuryhmää nopeammin.

Taulukko 8. OLS regressiotulokset: 3 vuoden opintopistekertymä

### OLS regressiotulokset: 3 vuoden opintopistekertymä

	(1)	(2)	(3)	(4)	(5)	(6)
Ylioppilastutkinnon indikaattorimuuttujat	x		x			
Valintakokeen persentiili		.0792** (.0366)	.1119*** (.0380)	.1055*** (.0376)	.1084*** (.0375)	.0820** (.0372)
Yo-arvosanojen keskiarvopersentiili				.1303*** (.0410)		
Yo-arvosanojen lähtöpisteytyspersentiili					.1481*** (.0410)	
Yo-indikaattorimuuttujien ennustepersentiili						.0680 (.0415)
Vuosikurssi	x	*	x	*	*	**
Tutkintovuosi	***		***	***	***	***
Vakio	135.6713*** (6.6299)	149.4796*** (3.4086)	134.2288*** (6.6198)	129.7428*** (3.2010)	129.5421*** (2.8810)	132.9908*** (4.2338)
Havaintoja	2270	2270	2270	2270	2270	2270
R <sup>2</sup>	.0775	.0120	.0814	.0613	.0628	.0576
Korjattu R <sup>2</sup>	.0490	.0085	.0526	.0445	.0460	.0407
Keskimääräinen neliövirhe	47.896	48.906	47.806	48.010	47.971	48.105

Taulukko raportoi jatkuvan muuttujan  $\beta$ -kertoimen tai indikaattorimuuttujan kokonaisvaikutuksen merkitsevyysasteen. Suluissa kertoimien robustit keskivirheet. Kertoimet ovat epämerkitseviä (x) tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa t-testissä. Malli ei sisällä muita kuin tässä mainittuja kontrolleja.

Valintakoe ei tuo opintopistekertymien ennustamiselle lisäarvoa. Valintakoe itsessään voidaan yhdistää vain prosenttiin opintopistekertymien vaihtelusta (malli (2)) ja ylioppilastutkinnon indikaattorimuuttujiin yhdistettynä valintakoe ei tuo selitysarvoa tämänkään vertaa lisää (malli (3)). On huomattava, että ylioppilastutkintoarvosanojen lisääminen regressioon malleissa (3) – (5) kasvattaa valintakoepersentiilin kerrointa, mikä on loogista, kun muistetaan näiden kahden välinen negatiivinen korrelaatio. Mallien (2) ja (6) valintakokeen kerroin lieneekin tutkielman tarkoituksiin nähden harhainen, kun ylioppilastutkinnon mittaaman kyvykkyyden merkitystä koepisteiden muodostumiseen ei kontrolloida. Mallien (3) – (5) valintakokeen kerroin 0,11 jää kuitenkin myös hyvin pieneksi, ja suurimmillaan valintakoe persentiilin ääripäiden välinen ero jää 11 opintopisteseen.

Voidaan todeta, että valintakoepisteiden ja kolmen vuoden opintopistekertymien välillä on kevyt merkitsevä positiivinen yhteys, jonka voimakkuus ei kuitenkaan riitä opintomenestyksen ennustamiseen.

Tulosten valossa vaikuttaa kokonaisuudessaan siltä, että opiskelijoiden kolmen vuoden opintopistekertymiä ei voida ennustaa valintahetkellä tiedossa olevien mittareiden perusteella.

#### 4.2.2 Kolmen vuoden arvosanakeskiarvo

Kolmen vuoden arvosanakeskiarvojen regressiotulokset on esitelty indikaattorimuuttujien kerrointen osalta taulukossa 9 ja mallien (1) – (6) osalta taulukossa 10. Arvosanakeskiarvojen ennustaminen ylioppilastutkintoarvosanojen perusteella vaikuttaa varsin mielekkäältä. Indikaattorimuuttujina käsiteltyinä ylioppilastutkintoarvosanat onnistuvat jo yksinään selittämään reilut 24 % aineiston opiskelijoiden arvosanakeskiarvojen vaihtelusta. Yhdessä valintakokeen kanssa valintakriteerien selitysarvo on noin 26 %. Yksittäisten ylioppilastutkinnon arvosanaryhmien välillä onkin havaittavissa pääosin ennustettavia, tilastollisesti merkitseviä eroja. Muutaman (mahdollisesti ylisovitetun) arvosanaryhmän poikkeusta lukuun ottamatta opiskelijoiden keskiarvoiset yliopistoaikaiset arvosanat laskevat jatkuvasti ylioppilastutkintoarvosanojen laskiessa. Voimakkain yhteys arvosanakeskiarvoihin vaikuttaa olevan matematiikan, heikoin pitkän kielen ja äidinkielen arvosanoilla. Pitkän kielen arvosanaryhmien L, E tai M välillä ei ole tilastollisesti merkitsevää eroa, mutta seuraavien arvosanaryhmien kohdalla ero on selkeästi laskeva. Äidinkielen arvosanaryhmien välille ei ole tehtävissä tilastollisesti selkeää yhteyttä, ja arvosanan C poikkeava kerroin lienee seurausta ylisovittamisesta. Reaaliaineiden ja lyhyen kielen arvosanojen yhteys arvosanakeskiarvoihin on selkeämpi: heikompi ylioppilastutkintomenestys on yhteydessä heikompaan menestymiseen korkeakoulussa.

Matematiikan muuttujalla kaikkien arvosanaryhmien erot ovat 1 % tasolla merkitseviä ja arvosanakeskiarvon yhteys aineen arvosanoihin on voimakkaasti positiivinen. Jo pitkän matematiikan laudaturin ja esimian kirjoittaneiden välillä arvosanakeskiarvot eroavat opiskelijoiden välillä keskimäärin 0,12 yksikköä, kun matalimpien arvosanaryhmien ja

pitkän matematiikan laudaturin välillä keskimääräinen ero opiskelijoiden välillä on jo puolen arvosanan luokkaa (ero esimerkiksi keskiarvojen 3,0 ja 3,5 välillä). Kenties mielenkiintoisin kerroin on lyhyen matematiikan laudaturin -0,28, joka on samaa kokoluokkaa pitkästä matematiikasta M tai C arvosanan kirjoittaneiden opiskelijoiden kanssa. Tulos antaa lähtökohdan pitkän ja lyhyen matematiikan vertailukelpoisen pisteitysasteikon luomiselle. Lyhyen matematiikan arvosanaryhmien E, M tai C kertoimet ovat hyvin saman suuruiset, eikä näiden keskiarvoisten kirjoittajien välillä ole siis havaittavaa eroa arvosanakeskiarvoissa. Kokonaisuudessaan ylioppilastutkintoarvosanojen indikaattorimuuttujilla vaikuttaa olevan vahva ennustettava yhteys opiskelijoiden arvosanakeskiarvoihin.

Taulukko 9. Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet 3 vuoden arvosanakeskiarvon OLS regressiomallissa (1)

### Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet

3 vuoden arvosanakeskiarvon OLS regressiomallissa (1)

	Äidinkieli	Pitkä matemat.	Lyhyt matemat.	Reaaliaineet	Pitkä kieli	Lyhyt kieli
B	-.1000 (.0680)	<b>-.4087***</b> (.0556)	<b>-.5906***</b> (.0993)	<b>-.2895***</b> (.0703)	<b>-.1806***</b> (.0642)	<b>-.2678***</b> (.0646)
C	<b>-.1190***</b> (.0427)	<b>-.3007***</b> (.0492)	<b>-.3954***</b> (.0810)	<b>-.2074***</b> (.0445)	<b>-.1645***</b> (.0438)	<b>-.1951***</b> (.0436)
M	-.0545 (.0362)	<b>-.2559***</b> (.0463)	<b>-.4148***</b> (.0622)	<b>-.1744***</b> (.0364)	-.0193 (.0353)	<b>-.1956***</b> (.0370)
E	-.0076 (.0338)	<b>-.1249***</b> (.0439)	<b>-.4014***</b> (.0530)	<b>-.1501***</b> (.0285)	-.0361 (.0311)	<b>-.1261***</b> (.0321)
L	x	x	<b>-.2775***</b> (.0514)	x	x	x
Muut	<b>.2653***</b> (.0942)	<b>-.4722***</b> (.0617)		<b>-.1925***</b> (.0567)	-.0867 (.1432)	<b>-.1940***</b> (.0380)

Kategoria x on indikaattorimuuttujan vertailuryhmä, jonka suhteen muita kertoimia tulkitaan. Sekä pitkän että lyhyen matematiikan arvosanoja verrataan pitkän matematiikan L-kategoriaan. Taulukko raportoi indikaattorimuuttujan ryhmäkohtaisen  $\beta$ -kertoimen. Suluissa kertoimien robustit keskivirheet. Kertoimet ovat epämerkitseviä (x) tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa t-testissä.

Taulukko 10. OLS regressiotulokset: 3 vuoden arvosanakeskiarvo

## OLS regressiotulokset: 3 vuoden arvosanakeskiarvo

	(1)	(2)	(3)	(4)	(5)	(6)
Ylioppilastutkinnon indikaattorimuuttajat	***		***			
Valintakokeen persentiili		.0008*	.0030***	.0033***	.0032***	.0027***
		(.0004)	(.0004)	(.0004)	(.0004)	(.0004)
Yo-arvosanojen keskiarvopersentiili				.0092***		
				(.0004)		
Yo-arvosanojen lähtöpisteytyspersentiili					.0093***	
					(.0398)	
Yo-indikaattorimuuttajien ennustepersentiili						.0092***
						(.0004)
Vuosikurssi	**	*	**	x	**	x
Tutkintovuosi	***		***	***	***	***
Vakio	3.2160***	3.5098***	3.1768***	1.9892***	2.0519***	1.9908***
	(.0653)	(.0358)	(.0659)	(.0323)	(.0281)	(.0331)
Havaintoja	2263	2263	2263	2263	2263	2263
R <sup>2</sup>	.2413	.0076	.2623	.2212	.2308	.2157
Korjattu R <sup>2</sup>	.2184	.0040	.2391	.2072	.2170	.2016
Keskimääräinen neliövirhe	.5002	.5646	.4935	.5038	.5006	.5055

Taulukko raportoi jatkuvan muuttujan  $\beta$ -kertoimen tai indikaattorimuuttujan kokonaisvaikutuksen merkitsevyyssasteen. Suluissa kertoimien robustit keskivirheet. Kertoimet ovat epämerkitseviä (x) tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyydellä kaksisuuntaisessa t-testissä. Malli ei sisällä muita kuin tässä mainittuja kontrolleja.

Ylioppilastutkintomuuttajat toimivat hyvänä arvosanakeskiarvojen ennustajana myös muussa muodossa kuin indikaattorimuuttujina: Ylioppilastutkintoarvosanamuuttajien kerroin on tuhannesosan tarkkuudella saman suuruinen persentiilin muodostamisen menetelmästä riippumatta. Mallien (4) – (6) kerroin 0,009 tarkoittaa käytännössä sitä, että opiskelijoiden keskimääräiset arvosanakeskiarvot eroavat persentiilin ääripäissä, sijaluvulta 0 sijaluvulle 100 siirryttäessä, noin 0,9 yksikköä. Prosentuaalisesti ero on noin 25 % arvosanakeskiarvojen havaitusta vaihtelusta ( $0,9/3,60 = 0,25$ ). Valintakoeopersentiilin kertoimeen verrattuna ylioppilastutkintoarvosanojen kerroin on kolminkertainen. Out-of-sample regressiolla luotu malli (6) toimii ylioppilastutkintoarvosanojen indikaattorimuuttajien mallin (1) ennustearvon validointina: indikaattorimuuttajien perusteella laskettujen ennusteiden paremmuusjärjestys sopii opiskelijoiden todelliseen opintomenestykseen yhtä hyvin muiden mallien kanssa. Validointi osoittaa, että kolmen vuoden arvosanakeskiarvojen ennustaminen ylioppilastutkintoarvosanojen perusteella on aineiston opiskelijoilla mielekästä. On myös huomattavaa, että ylioppilastutkinnon informaatiota eri tavalla käsittelevien mallien (4) – (6) välillä selitysarvo on korkeimmillaan lähtöpisteytystä hyödyntävässä mallissa (5) – vaikka



ero ei ole suuri, on tämä positiivinen signaali siitä, että ylioppilastutkintoarvosanojen painottaminen hyvin määritellyllä pisteytysmallilla on opiskelijavalinnan kannalta suotavaa.

Valintakoe sen sijaan ei tuo arvosanakeskiarvojen ennustamiseen juuri lisäarvoa: yksinään valintakoe ennustaa reilusti alle prosentin arvosanojen vaihtelusta (malli (2)), ja ylioppilastutkinnon indikaattorimuuttujiin lisättyinä valintakoe kasvattaa mallin selitystasetta vain noin 2 prosenttiyksikköä (malli (3)). Valintakoepersentiilin kerroin on mallin tulosten mukaan 1 % tasolla merkitsevä, mutta kooltaan varsin pieni. Mallien (3) – (6) valintakokeen kerroin 0,003 voidaan tulkita niin, että sadan sijan nousussa matalimmalta korkeimmalle sijalle arvosanakeskiarvon nousu olisi 0,30 yksikköä tai noin 8 % opiskelijoiden keskiarvon vaihtelusta ( $100 \cdot 0,0030 / 3,60 = 0,083$ ). Voidaan todeta, että valintakoepisteiden ja kolmen vuoden arvosanakeskiarvon välillä on hyvin pieni positiivinen yhteys. Suuruudeltaan yhteys on lähes mitätön, eikä valintakoe ole arvosanakeskiarvojen ennustamiselle tärkeä muuttuja.

Kokonaisuudessaan opiskelijoiden kolmen opintovuoden arvosanakeskiarvojen ennustaminen Kauppakorkeakoulussa vaikuttaa kuitenkin olevan melko mielekäästä.

#### 4.2.3 Kaikille yhteisten opintojen arvosanakeskiarvo

Kaikille yhteisten opintojen arvosanakeskiarvojen regressiotulokset on esitelty indikaattorimuuttujien kerrointen osalta taulukossa 11 ja mallien (1) – (6) osalta taulukossa 12. Tulokset ovat laadullisesti samankaltaiset kolmen vuoden arvosanakeskiarvojen kanssa, mutta eroavaisuuksiakin löytyy. Arvosanakeskiarvojen ennustaminen vaikuttaa siis varsin mielekkäältä muuttujan määrittelystä riippumatta. Indikaattorimuuttujien mallien (1) ja (3) selitystasot 24 % ja 27 % ovat molemmilla arvosanamuuttujilla lähes identtiset, mutta kaikille yhteisten opintojen mallien keskimääräinen neliövirhe on kolmen vuoden malleja lähes 40 % korkeampi. Sama havainto keskivirheen suuruudesta on todettavissa myös malleille (2) ja (4) – (6). Neliövirheiden erot selittynevät arvosanamuuttujien eroavilla vaihtelutasoilla: kolmen vuoden arvosanakeskiarvon muuttujan varianssi on matalampi kuin yhteisten opintojen muuttujalla (taulukko 4). Opiskelijoiden väliset erot ovat siis kokonaisuudessaan hieman suurempia yhteisten opintojen arvosanakeskiarvoa tarkasteltaessa. Opintojen edetessä opiskelijoiden väliset erot pienenevät ja keskimääräiset

arvosanakeskiarvot kasvavat: reilulla 60 % opiskelijoista kolmen vuoden arvosanakeskiarvo on yhteisten opintojen muuttujaa korkeampi. Tämä näkyy myös ennustekertoimissa, jotka ovat kauttaaltaan kaikille yhteisten opintojen malleissa aiempaa korkeampia. Selitysasteiden suuruudessa on huomattavaa, että mallien (4) ja (5) selitysasteet ovat kaikille yhteisten opintojen malleissa reilusti muita malleja matalammat. Mallin (6) korkea selitysaste ja ylioppilastutkinnon ennustepersentiilin kerroin sen sijaan osoittavat indikaattorimuuttujien soveltuvan varsin hyvin aineiston opiskelijoiden arvosanakeskiarvojen ennustamiseen.

Taulukko 11. Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet kaikille yhteisten opintojen arvosanakeskiarvon OLS regressiomallissa (1)

## Ylioppilastutkintoarvosanojen indikaattorimuuttujien kertoimet

Kaikille yhteisten opintojen arvosanakeskiarvon OLS regressiomallissa (1)

	Äidinkieli	Pitkä matemat.	Lyhyt matemat.	Reaaliaineet	Pitkä kieli	Lyhyt kieli
B	.0528 (.0896)	<b>-.7113***</b> (.0773)	<b>-1.104***</b> (.1494)	<b>-.3571***</b> (.0908)	-.0261 (.0852)	<b>-.2159**</b> (.0875)
C	<b>-.0458</b> (.0587)	<b>-.5212***</b> (.0688)	<b>-.7532***</b> (.1074)	<b>-.2963***</b> (.0615)	<b>-.0822</b> (.0597)	<b>-.1871***</b> (.0604)
M	<b>-.0178</b> (.0498)	<b>-.3605***</b> (.0644)	<b>-.7909***</b> (.0894)	<b>-.2274***</b> (.0498)	<b>.1087**</b> (.0493)	<b>-.1621***</b> (.0503)
E	<b>-.0145</b> (.0467)	<b>-.1229**</b> (.0608)	<b>-.7656***</b> (.0766)	<b>-.2054***</b> (.0391)	<b>.0255</b> (.0422)	<b>-.1511***</b> (.0444)
L	x	x	<b>-.5621***</b> (.0727)	x	x	x
Muut	<b>.4491***</b> (.1445)	<b>-.8168***</b> (.0845)		<b>-.3203***</b> (.0776)	<b>.1509</b> (.1822)	<b>-.1202**</b> (.0514)

Kategoria x on indikaattorimuuttujan vertailuryhmä, jonka suhteen muita kertoimia tulkitaan. Sekä pitkän että lyhyen matematiikan arvosanoja verrataan pitkän matematiikan L-kategoriaan. Taulukko raportoi indikaattorimuuttujan ryhmäkohtaisen  $\beta$ -kertoimen. Suluissa kertoimien robustit keskivirheet. Kertoimet ovat epämerkittäviä (x) tai merkittäviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyydellä kaksisuuntaisessa t-testissä.

Ylioppilastutkintoarvosanojen indikaattorimuuttujat käyttäytyvät kaikille yhteisten opintojen mallissa aiemmasta luvusta tutulla tavalla, mutta kertoimien kokoluokka on nyt lähes poikkeuksetta suurempi. Kun kolmen vuoden arvosanakeskiarvoilla suurin ero vahvimmin erottelevan matematiikan muuttujan ääripäiden välillä oli -0,59, on vastaava kerroin nyt -1,1 eli reilut 86 % korkeampi. Etenkin matemaattisen osaamisen vaikutus arvosanoihin vaikuttaakin korostuvan kaikille yhteisillä kursseilla: matematiikan arvosanaryhmien väliset erot ovat yhteisissä opinnoissa kolmen vuoden muuttujaa selkeästi

suurempia. Lyhyen matematiikan kirjoittajien opintomenestys suhteessa pitkän matematiikan kirjoittajiin vaikuttaa ensimmäisen vuoden kursseilla olevan seuraavia vuosia heikompaa: lyhyestä matematiikasta laudaturin kirjoittaneet opiskelijat asettuvat nyt pitkästä matematiikasta C:n kirjoittaneiden alapuolelle, kun kolmen vuoden muuttujalla vertailukohta löytyi M:n ja C:n väliltä. Myös lyhyen kielen ja reaaliaineiden arvosanaryhmien kertoimet ovat kolmen vuoden mallia suuremmat, mutta niiden väliset erot ovat edelleen hyvin samankaltaisia. Äidinkielen ja pitkän kielen arvosanaryhmien väliset erot sen sijaan vaikuttavat nyt lähinnä satunnaiselta, eikä ylisovittamisen mahdollisuutta mallissa voida sivuuttaa.

Taulukko 12. OLS regressiotulokset: Kaikille yhteisten opintojen arvosanakeskiarvo

### OLS regressiotulokset: Kaikille yhteisten opintojen arvosanakeskiarvo

	(1)	(2)	(3)	(4)	(5)	(6)
Ylioppilastutkinnon indikaattorimuuttujat	***		***			
Valintakokeen persentiili		.0023*** (.0006)	.0045*** (.0005)	.0050*** (.0005)	.0051*** (.0005)	.0041*** (.0005)
Yo-arvosanojen keskiarvopersentiili				.0101*** (.0006)		
Yo-arvosanojen lähtöpisteytyspersentiili					0.0109*** (.0553)	
Yo-indikaattorimuuttujien ennustepersentiili						.0121*** (.0006)
Vuosikurssi	***	*	***	***	***	x
Tutkintovuosi	***		***	***	***	***
Vakio	3.1156*** (.0916)	3.2245*** (.0497)	3.0588*** (.0909)	1.6757*** (.0448)	1.7003*** (.0391)	2.4303*** (.0077)
Havaintoja	2209	2209	2209	2209	2209	2209
R <sup>2</sup>	.2432	.0252	.2682	.1686	.1903	.2279
Korjattu R <sup>2</sup>	.2199	.0216	.2453	.1540	.1761	.2144
Keskimmääinen neliövirhe	.6797	.7612	.6686	.7078	.6986	.6821

Taulukko raportoi jatkuvan muuttujan  $\beta$ -kertoimen tai indikaattorimuuttujan kokonaisvaikutuksen merkitsevyyssasteen. Suluissa kertoimien robustit keskiarvot. Kertoimet ovat epämerkitseviä (x) tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyydellä kaksisuuntaisessa t-testissä. Malli ei sisällä muita kuin tässä mainittuja kontrolleja.

Ylioppilastutkintoarvosanat toimivat edelleen hyvinä arvosanakeskiarvojen ennustajina myös muussa muodossa kuin indikaattorimuuttujina, vaikka keskiarvo- ja lähtöpisteytyspersentiilien muuttujien kertoimet eivät ole kasvaneet samassa suhteessa

muiden mallien muuttujien kanssa. Mallien (4) ja (5) kerroin 0,01 vertautuu 1 arvosanayksikön nousuun persentiilin ääripäiden välillä; tämä on 29 % arvosanojen vaihtelusta ( $1/3,43 = 0,292$ ). Vaikuttaakin siltä, että kaikille yhteisten opintojen arvosanakeskiarvojen ennustamisessa indikaattorimuuttujat ovat muita järjestysmenettelyjä toimivampia: mallin (6) selitysaste on malleja (4) ja (5) selkeästi korkeampi, ja ennustepersentiilin kerroin poikkeaa muista menetelmistä positiiviseen suuntaan. Out-of-sample validointi osoittaa, että arvosanakeskiarvojen ennustaminen ylioppilastutkintoarvosanojen perusteella on aineiston opiskelijoilla mielekästä.

Valintakokeen yhteys arvosanakeskiarvoihin on nyt huomattavasti aiempaa selkeämpi: kaikille yhteisten opintojen ennustemalleissa valintakoe persentiilin kerroin on noin 1,5-kertainen kolmen vuoden malleihin verrattuna. Yhteys on tästä huolimatta hyvin pieni; ylioppilastutkintomuuttujien kertoimet ovat edelleen suhteessa kaksinkertaisia. Yksinään valintakoe kattaa nyt reilut 2 % arvosanakeskiarvojen vaihtelusta (malli (2)), mutta selitysasteen lisäys ylioppilastutkinnon indikaattorimuuttujiin yhdistettynä on kasvusta huolimatta vain 2 prosenttiyksikön luokkaa. Valintakoe persentiilin kerroin 0,005 vertautuu noin 0,5 arvosanayksikön vaihteluun persentiilin ensimmäisen ja viimeisen sijan välillä; aineiston opiskelijoiden arvosanakeskiarvon vaihtelusta tämä on vajaat 15 % ( $0,5/3,43 = 0,146$ ). Ero kolmen vuoden arvosanakeskiarvon vastaavalla tavalla laskettuun prosenttiin on noin 7 prosenttiyksikköä tai 88 %. Vaikuttaakin siltä, että valintakoe pisteiden yhteys arvosanoihin on ensimmäisenä opintovuonna korkeampi kuin sitä seuraavina vuosina, mutta molemmissa huomattavan pieni. Valintakoe ei näiden tulosten valossa tuo opintomenestyksen ennustamiseen merkittävää lisäarvoa.

Yhteisten opintojen arvosanakeskiarvon ennustettava yhteys mallin muuttujiin on hyvin samankaltainen pidemmän aikavälin arvosanakeskiarvon kanssa. Yhteys selittäviin muuttujiin on kuitenkin kauttaaltaan voimakkaampi. Kokonaisuudessaan arvosanakeskiarvojen ennustaminen vaikuttaa Aalto-yliopiston kauppakorkeakoulussa mielekkäältä.

### 4.3 Analyysi

Regressiotulosten perusteella vaikuttaa siltä, että ylioppilastutkintoarvosanat ennustavat kurssiarvosanojen muodostumista varsin mielekkäästi, mutta opintojen etenemisen ennustamiseen ei ylioppilastutkintoarvosanoista ole. Arvosanojen sopivimman ennustemallin 26 % selitysaste vastaa aiheen parissa tehdyn aiemman tutkimuksen tuloksia (luku 2.3), mutta opintopistekertymien 8 % selitysaste on aiempia tutkimuksia korkeampi. Arvosanakeskiarvojen regressiotulosten mukaan ylioppilastutkintoarvosanojen yhteys opinnoissa menestymiseen on varsin saman kaltainen opintovuosien ajan, kuitenkin niin, että ylioppilastutkinnon perusteella ennustettavat erot ovat voimakkaimmillaan opintojen alussa. Etenkin matemaattinen osaaminen vaikuttaa korostuvan opinnoissa menestymisen indikaattorina, mikä vahvistaa aiempien tutkimusten havaintoja (mm. Erkki Komulainen 1987 – 1996, Mieskonen 2016). Ylioppilastutkinnon indikaattorimuuttujien havaittu merkityksellisyys on linjassa samalla aineistolla aiemmin saatujen tulosten kanssa (Mieskonen 2016).

Valintakoe ei analyysin mukaan tuo ennustamiseen huomattavaa lisäarvoa arvosanakeskiarvojen tai opintopistekertymien regressioissa. Valintakoepercenttiin ja lähtöpisteytyspercenttiin havaittua merkitystä voi verrata samalla lailla laskettuja muuttujia hyödyntäneen Pietilän (2012) tuloksiin. Pietilän analyysin mukaan valintakoepercenttiin yhden sijan nousu nostaa kolmen vuoden keskimääräisiä opintopistekertymiä 0,43 yksikköä (nyt 0,11) ja arvosanakeskiarvoja 0,011 yksikköä (nyt 0,003). Valintakoepercenttiin kerroin on siis molemmilla muuttujilla Pietilän analyysissä nelinkertainen<sup>21</sup>. Lähtöpisteytyksen vaikutus opintopistekertymiin on Pietilän tuloksissa 0,26 (nyt 0,15) ja arvosanakeskiarvoihin 0,1 (nyt 0,1). Tulosten mukaan ylioppilastutkintoarvosanojen painottaminen hyvin

---

<sup>21</sup> Kertoimien suuruuserolle on vaikea löytää mielekästä selitystä. Tutkielman regressiomalli (5) on lähes identtinen Pietilän käyttämän ennustemallin kanssa. Ainoa lisätty muuttuja on ylioppilastutkinnon suoritusvuosi. Aineiston koko on kasvanut muutamalla opiskelijavuosikurssilla, mutta vuosikurssikohtainen vaihtelu on molemmissa analyyseissa vakioitu kontrollimuuttujalla. Lähtöpisteytykset on tässä laskettu nykyisen valintamenettelyn mukaisesti, kun Pietilän analyysissä on käytössä vanha pisteytys. Ero näiden välillä ei kuitenkaan ole suuri. Percenttimuuttujat on teknisesti laskettu hieman eri tavoin, mutta ovat muodoltaan ja funktioiltaan samanlaiset. Virheen mahdollisuutta ei voida sulkea pois.

määritellyllä pisteytysmallilla vaikuttaa olevan opintomenestyksen ennustamisen kannalta suotavaa, vaikka erot yksinkertaisempiin malleihin eivät ole suuria.

#### 4.3.1 Vaihtoehtoinen ylioppilastutkintoarvosanojen pisteytys

Tutkielman yhtenä tavoitteena on ollut esitellä ylioppilastutkintoarvosanojen ennustemerkityksellisyyteen perustuva valintapisteytysmalli, joka kuvataan taulukossa 13. Pisteytys on muodostettu kahdessa osassa: ensin eri aineiden laudatur -arvosanojen keskinäinen asetelma on selvitetty laudaturin kirjoittaneiden opiskelijoiden keskimääräisen arvosanakeskiarvon perusteella (liite 5), sitten muiden arvosanaluokkien etäisyys laudaturista on laskettu taulukoissa 9 ja 11 raportoitujen indikaattorimuuttujien kertoimien mukaan. Ylioppilastutkintoarvosanojen merkitystä opintojen etenemiseen ei pisteytyksessä ole huomioitu, sillä näiden välinen yhteys on regressiotulosten mukaan pieni ja pääosin epämerkittävä. Ristiriitatilanteissa on painotettu yhteisten opintojen arvosanakeskiarvon regressiotuloksia mallin korkeamman erottelevuuden ja suoraan vertailukelpoisten kurssisisältöjen vuoksi. Lopputuloksena on ylioppilastutkintopisteytys, joka perustuu suoraan ylioppilasarvosanojen ja opintomenestyksen yhteyteen. Pisteytyksessä eniten pisteitä antavat arvosanat ovat yhteydessä parhaaseen menestykseen kauppatieteellisissä opinnoissa, ja vähiten pisteitä antavat arvosanat esiintyvät useimmin aineiston keskiarvoisilla opiskelijoilla. Ennustemallien tulosten johtaminen pisteytysjärjestykseen on tässä esitettyjen periaatteiden nojalla yksiselitteistä, mutta itse pisteytyksen voi tehdä usealla tavalla valitsijan preferensseistä riippuen.

Vaihtoehtoinen pisteytysmalli on tehty samanmuotoiseksi kuin nyt käytössä oleva lähtöpisteiden laskentamalli (liite 1). Pisteytysmallin mukaan valintapistettä jaetaan 4 – 10 maksimissaan viidestä ylioppilaskokeesta. Määritelmällistä maksimipistemäärää ei kuitenkaan ole, vaan pisteet määräytyvät suoraan taulukon osoittamien arvojen mukaan. Matematiikan eri laajuuksien kokeet ovat toisensa poissulkevia, eli pisteitä voi saada vain toisesta, paremmat pisteet antavasta suorituksesta. Aiemmissä luvuissa (4.2.2 ja 4.2.3) todettiin, että lyhyen matematiikan laudaturin kirjoittaja vertautuu pitkästä matematiikasta C:n kirjoittaneeseen opiskelijaan, ja tämä vertaus pätee pisteytysmallissakin. Matemaattinen osaaminen vaikuttaa korostuvan kauppatieteellisissä opinnoissa menestymisessä etenkin

alkuvaiheen, mutta myös myöhemmän vaiheen opinnoissa, ja se arvotetaan myös nykyisessä opiskelijavalinnassa korkealle (liite 1). Regressiotuloksiin pohjautuva vaihtoehtoinen pisteytysmalli kannustaa opiskelijoita lukemaan lukiossa pitkän matematiikan laajan oppimäärän, vaikka lopullinen ylioppilaskirjoitustulos ei olisikaan skaalan yläpäässä. Tällaisella pisteytyksellä voisi olla Opetus- ja kulttuuriministeriön (2016b) tarkoittaman mukainen ohjaava vaikutus lukiolaisiin: lukioaikaisiin oppiainevalintoihin voi ja kannattaa pyrkiä vaikuttamaan, mikäli tietyn osaamisen merkitystä tulevilla yliopisto-opinnoissa halutaan korostaa.

Taulukko 13. Ylioppilastutkintoarvosanojen vaihtoehtoinen pisteytys

### Ylioppilastutkintoarvosanojen vaihtoehtoinen pisteytys

Äidinkieli	Pitkä matematiikka	Lyhyt matematiikka	Reaaliaineet	Pitkä kieli	Lyhyt kieli	
L						10
L			L		L	9
E	E			L		8
			E	E	E	7
M	M		M	M	M	6
	C	L				5
C	B	EMC	C	C	C	4

Pisteytys

Pisteytysmallissa on tehty kolme tietoista valintaa koskien pisteytyksen alapään rajausta, aineluokkien sisällyttämistä ja pisteytyksen erottelevuuden kasvattamista. Pisteytys ulottuu pääosin arvosanaan C asti kahdesta pääsyystä: 1) opiskelijoiden erottelemiseen toivotaan saatavan tarpeeksi tasoja ja syvyyttä, kun 60 % ylioppilastutkinnon arvosteluskaalan mahdollisista arvosanoista huomioidaan ja 2) rajausta on tehty niin, että alimman pistetason arvosana ylioppilaskokeesta on regressiotulosten mukaan mielekkäämpi suoritus kuin tämän aineluokan kokeen kirjoittamatta jättäminen. Kaikki aineluokat on päätetty ottaa mukaan kolmesta syystä: 1) heikoiten merkitsevien aineluokkien – äidinkielen ja pitkän kielen – poisjättäminen tuntuu väärältä, kun äidinkielen koe on ainoa kaikille hakijoille pakollinen

koe ja pitkän kielen koe on määritelmällisesti haastavuudeltaan lyhyttä kieltä suurempi<sup>22</sup>, 2) ylioppilasaineiden merkityksellisyyden selvittäminen kannustaa eri aineiden laajaan hyödyntämiseen tutkimusmielessä ja 3) opiskelijavalinnassa huomioidaan myös nykyisen valintamallin mukaan viisi koetta<sup>23</sup>. Lopuksi, täpärästi tasapisteisiin päätyvät arvosanalokat on eroteltu eri pisteluokkiin, jotta pisteytyksen perusteella on mahdollista tehdä opiskelijoiden välille todellisia eroja. Näiden valintojen vaikutus pisteytysmallille on merkittävä.

---

<sup>22</sup> On huomattava, että pitkä kieli käsittää tässä ja aiemmissa pisteytysmalleissa kaksi hyvin eri tyyppistä koetta. Aineiston opiskelijoista reilut 99 % on sisällyttänyt ylioppilastutkintoonsa A-englannin kokeen, joka on kielten kokeista kansallisesti suosituin (Ylioppilastutkintolautakunta 2014) ja helpoimmaksi kielen kokeeksi koettu (Kupiainen & Marjanen 2014). Tämän lisäksi vain vajaat 8 % (174 opiskelijaa) on suorittanut muun A-laajuisen kielen, käytännössä A-venäjän, A-ranskan tai A-saksan, joiden arvosanajakauman on väitetty olevan vääristynyt (alaviite 2). Aineistossa vain 28 opiskelijaa on saanut korkeimman pitkän kielen arvosanansa muusta kuin englannin kielen kokeesta. Lyhyen kielen luonteesta on käyty keskustelua alaviitteessä 18. On siis mahdollista argumentoida, että aineiston opiskelijoilla lyhyt kieli on todellisuudessa koettu opiskelijoiden keskuudessa ”haastavammaksi”, vaikka pitkä kieli määritelmällisesti kattaakin suuremman laajuuden opintoja.

<sup>23</sup> Aineiston opiskelijoista 95 % on suorittanut vähintään viisi ylioppilaskoetta (keskiarvo 5,85, mediaani 6).



## 5 Opiskelijavalinnan vaikutuksen arviointi

Toisen empiirisen vaiheen tarkoituksena on tutkia, kuinka eri ihmiset tulisivat valituksi ja miten opiskelijoiden keskimääräinen opintomenestys muuttuisi, jos Aalto-yliopiston kauppakorkeakoulun opiskelijavalinnan menettelyjä muutettaisiin. Vertailtavina valintamenettelyinä toimivat kauppatieteellisen alan nykyinen valintamenettely sekä puhtaasti ylioppilastutkintoarvosanoihin perustuva vaihtoehtoinen valinta. Toinen vaihe pohjautuu ensimmäisessä empiirisessä vaiheessa (luku 4) luotuun ylioppilastutkintoarvosanojen vaihtoehtoiseen pisteytysmalliin, joka perustuu suoraan ylioppilastutkintoarvosanojen havaittuun merkityksellisyyteen opintomenestyksen ennustamisessa.

Luku pyrkii vastaamaan seuraaviin tutkimuskysymyksiin:

- 3) Kuinka eri ihmiset tulisivat valituksi, jos käytössä olisi pelkästään ylioppilastutkintoon perustuva valinta?
- 4) Miten ylioppilastutkinnon perusteella valitut opiskelijat menestyisivät verrattuna niihin opiskelijoihin, jotka eivät olisi tulleet valituksi pelkkään ylioppilastutkintoon perustuvassa valinnassa?

Analyysin käytännöllisenä tavoitteena on selvittää mikä kauppatieteellisen alan valintakokeen mahdollinen lisäarvo opiskelijavalinnassa on.

### 5.1 Metodologia

Valintamenettelyn vaikutuksen arviointi perustuu tutkielmassa keinotekoisen opiskelijavalinnan mallinnukseen: jotta valintamenettelyjen vaikutusta valittuihin opiskelijoihin ja näiden keskiarvoiseen opintomenestykseen voi arvioida, on ensin selvitettävä, ketkä hakijat tulisivat valituksi eri valintamenettelyillä. Arvioitavia valintamenettelyjä on kaksi: kevästä 2018 alkaen käytössä oleva nykyinen valintamenettely (liite 1) ja vaihtoehtoinen, luvussa 4.3 luotuun pisteytysmalliin perustuva puhdas todistusvalinta. Valintamenettelyjen tarkemmat yksityiskohdat löytyvät liitteestä 6.

Keinotekoisen opiskelijavalinnan suorittamista varten aineiston opiskelijoiden ajatellaan koostavan hakijajoukon, joiden keskuudesta uudet opiskelijat valitaan vaihtoehtoisilla

valintamenettelyillä. Näin toimitaan, koska tutkielman käytössä oleva aineisto sisältää ainoastaan Aalto-yliopiston kauppakorkeakouluun hyväksytyt ja paikan vastaanottaneet opiskelijat. Muista potentiaalisista hakijoista ei ole dataa. Tästä johtuen keinotekoinen opiskelijavalinta suoritetaan rajattuna niin, että yhden vuosikurssin kooksi asettuu todellisten vuosikurssikokojen sijaan tätä pienempi opiskelijajoukko. Opiskelijavalinnassa hakijat hyväksytään opiskelijoiksi valintamenettelyn osoittamassa paremmuusjärjestyksessä vuosikurssi kerrallaan. Marginaalisten opiskelijoiden joukko on suurimmillaan, kun vuosivalinta rajataan 150 opiskelijaan; seuraavien lukujen opintomenestyksen tarkastelussa raportoidaan tämän valintarajauksen tulokset. Tulokset ovat laadullisesti robusteja valinnan rajaukselle<sup>24</sup> (N = 120, N = 150, N = 200, N = 250 vuosittain) ja ilman pienimpiä vuosikursseja 2011 ja 2012.

Kun keinotekoinen opiskelijavalinta on suoritettu, analyysin tarkoituksena on selvittää miten valitut opiskelijajoukot eroavat toisistaan. On todennäköistä, että valintamenettelystä riippumatta iso osa hakijoista tulisi valituksi joka tapauksessa. Molemmat valintamenettelyt perustuvat laajalti ylioppilastutkinnon arvottamiseen, jolloin parhaiten ylioppilaskokeissa menestyneet hakijat muodostanevat hakijoiden kärkijoukon, joka tulisi valituksi valintamenettelystä huolimatta. Näiden opiskelijoiden tutkiminen ei ole tutkielman tavoitteisiin nähden mielenkiintoista. Sen sijaan kiinnostavaa on pyrkiä ymmärtämään, miten ne opiskelijat eroavat toisistaan, jotka tulisivat valituksi yhdellä, mutta eivät toisella valintamenettelyllä. Näiden marginaalisiksi kutsuttujen opiskelijajoukkojen vertaileva tarkastelu vastaa kysymykseen siitä, miten valittujen opiskelijoiden joukko ja näiden keskimääräinen opintomenestys muuttuisi valintamenettelyä muuttamalla.

---

<sup>24</sup> Valinnan rajausta muuttamalla tarkasteltavan marginaalisten opiskelijoiden joukon opintomenestystä suhteessa koko tarkastellun opiskelijajoukon menestymiseen. Pienemmällä rajauksella (esimerkiksi 120 opiskelijan vuosivalinnalla) tarkasteltava joukko on lähempänä kuvion 2 nollapistettä, ja suuremmalla rajauksella joukko asettuu lähemmäksi aineiston todellisia marginaalisia opiskelijoita 45 asteen suoraa mukailleen. Analyysissä havaitaan, että opiskelijoiden keskimääräisen opintomenestyksen erotus kasvaa hieman opiskelijavalinnan rajauksen kasvaessa. Todellisessa opiskelijavalinnassa viimeisten, marginaalisten hakijoiden välisten erojen voidaan odottaa olevan suurempia kuin tässä esitellyn rajatun valinnan tuloksista voi suoraan päätellä. Kuvio 2 havainnollistaa ilmiön taustan: joukossa A vahvasti keskittynyt opiskelijajoukko hajaantuu sitä enemmän mitä kauemmaksi kuvion nollapistestä siirrytään. Tutkimusaineistolla rajausten väliset erot ovat kuitenkin niin marginaalisia, ettei valitulla rajauksella ole tulosten raportoinnille laadullista merkitystä. Analyysin tulokset esitellään rajauksella N = 150, jolloin marginaalisten opiskelijoiden joukko on suurimmillaan.

Opiskelijajoukkojen keskimääräisen opintomenestyksen eroja testattiin kahden otoksen kaksitahoisella Studentin t-testillä. Testin nollahypoteesi on, että vertailtavien ryhmien keskiarvon erotus on nolla eli ryhmät ovat samankaltaiset, jolloin hypoteesin hylkääminen ehdottaa, että ryhmien keskiarvot poikkeavat tilastollisesti toisistaan. T-testistä raportoidaan keskiarvojen erotus sekä sen keskivirhe ja tilastollinen merkitsevyys. Tutkimuksen aineistossa vertailtavat otosryhmät ovat toisistaan itsenäisiä, eli niillä ei ole päällekkäisyyksiä. Tällöin Studentin t-testi perustuu kahteen oletukseen: 1) otossuure on jakautunut normaalisti sekä ryhmien välillä samoin ja 2) ryhmien otosvarianssit ovat yhtenevät. Tutkimuksen opintomenestysmittarit eivät täysin noudata normaalijakaumaa, mutta ovat otosryhmien välillä samoin tavoin jakautuneet (jakaumakuviot liitteessä 2); otossuureiden varianssit ovat lähes identtiset. Vahvasti normaalijakaumasta poikkeavilla pienillä otoksilla t-testi voi olla harhainen. Kahden samankaltaisen otoksen t-testit ovat kuitenkin hyvin robusteja kohtuullisille poikkeamille näistä oletuksista, ja t-testi tarjoaa analyysille tarpeeksi luotettavan keskiarvojen erotuksen mittarin. Seuraavissa luvuissa esitellyt tulokset on lisäksi toistettu jakaumaneutraalilla Mann-Whitneyn U-testillä sekä erisuuruista varianssia olettavalla Welchin t-testillä ja ovat (yhtä erikseen osoitettua poikkeusta lukuun ottamatta) robusteja poikkeamalle normaalijakauman<sup>25</sup> sekä yhtenevän varianssin oletuksesta.

## 5.2 Arviointitulokset

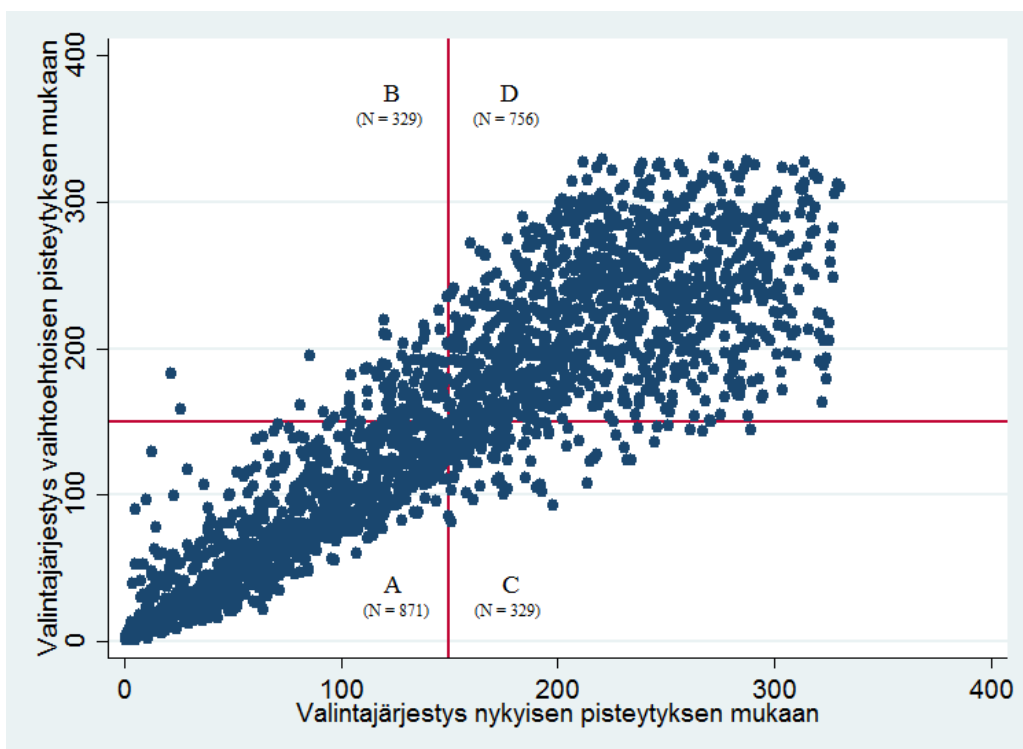
Seuraavat alaluvut esittelevät keinotekoisien opiskelijavalinnan vaikutuksen arvioinnin tulokset.

---

<sup>25</sup> Normaalijakaumasta poikkeavien muuttujien keskiarvon epäparametrinen Mann-Whitneyn U-testi eli Wilcoxonin järjestyssummatesti ei vaadi jakaumaoletuksia, mutta approksimoi yli 20 havainnon ryhmien jakauman normaaliksi keskeiseen raja-arvolauseeseen perustuen (todennäköisyyslaskennan periaate, jonka mukaan riippumattomien satunnaismuuttujien keskiarvo on likimain normaalisti jakautunut, kun  $N$  on tarpeeksi suuri). Näin ollen U-testi perustuu analyysin otoksilla ( $N > 300$ ) tosiasiassa samaan normaalijakauman oletamaan kuin Studentin t-testi. Tämän havainnon nojalla lienee todettavissa, että analyysi-aineisto noudattaa t-testille tarvittavalla tasolla normaalisti jakautuneen muuttujan ominaisuuksia.

### 5.2.1 Valintamenettelyn erottelevuus

Valintamenettelyjen eroja arvioitaessa on ensimmäiseksi selvittävä, mikä valintamenettelyjen vaikutus valituksi tulevaan opiskelijajoukkoon on. Kuvio 2 havainnollistaa, miten nykyisellä ja vaihtoehtoisella valintamenettelyllä paremmuusjärjestykseen laitettujen opiskelijoiden mallinnettu valintajärjestys hajaantuu valinnan edetessä. Kuviossa yksi piste vastaa yhtä hakijaa ja kuviossa on esitettyä koko hakijajoukon jakautuminen valinnan tuloksen mukaisiin joukkoihin. Joukko (A) kuvaa opiskelijavalinnan kärkihakijoiden joukkoa: hakijoita, jotka tulisivat valituksi molemmilla valintamenettelyillä. Joukko (B) kuvaa hakijoita, jotka tulisivat valituksi nykyisellä valintamenettelyllä, mutta joiden sijaluku ei riitä valintaan ylioppilastutkintoon puhtaasti perustuvalla vaihtoehtoisella valintamenettelyllä. Joukko (C) päinvastoin tulisi valituksi vaihtoehtoisella menettelyllä, mutta ei nykymallin mukaan. Joukko (D) koostuu opiskelijoista, jotka eivät tulisi valituksi kummallakaan valintamenettelyllä. Nykyisellä valintamenettelyllä valittujen opiskelijoiden koko joukko koostuu siis osajoukoista (A) ja (B); vaihtoehtoisella menettelyllä valituksi tulevat joukot (A) ja (C).



Kuvio 2. Valintamenettelyn erottelevuus

Hakijoiden joukko hajaantuu viuhkamaisesti nollapisteestä kauemmas siirryttäessä: tämä indikoi, että hakijoiden kärkijoukossa opiskelijoiden paremmuusjärjestys on molemmilla valintamenettelyillä hyvin samankaltainen, mutta opiskelijavalinnan edetessä opiskelijoiden väliset erot kasvavat. Kuviossa hajaantunut joukko (D) koostuu opiskelijoista, joiden valintajärjestyksen sijaluku poikkeaa suuresti sen perusteella, mitä kriteerejä valinnassa painotetaan. Kuten odotettua, suuri osa uusista opiskelijoista ( $871/1200 = 0,726$ , eli 73 %) tulisi valituksi molemmilla valintamenettelyillä. Loput 27 % eli vajaa kolmasosa uusista opiskelijoista valitaan eri tavalla valintamenettelystä riippuen. Näiden joukkojen osuus säilyy samankaltaisena keinotekoisen opiskelijavalinnan rajauksesta riippumatta. Valitsematta jäävien hakijoiden (D) suhteellinen osuus riippuu hakijoiden määrästä ja valinnanrajauksesta, eikä ole tutkielman analyysille merkityksellinen. Tulosten valossa vaikuttaa siltä, että valintamenettelyt erottelevat opiskelijoita onnistuneesti. Käytetyllä valintamenettelyllä on siis opiskelijavalinnassa merkitystä, ja noin kolmasosa hakijoista vaihtaa valintastatustaan käytetystä valintamenettelystä riippuen.

### 5.2.2 Valintamenettelyn yhteys keskimääräiseen opintomenestykseen

Yhtenä tutkielman keskeisimpänä kiinnostuksen kohteena on ollut selvittää, miten ylioppilastodistuksen perusteella valitut opiskelijat menestyisivät verrattuna niihin opiskelijoihin, jotka eivät olisi tulleet valituksi pelkkään ylioppilastutkintoon perustuvassa valinnassa. Kysymykseen vastataan valintamenettelyjen marginaalisia – siis valintastatustaan valintamenettelyn perusteella muuttavia – opiskelijoita tarkastelemalla. Taulukko 14 esittelee yhteenvedon marginaalisten opiskelijoiden keskimääräisen opintomenestyksen eroista. Ensimmäinen sarake (B) koskee opiskelijoita, jotka olisivat tulleet valituksi nykyisellä valintamenettelyllä mutta eivät puhtaassa todistusvalinnassa ja toinen sarake (C) opiskelijoita, jotka olisivat tulleet valituksi todistusvalinnassa mutta eivät nykyisellä valintamenettelyllä. Näiden ryhmien välinen, t-testillä laskettu erotus sekä sen keskivirhe ja merkitsevyystaso on raportoitu kolmannessa sarakkeessa. Kolmannen sarakkeen tulokset kertovat siis, miten valintamenettelyjen marginaalisten opiskelijoiden keskimääräinen opintomenestys eroaa näiden valintamenettelyiden välillä.

Taulukko 14. Valintamenettelyjen marginaalisten opiskelijoiden yhteenveto

## Valintamenettelyjen marginaalisten opiskelijoiden yhteenveto

Vertailussa nykyinen valintamenetelmä ja täysin ylioppilastutkintoon perustuva menetelmä

	Nykyinen (B)	Todistusvalinta (C)	Erotus (B) - (C)
Havaintojen määrä	329	329	
Osuus hakijoista	.14	.14	
Osuus valituista	.27	.27	
Naisten osuus	.3009	.5046	-.2036*** (.0375)
3 vuoden opintopistekertymä	148.891	153.043	-4.1522 (3.7606)
3 vuoden arvosanakeskiarvo	3.506	3.596	-.0899** (.0416)
Yhteisten opintojen arvosanakeskiarvo	3.399	3.397	.0021 (.0564)
Keskeyttämisprosentti	.0304	.0395	-.0091 (.0143)
Kolmen ensimmäisen vuoden aikana keskeyttäneet			
Valmistuneet 5 vuoden tavoiteajassa	.4255	.3421	.0834 (.1072)
Osuus vuosikurssien 2005 - 2007 opiskelijoista			
Valmistuneet 6 vuodessa	.6538	.6667	-.0128 (.0771)
Osuus vuosikurssien 2005 - 2007 opiskelijoista			
Valmistuneet ylipäänsä	.8000	.8148	-.0148 (.0481)
Osuus vuosikurssien 2005 - 2007 opiskelijoista			

Taulukko raportoi annettujen ryhmien keskiarvot ja näiden erotuksen. Suluissa keskiarvojen erotuksen keskivirheet. Kertoimet ovat epämerkitseviä tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa t-testissä.

Kokonaisuudessaan ryhmät (B) ja (C) ovat hyvin samankaltaiset. Opintomenestysmittareista ainoastaan opiskelijoiden kolmen vuoden arvosanakeskiarvot eroavat ryhmien välillä tilastollisesti merkitsevästi: todistusvalinnassa valittavien opiskelijoiden arvosanakeskiarvot ovat keskimäärin vajaat 0,1 yksikköä nykyisellä valintamenettelyllä valittavia opiskelijoita korkeampia. Ero on hyvin pieni. Myös opintopistekertymät ovat todistusvalinnan joukolla

hieman korkeammat, mutta ryhmien välistä eroa ei t-testillä voida tilastollisesti merkitsevästi todeta<sup>26</sup>. Yhteisten opintojen keskiarvo on ryhmillä sadasosan tarkkuudella sama, eikä myöskään keskeyttämisprosenttien tai valmistumisasteiden perusteella ryhmien eroista voida tehdä johdonmukaisia päätelmiä<sup>27</sup>. Tulosten perusteella valintakoe ei ole tuottanut lisäarvoa tutkielman keinotekoisessa opiskelijavalinnassa: hyvin samankaltainen, jopa hieman paremmin menestynyt opiskelijajoukko olisi saatu valittua ilman valintakokeen tuottamaa informaatiota.

Ainoa merkittävä ero marginaalisten opiskelijajoukkojen välillä on näiden sukupuolijakaumassa: ylioppilastutkintoon perustuvassa valinnassa (C) 50 % marginaalisista opiskelijoista olisi naisia, kun nykyisellä valintamenettelyllä valittujen joukossa on selvästi enemmän miehiä. Huomattavaa on, että myös molemmilla valintamenettelyillä valituksi tulevien opiskelijoiden joukossa (A) naisia on noin 50 % (taulukko 15). Hakijoiden kärkijoukossa sukupuolet ovat siis tasaisesti edustettuna. Marginaalisten opiskelijoiden välisten erojen voidaan tulkita johtuvan naisten ja miesten poikkeavasta suoriutumisesta ylioppilas- ja valintakokeissa. Naiset pärjäävät keskimäärin miehiä paremmin ylioppilaskokeissa, kun valintakokeissa miehet suoriutuvat naisia paremmin (Opetus- ja kulttuuriministeriö 2016b). Näiden tulosten pohjalta vaikuttaa siltä, että kauppatieteellisellä alalla paljon puhuttanut naisten kato<sup>28</sup> saataisiin korjattua ylioppilastutkintoa painottavalla valinnalla.

---

<sup>26</sup> Mann-Whitney'n U-testillä opintopistekertymien välinen ero todetaan tilastollisesti merkitsevästi 10 % merkitsevyystasolla.

<sup>27</sup> Opiskelija määritellään tutkielmassa keskeyttäneeksi, kun opiskelijan opinto-oikeus on Aalto-yliopiston kauppakorkeakoulun opiskelijarekisterissä merkitty päättyneeksi tämän kolmen ensimmäisen opintovuoden aikana. Keskeyttäneen opiskelijan opinto-oikeuden päättymisen syynä on "Jättänyt ilmoittautumatta" (yli 90 % keskeyttäneistä), "Luopunut" tai "Opintoaika käytetty" (vs. "Tutkinto valmistunut"). Valmistumisaste on laskettu vuosikurssien 2005 – 2007 opiskelijoille, jotka ovat tutkielman aineiston puitteissa ehtineet opiskelemaan sallitun tutkinnon maksimiajan seitsemän vuotta, ja siten saavuttaneet lopulliset valmistumisasteensa. Valmistumisaika on opiskelijoille laskettu opintojen aloittamisen ja kauppatieteiden maisteriksi valmistumispäivämäärän erotuksena.

<sup>28</sup> Sukupuolten välisestä jakaumasta on kauppatieteellisellä alalla puhuttu paljon, ja keinoja naisten houkuttelemiseksi alalle on pohdittu sekä Aalto-yliopiston kauppakorkeakoulun sisällä, että alan ammattiliitossa Suomen Ekonomieissa. Kauppatieteellisen alan yhteisvalinnan valintatilastojen mukaan vuoden 2016 (2015) valintakierroksella Aalto-yliopiston kauppakorkeakouluun hyväksyttiin 30 % (38 %) naisia. Yleisesti kauppatieteellisellä alalla sukupuolten jakauma on hieman tasapainoisempi: yhteensä kauppatieteitä opiskelemaan valituista 41 % (41 %) oli naisia.

Taulukko 15. Muiden kuin marginaalisten opiskelijoiden yhteenveto

## Muiden kuin marginaalisten opiskelijoiden yhteenveto

Vertailussa nykyinen valintamenetelmä ja täysin ylioppilastutkintoon perustuva menetelmä

	Ei valittu (D)	Aina valittu (A)	Erotus (D) - (A)
Havaintojen määrä	756	871	
Osuus hakijoista	.33	.38	
Osuus valituista	0	.73	
Naisten osuus	.3849	.5029	-.1179*** (.0246)
3 vuoden opintopistekertymä	143.179	149.822	-6.6430*** (2.463)
3 vuoden arvosanakeskiarvo	3.349	3.866	-.5170*** (.0258)
Yhteisten opintojen arvosanakeskiarvo	3.095	3.761	-.6661*** (.0364)
Keskeyttämisprosentti	.0384	.0746	-.0363*** (.0116)
Kolmen ensimmäisen vuoden aikana keskeyttäneet			
Valmistuneet 5 vuoden tavoiteajassa	.2977	.3147	-.0170 (.0560)
Osuus vuosikurssien 2005 - 2007 opiskelijoista			
Valmistuneet 6 vuodessa	.5400	.5288	.0112 (.0495)
Osuus vuosikurssien 2005 - 2007 opiskelijoista			
Valmistuneet ylipäänsä	.7325	.6889	.0437 (.0353)
Osuus vuosikurssien 2005 - 2007 opiskelijoista			

Taulukko raportoi annettujen ryhmien keskiarvot ja näiden erotuksen. Suluissa keskiarvojen erotuksen keskivirheet. Kertoimet ovat epämerkitseviä tai merkitseviä \*10 %, \*\*5 % tai \*\*\*1 % merkitsevyystasolla kaksisuuntaisessa t-testissä.

Taulukossa 15 esitellään niiden opiskelijoiden keskimääräinen opintomenestys, jonka keinotekoisin valintatuloksiin käytetyllä valintamenettelyllä ei ole merkitystä: joukko (A) koostuu opiskelijoista, jotka tulisivat valituksi kummassakin valinnassa ja (D) opiskelijoista, joita ei valittaisi kummallakaan menettelytavalla. Kolmas sarake sisältää taas näiden ryhmien välisen, t-testillä lasketun erotuksen sekä sen keskivirheen ja merkitsevyystason.



Keskiarvotestien tuloksia tarkastellessa huomataan, että hakijoiden kärkijoukon (A) ja valinnan ulkopuolelle jääneiden opiskelijoiden (D) välillä keskimääräisen opintomenestyksen erot ovat todellisia ja tilastollisesti merkitseviä. Hakijoiden kärkijoukolla on kolmen vuoden tarkastelujaksolla keskimäärin vajaat 7 opintopistettä enemmän ja reilut 0,5 yksikköä korkeampi keskiarvo kuin joukolla (D) – yhteisten opintojen arvosanakeskiarvojen ero on vielä hieman suurempi. Valmistumistodennäköisyyden mittareilla ei kuitenkaan näilläkään opiskelijajoukoilla voida tehdä tilastollisesti päteviä päätelmiä. Keskeyttämistodennäköisyys sen sijaan on hieman suurempi kärkihakijoiden joukossa. Opiskelijoiden poissaolokäyttäytymistä ja keskeyttämistä tutkineet Saarenmaa, Saari ja Virtanen (2010) esittävät, että opintojen keskeyttämisen taustalla on useimmiten toisen, mieluisamman opiskelupaikan saaminen. Tähän peilaten on luonnollista, että opiskelijavalinnoissa hyvässä asemassa olevat, parempiin ylioppilastutkintoarvosanoihin yltäneet opiskelijat keskeyttävät opintonsa muita useammin. Kokonaisuudessaan vertailun tulokset vahvistavat aiemmassa luvussa tehdyn päätelmän valintakriteereissä ja opinnoissa menestymisen positiivisesta yhteydestä.

Opiskelijajoukkojen erot osoittavat, että opiskelijavalinnan kärkihakijoilla ja viimeisten joukossa valituilla on keskimäärin todellisia opintomenestyseroja. Joukko (D) kuvastaa tässä Aalto-yliopiston kauppakorkeakoulun opiskelijavalinnan todellisia marginaalisia opiskelijoita: hakijoita, jotka ovat juuri tulleet valituksi, vaikka heidän valintapisteensä eivät ole olleet huomattavasti seuraavia, valitsematta jääneitä opiskelijoita korkeammat. Valintamenettelyn muuttaminen vaikuttaisi siis tosi asiassa joukon (D) kaltaisiin hakijoihin. Todellisessa opiskelijavalinnassa viimeisten, marginaalisten hakijoiden välisten erojen voidaankin odottaa olevan suurempia kuin taulukossa 14 esitellyn rajatun valinnan tuloksista voi suoraan päätellä<sup>29</sup>.

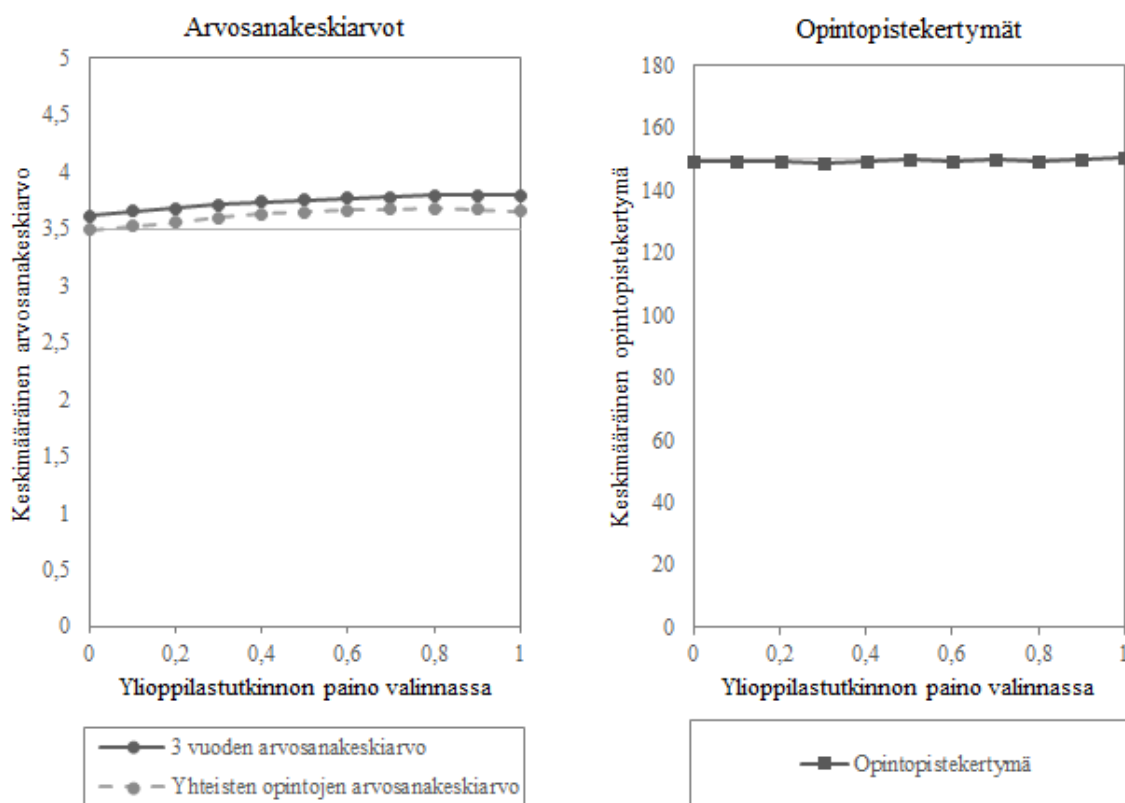
Nykyisen valintamenettelyn ja puhtaasti ylioppilastutkintoon perustuvan valinnan lisäksi valintakoetta ja ylioppilastutkintoa on mahdollista painottaa myös toisenlaisessa suhteessa toisiinsa. Kuvio 3 havainnollistaa, miten opiskelijoiden keskimääräinen opintomenestys muuttuisi, kun ylioppilastutkinnon painotusta valinnassa muutetaan välillä 0 – 100 %.

---

<sup>29</sup> Katso alaviite 24.

Vertailussa 150 opiskelijan vuosittainen opiskelijavalinta suoritetaan niin, että ensin valitaan X % opiskelijoista nykyisin käytetyn ylioppilastutkinnon lähtöpisteityksen perusteella, ja sitten loput opiskelupaikat (100 – X %) täytetään valintakokeen osoittamassa paremmuusjärjestyksessä. Kuviosta nähdään, että valintamenettelyllä ei ole havaittavaa vaikutusta kolmen vuoden opintopistekertymiin, jotka ovat ylioppilastutkinnon painotuksesta riippumatta keskimäärin noin 150. Arvosanakeskiarvoissa sen sijaan on havaittavissa nouseva yhteys arvosanojen ja ylioppilastutkinnon valintapainotuksen välillä. Erotus ääripäiden välillä on sekä kolmen vuoden että yhteisten opintojen arvosanakeskiarvoilla tarkasteltuna vajaat 0,2 yksikköä. Erot valintamenettelyjen välillä ovat siis pieniä. Voidaan kuitenkin todeta, että puhtaasti ylioppilastutkinnon perusteella valittu opiskelijajoukko vaikuttaa menestyvän keskimäärin valintakoetta hyödyntävien valintamenettelyjen opiskelijajoukkoja paremmin.

### Ylioppilastutkinnon painotuksen merkitys keskimääräiselle opintomenestykselle



Kuvio 3. Ylioppilastutkinnon painotuksen merkitys keskimääräiselle opintomenestykselle

Luvussa 2.2 määriteltiin, että taloustieteen näkökulmasta erillisen valintakokeen järjestäminen on opiskelijavalinnan kannalta kannattavaa ainoastaan, mikäli se antaa hakijoista sellaista parhaan valinnan kannalta arvokasta informaatiota, jota ei saada muutoin olemassa olevan tiedon perusteella pääteltyä. Määritelmän mukaan valintakokeen järjestäminen on siis perusteltua ainoastaan, mikäli valintakokeen perusteella valituiksi tulevat opiskelijat menestyvät opinnoissaan keskimäärin tarpeeksi paljon paremmin kuin ilman valintakokeen järjestämistä valitut opiskelijat. Tässä esiteltyjen tulosten valossa perusteita erillisen valintakokeen järjestämiselle Aalto-yliopiston kauppakorkeakoulussa ei näiden kriteerien perusteella havaita. Päinvastoin, on mahdollista, että puhtaalla todistusvalinnalla saataisiin valittua keskimäärin hieman paremmin menestyviä opiskelijoita. On kuitenkin huomattava, että saadut tulokset ovat hyvin riippuvaisia sekä esitellyn mallin rajoitteista, että tehdyistä menestyksen määrittelyistä ja asetetuista opiskelijavalinnan tavoitteista.

## 6 Keskustelu ja rajoitteet

Tutkielma on pyrkinyt arvioimaan kauppatieteellisen alan opiskelijavalinnan valintamenettelyn toimivuutta etenkin Opetus- ja kulttuuriministeriön esittämän opiskelijavalinnan uudistuksen tarpeiden näkökulmasta. Tavoitteena on ollut ymmärtää, miten ylioppilastutkintoarvosanojen ja valintakokeen perusteella voidaan ennustaa opiskelijoiden opintomenestystä, ja miten näihin havaintoihin pohjautuen opiskelijavalinta voitaisiin järjestää niin, että parhaat potentiaalisimmat opiskelijat tulisivat valituiksi mahdollisimman pienin kustannuksin. Lisäksi tavoitteena on ollut esitellä rajallinen malli opiskelijavalinnan arviointiin, johon pohjautuen on selvitetty, miten valintakokeesta luopuminen tai sen painoarvon pienentäminen voisi vaikuttaa opiskelijavalinnan tuloksiin Aalto-yliopiston kauppakorkeakoulussa. Tutkielman tarkoituksena ei kuitenkaan ole ollut tarjota valmista mallia ylioppilastutkintoarvosanojen hyödyntämiseen opiskelijavalinnoissa, eivätkä tutkielman tulokset sellaisenaan sovellu kauppatieteellisen alan opiskelijavalinnan suunnitteluun. Seuraavat kappaleet keskustelevat tutkielman tavoitteiden saavuttamisesta, tulosten sovellettavuudesta sekä tutkielman rajoitteista.

Tutkielman ensimmäisenä tavoitteena on ollut ymmärtää, miten hakijoiden valintahetkellä tiedossa olevien ominaisuuksien perusteella voidaan ennustaa opiskelijoiden opintomenestystä Aalto-yliopiston kauppakorkeakoulussa. Saatavilla oleva data on sisältänyt vain Kauppakorkeakouluun valituksi tulleet ja paikan vastaanottaneet opiskelijat, ei siis dataa valitsematta jääneistä tai tarjotun paikan hylänneistä hakijoista. Kiinnostuksen kohteena analyysille on siis ollut koko potentiaalisten hakijoiden joukko, mutta todellinen opintomenestys on voitu havaita vain Kauppakorkeakoulussa opintonsa aloittaneiden opiskelijoiden osalta. Tutkimusjoukko on selektiivinen, mikä asettaa rajoitteita tulosten käytettävyydelle. Opintomenestyksen selittämiseksi valikoituminen ei ole ongelma: kun tarkoituksena on selvittää, onko tämän aineiston opiskelijoiden aiemmalla opintomenestyksellä yhteyttä tulevaan menestykseen, ei sopivan selitysmallin löytämisen kannalta ole merkitystä sillä, miten opiskelijat on aineistoon valittu. Käytetty out-of-sample mallintaminen lisää ennusteiden validiteettia myös käytetyn aineiston ulkopuolella. Aineistoon mukautumiseen liittyvästä ylisovittamisen ongelmasta rajatussa aineistossa on käyty keskustelua luvussa 4.1.2. Tutkielman tulosten pohjalta voidaankin ennustaa, miten

ylioppilastutkintoarvosanat tai valintakoe ovat yhteydessä tämän aineiston kaltaisten opiskelijoiden opintomenestykseen.

Opintomenestyksen ennustetulosten yleistämisen kannalta aineiston valikoitumisella on kuitenkin merkitystä. Tutkielman aineistona on opiskelijajoukko, joka on opiskelijavalinnassa todettu käytössä olevalla mittaristolla kaikista potentiaalisimmaksi osajoukoksi laajasta hakijoiden joukosta. Rajatun aineiston tulosten perusteella ei voida tehdä pitkälle kantavia päätelmiä opiskelijoista, jotka ovat hyvin erilaisia nykyisen aineiston opiskelijoihin verrattuna. Voidaan hyvin perustein olettaa, että valitsematta jääneet hakijat poikkeavat valituksi tulleista opiskelijoista monella mittarilla, kategorisesti ainakin valintakriteeriensä osalta. Tulokset eivät liene sovellettavissa näiden hakijoiden opintomenestyksen ennustamiseen. Saadut OLS-estimaatit voitaneen kuitenkin yleistää tulevien valintakierrosten hakijajoukon parhaimmiston: mikäli aineiston opiskelijat ja tulevat kärkihakijat ovat tarpeeksi samankaltaisia ja eroavat toisistaan ainoastaan havainnoitavissa olevien ominaisuuksiensa perusteella, voidaan tutkimuksen tulosten opintomenestyksen ennustamisesta odottaa olevan sovellettavissa näihin tuleviin opiskelijoihin. Luvussa 2.3 esitelty aiempi kirjallisuus antaa ymmärtää, että Aalto-yliopiston kauppakorkeakoulun opiskelijoiden opintomenestyksen yhteys käytettyihin valintakriteereihin on säilynyt vuosikymmenestä toiseen lähes muuttumattomana.

Tutkielman toisena tavoitteena on ollut ennustemallien havaintoihin perustuen luoda ylioppilastutkintoarvosanoille vaihtoehtoinen valintapisteytysmalli, johon perustuen tunnistettaisiin jatkossa paremmin ne hakijat, joiden potentiaali menestyä opinnoissaan on suurin. Pisteytysmalli on luotu yksittäisten aine-arvosana kombinaatioiden ennustemerkityksellisyyteen perustuen, ja kärsii samoista yleistettävyyden rajoitteista kuin ennustemallit itsessään. Aineiston keinotekoisessa opiskelijavalinnassa pisteytysmallin avulla poimitaan hakijajoukosta onnistuneesti valintakokeeseen perustuvaa valintaa keskimäärin menestyneempi opiskelijajoukko. Luotu pisteytysmalli painottaa onnistuneesti kauppatieteellisellä alalla menestymistä ennustavia tekijöitä ja on nykyisin käytössä olevan pisteytysmallin mukainen.

Nykyisessä opiskelijavalinnassa käytettyjä monimutkaisia ylioppilastutkinnon pisteytysmalleja on kuitenkin kritisoitu siitä, että ne ovat lukiolaisille vaikeasti ymmärrettäviä ja korkeakouluille vaikeasti laskettavia (Opetus- ja kulttuuriministeriö 2016b) ilman että ne tuovat opiskelijavalintaan todellista hyötyä (Kupiainen 2014). Opetus- ja kulttuuriministeriössä onkin esitetty, että tulevien ylioppilastutkinnon pisteytysmallien olisi tärkeää olla yksinkertaisia ja yleisiä niin, että kaikissa opiskelijavalinnoissa hyödynnettäisiin samanlaista pisteiden laskentaa. Näitä geneerisiä pisteytysmalleja olisi tarvittaessa mahdollista painottaa alakohtaisesti, mikäli esimerkiksi matemaattisen osaamisen tai tietyn ainereaalikokeen tuloksen koettaisiin olevan kyseisellä alalla erityisen tärkeää. Tutkielman tulosten perusteella ei voida kuitenkaan suoraan tehdä päätelmiä tällaisen mahdollisen yleispisteytyksen soveltuvuudesta kauppatieteelliselle alalle. Tulokset ehdottavat kuitenkin, että ylioppilastutkinnon erilaisilla huomiointitavoilla on merkitystä sen ennustearvoon: valikoituihin, yksittäisiin ylioppilastutkintoarvosanoihin perustuva selitysmalli ennustaa opintomenestystä geneeristä keskiarvomuuttujaa paremmin. Tietyn yleisen pistemallin soveltuvuuden arviointi vaatisi tutkimuksen toistamista kyseisellä pisteytysmallilla.

Opiskelijavalinnan uudistamisen yhteydessä on myös ehdotettu, että todistusvalintoihin siirryttäessä osa hakijoista voitaisiin jatkossa valita valintakokeen sijasta yhteen tai kahteen erilliseen ylioppilaskokeeseen perustuen. Tällainen ylioppilastutkinnon kokeita kapeasti hyödyntävä pisteytys voisi esityksen mukaan toimia valintaperusteena myös hakijoille, jotka eivät ole suorittaneet ylioppilastutkintoa. Aineiston opiskelijoilla muutama ylioppilaskokeeseen perustuvan valinnan tutkiminen ei kuitenkaan olisi ollut mielekäästä, sillä julkilausuttujen valintamenettelyjen voidaan uskoa vaikuttavan merkittävästi hakijoiden käyttäytymiseen (Häkkinen 2004; Pietilä 2012). Mikäli esimerkiksi kauppatieteelliselle alalle jatkossa valittaisiin osa opiskelijoista ainoastaan matematiikan ja yhteiskuntaopin tai historian ylioppilaskokeeseen<sup>30</sup> perustuen, voitaisiin alalle odottaa saatavan joukko hakijoita, jotka olisivat keskittyneet erityisesti näiden alojen opiskeluun ja koearvosanojen korottamiseen. Tällaisen erikoistuneen hakijajoukon tulevaa

---

<sup>30</sup> Nykyinen valintakoe perustuu lukion opetussuunnitelman mukaisiin, määrättyjen yhteiskuntaopin, historian ja matematiikan kurssien sisältöön.

opintomenestystä ei voida käytössä olevan opiskelija-aineiston perusteella mielekkäästi ennustaa. Voidaan kuitenkin todeta, että kauppatieteellisen alan opinnoissa menestymistä vaikuttaa luvun 4 tulosten mukaan indikoivan erityisesti matemaattinen osaaminen. Erillinen kysymys on, vaativatko kaikki kauppatieteellisen alan erikoistumisohjelmat samanlaista osaamista, vai kannattaisiko esimerkiksi taloustieteen ja johtamisen opiskelijat valita eri perustein – tähän ei tämän tutkielman perusteella voida vastata.

Opiskelijavalinnan uudistaminen ylioppilastutkintopainotteiseksi tulee muuttamaan hakijoiden käyttäytymistä lopullisesta pisteystymallista riippumatta. Puhtaasti ylioppilastutkintoon perustuva valinta kannustaa hakijoita keskittymään jo lukio-opinnoissaan eniten pisteitä antaviin aineisiin. Opetus- ja kulttuuriministeriö (2016b) huomauttaakin esityksessään, että ylioppilastutkinnon pisteystymalleilla ja muilla valintavaatimuksilla on mahdollista myös tietoisesti ohjata lukiolaisten suuntautumista tietynkaltaisiin opintoihin. Tästä hyvä esimerkki on lääketieteellisen tiedekunnan opiskelijavalinnan selkeä ja pitkäaikainen painotus fysiikkaan, kemiaan ja biologiaan, joita lääketieteestä kiinnostuneet lukiolaiset osaavat valita opintoihinsa jo itseisarvona. Opiskelijavalinnan uudistuessa olisi hyvä mahdollisuus muodostaa myös kauppatieteelliselle alalle selkeästi viestityt, perustellut ja muuttumattomat valintavaatimukset, joilla voidaan signaloida lukiolaisille ja näiden opinto-ohjaajille alalla tarvittavia kyvykkyyksiä. Edelleen, tiettyjen ylioppilaskokeiden painottaminen valinnoissa tulee jatkossa näkymään näiden kokeiden arvosanojen korottamisen sekä näistä aineista huippuarvosanoja saaneiden hakijoiden suhteellisen kasvuna. Todistusvalintojen suunnittelussa onkin tarve panostaa myös siihen, että nykyisen ylioppilastutkinnon arvosanajärjestelmän rinnalle tai tilalle saadaan hakijoita entistä paremmin erotteleva pisteytys. Tasavertaisten laudaturin hakijoiden vyörystä on epämielikästä suorittaa reilua opiskelijavalintaa ilman nykyistä tarkempaa arvosteluasteikkoa. Opetus- ja kulttuuriministeriön opiskelijavalinnan työryhmässä (2016b) on jo esitetty muun muassa raakapisteisiin, persentiileihin, pistekoreihin ja valtakunnallisten kurssien määrää painottaviin pisteytyksiin perustuvia ratkaisuja.

Tutkielman kolmantena tavoitteena on ollut tarjota sovellettava ja käytännöllinen menetelmä opiskelijavalintojen arviointiin ja vaihtoehtoisten valintamenettelyjen vertailuun. Toisessa empiirisessä vaiheessa suoritettujen keinotekoisien opiskelijavalinnan tulokset antavat ymmärtää, että valintamenettelyillä on merkitystä opiskelijajoukon valintaan, ja että vuosien 2005 – 2012 opiskelijavalinnoissa erillisen valintakokeen järjestäminen ei olisi ollut Aalto-yliopiston kauppakorkeakoulussa kannattavaa, sillä samankaltainen opiskelijajoukko oltaisiin saatu valittua puhtaaseen todistusvalintaan perustuen. Tulos on aiemmin esiteltyihin rajoituksiin perustuen heikosti yleistettävissä nykyisestä aineistosta laajasti poikkeaville hakijoille. On kuitenkin epäselvää, miten valintakokeesta luopuminen tulisi vaikuttamaan seuraavien hakijajoukkojen muodostumiseen. Tutkielman opiskelijat ovat kaikki aikanaan suorittaneet erillisen valintakokeen ja tulleet valituksi puhtaasti siihen tai sen perusteella laskettuihin yhteispisteisiin perustuen. Pietilä (2012) argumentoi, että valintakokeella on ollut hakijoihin itsevalitseva vaikutus: valintakokeen olemassaolo on karsinut joukosta ne hakijat, jotka eivät ole kokeneet kauppatieteellistä alaa tarpeeksi kiinnostavaksi käyttääkseen rajallisia resurssejaan valintakokeeseen valmistautumiseen. Pekkarinen ja Sarvimäki (2016) argumentoivat edelleen, että valintakokeita painottava opiskelijavalintajärjestelmä on ajanut opiskelijat hakemaan todellisten kiinnostustensa sijaan niihin oppilaitoksiin, joihin uskovat todennäköisimmän tulevansa valituksi.

Valinnan muuttuessa hakijoiden itsevalinnan perusteet tulevat muuttumaan. Pekkarinen ja Sarvimäki (2016) esittävätkin, että yleisesti todistusvalintaan perustuvassa haussa hakijoiden kannattaa pyrkiä reilusti itselleen kiinnostavimpaan koulutusohjelmaan. Riskinä opiskelijavalinnan uudistamisen siirtymävaiheessa kuitenkin on, että ensimmäisenä tai ainoana todistusvalinnan käyttöön ottaneesta korkeakoulusta tulee hakijoiden varavaihtoehto, johon haetaan ensimmäiseksi priorisoitavan hakuvaihtoehdon ohella. Näiden havaintojen pohjalta lienee reilua odottaa, että valintamenettelyjen muuttuessa tulevaisuuden hakijajoukot voivat näyttää erilaisilta kuin aineiston opiskelijoiden valintahetkellä, eikä analyysin tuloksia voi siten ongelmitta yleistää tuleviin valintoihin. Tutkielman havaintoihin perustuen ei siis voida suoraan todeta, että valintakokeesta luopumalla Kauppakorkeakouluun saataisiin valittua aiempaa parempia opiskelijoita. Tulokset antavat kuitenkin viitteitä siitä, että muiden tekijöiden pysyessä samana, todistusvalinnalla saataisiin valittua vähintään yhtä hyvä joukko opiskelijoita kuin nykyisillä valintamenettelyillä.



Opiskelijoiden opintojen etenemistä ei tutkielman havaintojen perusteella kuitenkaan voida valintakriteerejä muuttamalla vaikuttaa. Opiskelijoiden valmistumisen jouduttaminen on kuitenkin Sipilän hallituksen kärkihankkeiden joukossa, ja opintojen nopean etenemisen tukemiseen kannustetaan vahvasti yliopistojen strategisen rahoituksen allokoinnissa. Yksi keino nuorempien maistereiden saamiseen on suosia opiskelijavalinnoissa ensikertalaisia hakijoita muiden pyrkijöiden kustannuksella. Toinen mahdollinen ratkaisu voisi opiskelijavalinnan sijaan löytyä itse opiskeluaikaan vaikuttavista tekijöistä. Tutkielman luvussa 2.3 viitattiin Aalto-yliopiston kauppakorkeakoulussa opintojen etenemistä tutkineen Arja Gillbergin (1987) tuloksiin, joiden mukaan opiskelijoiden valmistumisaikaa ennustivat parhaiten opiskelijoiden tehokas opiskelutyyli, opintoihin sitoutuminen ja opiskelun palkitsevuus, sekä opiskeluympäristö, opintojen organisointi ja opiskelijoiden sosiaaliset suhteet. Tuoreemman tutkimuksen mukaan heikko opiskelumotivaatio, opintojen työläys, puutteellinen opintojen ohjaus ja puutteelliset opiskelutaidot hidastavat opiskelijoiden valmistumista huomattavasti (Saarenmaa 2010). Nämä kaikki ovat tekijöitä, joihin voidaan ja joihin tulisi opiskelijoiden opintojen aikana vaikuttaa opetusta antavan korkeakoulun puolelta. Opintomenestyksen tutkimuksen tulevana kiinnostuksenkohteena voisikin olla selvittää millaisin keinoin opiskelijoiden menestykseen voidaan jo koulutukseen valittujen opiskelijoiden osalta vaikuttaa.

Lopulta opiskelijavalintaa arvioitaessa on tärkeää ymmärtää, että arvioinnin tulokset perustuvat siihen, mihin tavoitteisiin menestymistä peilataan. Tässä tutkielmassa opintomenestyksen kriteereiksi valikoitui Aalto-yliopiston kauppakorkeakoulun itse käyttämiin opintomenestysindekseihin perustuen opintojen laatua ja niiden etenemistä mittaavat muuttujat: arvosanakeskiarvot ja opintopistekertymät. Opintomenestystä on kuitenkin mahdollista mitata myös muilla muuttujilla. Isossa kuvassa vielä tärkeämpää olisi arvioida yksilön menestystä koulutuksen jälkeen. Opiskelijavalintaa arvioitaessa olisikin tärkeää tutkia esimerkiksi millainen yhteys valintakriteerien ja yksilön tulevan tulotason välillä on, tai millainen vaikutus valintamenettelyn muuttamisella on koulutettavaan opiskelijajoukkoon yhteiskunnallisesti. Voiko valintamenettelyillä vaikuttaa siihen millaisia ekonomieja tai lääkäreitä yhteiskuntaan tuotetaan? Ahola (2004) huomauttaakin, että opiskelijavalinnalla on muitakin tehtäviä kuin hakijajoukon karsinta. Tulevia

opiskelijavalinnan käytänteitä suunniteltaessa ja arvioitaessa on syytä pitää kirrkaasti mielessä valinnan tavoitteet, tarkoituksenmukaisuus ja tutkittuun tietoon perustuva toimivuus. Vanhan sanonnan mukaisesti: sitä saa, mitä mittaa.

## 7 Johtopäätökset

Tutkielma on pyrkinyt kauppatieteellisen alan opiskelijavalinnan näkökulmasta esittelemään opiskelijavalinnan ongelman, löytämään ylioppilastutkintoarvosanojen ja valintakokeen ennustearvon opintomenestykselle, muodostamaan ylioppilastutkintoarvosanoille niiden ennustearvoon perustuvan opiskelijavalinnan pisteytysmallin sekä selvittämään miten valintamenettelyjen muuttaminen vaikuttaisi valituksi tulevien opiskelijoiden joukkoon ja keskimääräiseen opintomenestykseen. Käytännöllisenä tavoitteena on ollut esitellä rajallinen malli opiskelijavalinnan arviointiin, johon pohjautuen on selvitetty, miten valintakokeesta luopuminen tai sen painoarvon pienentäminen voisi vaikuttaa opiskelijavalinnan tuloksiin Aalto-yliopiston kauppakorkeakoulussa. Tutkielman tarkoituksena ei kuitenkaan ole ollut tarjota valmista mallia ylioppilastutkintoarvosanojen hyödyntämiseen opiskelijavalinnoissa, eivätkä tutkielman tulokset sellaisenaan sovellu suoraan yleistettäväksi tai kauppatieteellisen alan opiskelijavalinnan suunnitteluun.

Tutkielman tulokset ovat linjassa aiemman kirjallisuuden kanssa osoittaessaan, että ylioppilastutkintoarvosanat ennustavat korkeakouluopintojen kurssiarvosanojen muodostumista varsin mielekkäästi. Erityisesti matemaattinen osaaminen vaikuttaa indikoivan menestymistä kauppatieteellisissä opinnoissa ja ylioppilastutkintoarvosanojen painottaminen hyvin määritellyllä pisteytysmallilla vaikuttaa olevan arvosanakeskiarvojen ennustamisen kannalta suotavaa. Opiskelijavalinnan uudistuessa onkin hyvä mahdollisuus muodostaa kauppatieteelliselle alalle perustellut ja toimivat valintavaatimukset, joilla voidaan selkeän viestinnän kautta ohjata tulevien hakijoiden valmentautumista alalle. Opintopistekertymien ennustaminen vaikuttaa kuitenkin haastavammalta, ja aiempaan kirjallisuuteen peilaten todetaankin, että opintojen etenemiseen vaikuttanevat valintakriteerien mittaamien ominaisuuksien sijaan opintojen aikaiset tekijät. Opiskelijavalinnan uudistamisen sijaan nopeammin valmistuvia opiskelijoita saataneen siis kiinnittämällä huomiota jo valittujen opiskelijoiden edistymiseen. Valintakokeen merkitys opintomenestyksen ennustamiseen jää sekä opintopistekertymillä että arvosanakeskiarvoilla mitättömän pieneksi.

Tutkielman tulosten valossa vaikuttaakin siltä, että erillisen valintakokeen järjestäminen ei olisi ollut Aalto-yliopiston kauppakorkeakoulussa kannattavaa tehokkaan opiskelijavalinnan määrittelyyn nähden. Tutkielman keinotekoisessa opiskelijavalinnassa valintakoe ei onnistu

tuomaan valintaan lisäarvoa: puhtaalla todistusvalinnalla olisi valittu hyvin samankaltainen, jopa hieman paremmin menestynyt opiskelijajoukko. Todellisessa opiskelijavalinnassa viimeisten, marginaalisten hakijoiden välisten erojen voidaan odottaa olevan suurempia kuin tutkielman tuloksista voi suoraan päätellä. Tutkielma on Kauppakorkeakoulussa ensimmäinen laatuaan: valintamenettelyjen vaikutusta opiskelijoiden keskimääräiseen opintomenestykseen ei ole ennen koulussa tutkittu. On kuitenkin huomattava, että tutkielman tulokset ovat hyvin riippuvaisia sekä mallin ja aineiston asettamista rajoitteista, että opintomenestyksen ja opiskelijavalinnan tavoitteiden määrittelyistä.

Tuloksia ei tulisikaan sellaisenaan yleistää tai soveltaa opiskelijavalinnan suunnitteluun ilman rajoitteiden syvällistä ymmärtämistä. Tutkielman aineistosta ja käytetyistä menetelmistä johtuen saadut tulokset sopivat aineiston opiskelijoiden opintomenestyksen ennustamiseen varsin hyvin, mutta eivät välttämättä sovellu sellaisenaan yleistettäväksi muihin opiskelijapopulaatioihin. Rajatun ja valikoituneen aineiston tulosten perusteella ei voida tehdä pitkälle kantavia päätelmiä opiskelijoista, jotka ovat hyvin erilaisia nykyisen aineiston opiskelijoihin verrattuna. Tietyin rajoituksin ja oletuksin tutkielman tuloksia voidaan kuitenkin soveltaa Kauppakorkeakoulun tuleviin hakijoihin, jotka ovat tarpeeksi samankaltaisia nykyisten opiskelijoiden kanssa. Toistaiseksi on epäselvää, miten opiskelijavalinnan uudistaminen tulee vaikuttamaan tulevien hakijajoukkojen muodostumiseen.

Aineiston ja menetelmien asettamista rajoitteista huolimatta tutkielma on esitellyt käytännöllisen empiirisen mallin opiskelijavalinnan arviointiin. Tutkielma on osoittanut, miten kauppatieteellisen alan yhteisvalinnassa käytetyt valintakriteerit ennustavat opintomenestystä Aalto-yliopiston kauppakorkeakoulussa ja miten ylioppilastutkintoarvosanat tulisi pisteyttää, jotta niiden merkityksellisyys opiskelijavalinnassa olisi niiden ennustearvon mukainen. Lopulta tutkielmassa on selvitetty, miten valintakokeesta luopuminen tai sen painoarvon pienentäminen voisi vaikuttaa opiskelijavalinnan tuloksiin Aalto-yliopiston kauppakorkeakoulussa. Tulokset tarjoavat lähtökohdan kauppatieteellisen alan opiskelijavalinnan tarkoituksenmukaisuuden arviointiin.

## Lähteet

Aalto-yliopiston kauppakorkeakoulu (2017a). Kauppatieteiden kandidaattiohjelman erikoistumisalueen valintaperusteet ja jatkaminen maisteritasolle. 25.08.2017. Saatavilla 23.9.2017:

<https://into.aalto.fi/download/attachments/12331237/Erikoistumisaluevalintakriteerit.KAAK%2025082017.pdf?version=1&modificationDate=1503993525394&api=v2>.

Aalto-yliopiston kauppakorkeakoulu (2017b). Hakeminen vaihtoon:

Opintomenestysindeksi. 31.08.2017. Saatavilla 23.9.2017:

<https://into.aalto.fi/display/fimasterbiz/Hakeminen+vaihtoon>.

Aalto-yliopiston kauppakorkeakoulun opintoasiaintoimisto (2010). Opinto-opas 2010-2011: KTK- ja KTM-tutkintoja suorittaville. Aalto-yliopiston kauppakorkeakoulu. Aalto Print

Aalto-yliopiston kauppakorkeakoulun opintoasiaintoimisto (2011). Opinto-opas 2011-2012: KTK- ja KTM-tutkintoja suorittaville. Aalto-yliopiston kauppakorkeakoulu. Aalto Print.

Aalto-yliopiston kauppakorkeakoulun opintoasiaintoimisto (2012). Opinto-opas 2012-2013: KTK- ja KTM-tutkintoja suorittaville. Aalto-yliopiston kauppakorkeakoulu. Unigrafia.

Ahola, Sakari & Anu Kokko (2001). Finding the Best Possible Students: Student selection and its problems in the field of business. *Journal of Higher Education Policy and Management*. Vol. 23, No. 2, pp. 191-203.

Ahola, Sakari (2004). Yhteishausta yhteisvalintaan: Yliopistojen opiskelijavalintojen kehittäminen. Opetusministeriö. Opetusministeriön työryhmämuistioita ja selvityksiä 2004:9. Saatavilla 25.9.2017:

<https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/80440/tr09.pdf?sequence=1>.

Angrist, Joshua & Jörn-Steffen Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Betts, Julian & Darlene Morell (1999). The determinants of undergraduate grade point average: the relative importance of family background, high school resources, and peer group effects. *The Journal of Human Resources*. Vol. 34, pp. 268-293.

Burnham, Kenneth P. & David R. Anderson (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second Edition. Springer-Verlag New York. USA.

Gillberg, Arja (1987). Tutkimus Helsingin kauppakorkeakoulun vuosina 1977-1979 sisäänpäässeiden opiskelijoiden opintomenestykseen vaikuttaneista tekijöistä. Helsingin kauppakorkeakoulun julkaisuja D-94. Helsinki.

Hakoniemi, Janne (1995). Valmistuuko kuuden laudaturin ylioppilas myös yliopistosta erinomaisin arvosanoin?. Turun yliopiston historian laitoksen julkaisuja 33. Turku: Turun yliopiston offset-paino.

Helsingin kauppakorkeakoulun opintoasiaintoimisto (2005). Opinto-opas 2005-2006: KTK- ja KTM-tutkintoja suorittaville. Helsingin kauppakorkeakoulu. HSE Print.

Helsingin kauppakorkeakoulun opintoasiaintoimisto (2006). Opinto-opas 2006-2007: KTK- ja KTM-tutkintoja suorittaville. Helsingin kauppakorkeakoulu. HSE Print.

Helsingin kauppakorkeakoulun opintoasiaintoimisto (2007). Opinto-opas 2007-2008: KTK- ja KTM-tutkintoja suorittaville. Helsingin kauppakorkeakoulu. HSE Print.

Helsingin kauppakorkeakoulun opintoasiaintoimisto (2008). Opinto-opas 2008-2009: KTK- ja KTM-tutkintoja suorittaville. Helsingin kauppakorkeakoulu. HSE Print.

Helsingin kauppakorkeakoulun opintoasiaintoimisto (2009). Opinto-opas 2009-2010: KTK- ja KTM-tutkintoja suorittaville. Helsingin kauppakorkeakoulu. HSE Print.

Häkkinen, Iida (2004). Do University Entrance Exams Predict Academic Achievement?. Working Paper. Department of Economics, Uppsala Universitet. No 2004:16

Jussila, M (1996). Korkeakoulujen opiskelijavalinnat. Kehittämisen haasteita ja mahdollisuuksia. Koulutus- ja tiedepolitiikan linjan julkaisusarja 1996:34. Opetusministeriö.

Kauppätieteellisen alan yhteisvalinta (2015). Valintatilasto 2015. Saatavilla 6.11.2017: <http://www.kauppätieteet.fi/wp-content/uploads/2015/12/Sukupuolijakauma-2015.pdf>.

Kauppatieteellisen alan yhteisvalinta (2016). Valintatilasto 2016. Saatavilla 6.11.2017:  
<http://www.kauppatieteet.fi/wp-content/uploads/2016/09/Hakijat-hyv%C3%A4ksytyt-naiset-ensikertalaiset.pdf>.

Kauppatieteellisen alan yhteisvalinta (2017). Ylioppilastutkinnon ja valintakokeen pisteytys. Saatavilla 13.9.2017:  
<http://www.kauppatieteet.fi/hakeminen/valintamenettely/pisteytys/>.

Komulainen, Erkki. (1987) Helsingin kauppakorkeakoulun opiskelijavalinnan toimivuuden tarkastelua kesän 1987 valinnan ylioppilasaineistolla. Helsingin kauppakorkeakoulun selvityksiä E-51.

Komulainen, Erkki. (1990a) Helsingin kauppakorkeakoulun kesän 1989 opiskelijavalinnan palauteraportti. Helsingin kauppakorkeakoulun selvityksiä E-63.

Komulainen, Erkki. (1990b) Helsingin kauppakorkeakoulun kesän 1990 opiskelijavalinta. Taustamuuttujien ja valintakokeen psykometrinen tarkastelu. Helsingin kauppakorkeakoulun selvityksiä E-67.

Komulainen, Erkki. (1992) Helsingin kauppakorkeakoulun kesän 1992 opiskelijavalinta. Taustamuuttujien ja valintakokeen psykometrinen tarkastelu ylioppilasaineistolla. Helsingin kauppakorkeakoulun selvityksiä E-79.

Komulainen, Erkki. (1996) Helsingin kauppakorkeakoulun opiskelijavalinta 1995. Taustamuuttujien ja valintakokeen psykometrinen tarkastelu ylioppilasaineistolla. Helsingin kauppakorkeakoulun selvityksiä E-92.

Komulainen, Erkki & Anne-Marja Nevgi (1993). Helsingin kauppakorkeakoulun opiskelijoiden minäkuvan, ammatillisten mielikuvien ja motivaation yhteydet opintomenestykseen. Helsingin kauppakorkeakoulun selvityksiä E-82.

Kupiainen, Sirkku (2014). Hakija valintojen pyörteessä: Ylioppilastutkinnon arvosanojen hyödyntäminen korkeakoulujen opiskelijavalinnassa. Seminaari korkeakoulujen opiskelijavalintojen uudistamisen etenemisestä.

- Kupiainen, Sirkku & Jukka Marjanen (2014). Ylioppilas valintojen pyörteessä: Ylioppilastutkinto ja korkeakoulujen opiskelijavalinta. Tutkimusprojektin loppuraportti. Koulutuksen arviointikeskus.
- Kupiainen, Sirkku, Jukka Marjanen & Jarkko Hautamäki (2016). The problem posed by exam choice on the comparability of results in the Finnish matriculation examination. *Journal for Educational Research Online*. Vol. 8, No. 2, pp. 87-106.
- Mieskonen, Emma (2015). Menestysanalyysi: Aalto-yliopiston kauppakorkeakoulun opiskelijoiden opintomenestys 2005 – 2014. Selvitys Aalto-yliopiston kauppakorkeakoulun johtoryhmälle. Elokuu 2015.
- Mieskonen, Emma (2016). Menestysanalyysi: Opintomenestyksen ennustaminen ylioppilastutkintoarvosanoilla Aalto-yliopiston kauppakorkeakoulussa. Selvitys Aalto-yliopiston kauppakorkeakoulun johtoryhmälle. Maaliskuu 2016.
- Opetus- ja kulttuuriministeriö (2016a). Yliopistojen rahoitusmalli uudistuu. Tiedote. 29.4.2016. Saatavilla 14.9.2017: [http://minedu.fi/artikkeli/-/asset\\_publisher/yliopistojen-rahoitusmalli-uudistuu](http://minedu.fi/artikkeli/-/asset_publisher/yliopistojen-rahoitusmalli-uudistuu).
- Opetus- ja kulttuuriministeriö (2016b). Valmiina valintoihin: Ylioppilastutkinnon parempi hyödyntäminen korkeakoulujen opiskelijavalinnoissa. Opetus- ja kulttuuriministeriön julkaisuja 2016:37. 4.11.2016. Saatavilla 12.9.2017: <http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/79291/okm37.pdf>.
- Pekkarinen, Tuomas & Matti Sarvimäki (2016). Parempi tapa valita korkeakouluopiskelijat. VATT Policy Brief 1-2016. 21.4.2016. Saatavilla 13.9.2017: <http://vatt.fi/parempi-tapa-valita-korkeakouluopiskelijat>.
- Peltonen, Riitta-Liisa, Arja Gillberg & Erkki Komulainen (1990). Kauppätieteellisen alan opiskelijavalinnassa vaikuttavat tekijät. Helsingin kauppakorkeakoulun selvityksiä E-63.
- Pietilä, Lotta (2012). Predicting academic performance: Does entrance exam matter? – Evidence from Helsinki School of Economics. Maisterin tutkielma. Taloustieteen laitos. Aalto-yliopiston kauppakorkeakoulu.



Rothstein, Jesse (2004). College Performance Predictions and the SAT. *Journal of Econometrics*, 121(1–2), 297–317.

Saarenmaa, Kaisa, Katja Saari & Vesa Virtanen (2010). *Opiskelijatutkimus 2010: Korkeakouluopiskelijoiden toimeentulo ja opiskelu*. Opetus- ja kulttuuriministeriön julkaisuja 2010:18.

Valmennustiimi Eximia Oy (2009). *Haku yliopistoihin ja korkeakouluihin: Valintakokeet 2010, Yhteenveto opiskelijavalinnoista 2009*. 18.12.2009. Saatavilla 22.9.2017:  
[http://www.eximia.fi/uploads/opo-apu/OPO-PAKETTI\\_Yhteenveto\\_valintakoetiedoista\\_2009-2010.pdf](http://www.eximia.fi/uploads/opo-apu/OPO-PAKETTI_Yhteenveto_valintakoetiedoista_2009-2010.pdf).

Valtioneuvoston kanslia (2004). *Valtioneuvoston asetus yliopistojen tutkinnoista*. Finlex 794/2004. 19.08.2004. Saatavilla 26.10.2017:  
<http://www.finlex.fi/fi/laki/alkup/2004/20040794>.

Valtioneuvoston kanslia (2015a). *Ratkaisujen Suomi: Pääministeri Juha Sipilän hallituksen strateginen ohjelma*. Hallituksen julkaisusarja 10/2015. 29.5.2015.

Valtioneuvoston kanslia (2015b). *Toimintasuunnitelma strategisen hallitusohjelman kärkihankkeiden ja reformien toimeenpanemiseksi*. Hallituksen julkaisusarja 13/2015. 28.9.2015.

Ylioppilastutkintolautakunta (2014). *Arvosanojen vertailtavuuden parantaminen*. Ylioppilastutkintolautakunnan esitys. 20.5.2014. Saatavilla 14.9.2017:  
[https://www.ylioppilastutkinto.fi/images/sivuston\\_tiedostot/Kehittaminen/arvosanojen\\_vertailtavuus\\_fi.pdf](https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/Kehittaminen/arvosanojen_vertailtavuus_fi.pdf).

Ylioppilastutkintolautakunta (2016). *Ylioppilastutkintoa tarvitaan edelleen* (HS 28.5.2016). Ylioppilastutkintolautakunta. Mieli-pide. Saatavilla 4.11.2017:  
<https://www.ylioppilastutkinto.fi/ajankohtaista/uutiset-tapahtumat/10-uutinen/337-ylioppilastutkintoa-tarvitaan-edelleen-hs-28-5-2016-114446-15022017>.

Öckert, Björn (2001). *Does pre-university background matter? In effects of higher education and the role of admission selection*. Dissertation series no. 52. Swedish Institute for Social Research. Stockholm University.

## Liitteet

### Liite 1: Kauppatieteellisen alan yhteisvalinta

Suomen kauppatieteelliset korkeakoulut suorittavat opiskelijavalintansa yhdessä sovituin kriteerein kauppatieteellisen alan yhteisvalinnassa. Yhteisvalinnassa valituksi tuleminen oikeuttaa suorittamaan sekä kauppatieteiden kandidaatin, että maisterin tutkinnon opiskelijan hyväksyvässä oppilaitoksessa. Aalto-yliopiston kauppakorkeakoulu on ollut mukana kauppatieteellisen alan yhteisvalinnassa vuodesta 2005 alkaen (Valmennustiimi Eximia Oy 2009). Kauppatieteellisen alan valintamenettelyä on vuoden 2005 jälkeen muutettu useita kertoja, viimeksi tulevan vuoden 2018 opiskelijavalintaa varten. Uusimmassa uudistuksessa opiskelijavalintaa on muokattu vastaamaan Opetus- ja kulttuuriministeriön toimenpideohjelmaa (2016b): valintamenettelyyn on lisätty mahdollisuus tulla valituksi pelkkään ylioppilastutkintomenestykseen perustuen, valintakokeen painoarvoa on pienennetty ja siihen valmentautumiseen kuluva aika on pyritty vähentämään keventämällä koesisältöjä. Valintakoe kirjallisuus on aiempienkin vuosien aikana vaihdellut, ja esimerkiksi vuonna 2010 mukana oli kirjallisuutta kauppatieteiden eri aloilta: organisaatiot ja johtaminen, laskentatoimi, markkinointi sekä taloustiede (Valmennustiimi Eximia Oy 2009). Matematiikka on ollut valintakokeissa mukana vuosien 2010 ja 2011 valintoja lukuun ottamatta. Nyt alakohtaisesta valintakoe kirjallisuudesta on siirrytty lukion opetussuunnitelman mukaisiin oppimääriin, ja valintakoe perustuu nykyisin määrättyjen yhteiskuntaopin, historian ja matematiikan kurssien sisältöön. Valinta pyrkii entistä paremmin hyödyntämään ylioppilastutkinnon tarjoamaa informaatiota hakijoiden kyvykkyydestä.

Kauppatieteellisen alan nykyinen, kevästä 2018 alkaen sovellettava yhteisvalinta perustuu kahteen valintatapajonoon: todistuspistejonoon ja koepistejonoon. Pitkään käytössä olleesta yhteispistevalinnasta on siis luovuttu, ja tilalle on tuotu suora todistusvalinta, joka on yhteisvalinnassa ensisijainen, ainoastaan ensikertalaisille hakijoille suunnattu valintatapajono. Yhteispistevalinnassa opiskelijan valintapisteet muodostettiin kahden koekokonaisuuden perusteella: lähtöpisteet laskettiin ylioppilastutkintoarvosanojen pisteytysäännöillä ja koepisteet muodostuivat valintakokeen kysymyksistä laskusäännön avulla. Yhteispistevalintajonossa näiden kahden kokonaisuuden pisteet laskettiin yhteen

yhdeksi valintapistekokonaisuudeksi; valintakoepistejonossa opiskelijat laitettiin paremmuusjärjestykseen suoraan valintakoepisteiden perusteella. Noin 60 % tutkielman aineiston opiskelijoista on hyväksytty opiskelemaan yhteispistevalinnassa; loput 40 % on valittu valintakoepistejonosta.

Todistuspistejono otettiin käyttöön vuoden 2017 opiskelijavalinnassa, jolloin 20 % uusista opiskelijoista valittiin uuden valintamenettelyn mukaisesti puhtaasti ylioppilastutkintoon perustuen. Loput opiskelijat valittiin tuttuun tapaan yhteispisteisiin (50 %) ja valintakoepisteisiin (30 %) perustuen. Nykyisellä valintamenettelyllä vuoden 2018 opiskelijavalinta tehdään ensin todistusvalintajonosta, jolla täytetään 60 % kunkin hakukohteen aloituspaikoista. Todistusvalinnan valintajono muodostetaan asettamalla hakijat todistuspisteiden mukaiseen paremmuusjärjestykseen, josta parhaat pisteet saaneet hyväksytään suoraan opiskelemaan ilman tarvetta valintakokeen tekemiseen. Loput 40 % aloituspaikoista täytetään koepistejonon mukaisessa järjestyksessä. Koepistejono muodostetaan uudistuneessa valintajärjestelmässä pelkästään valintakoepisteiden perusteella, eikä ylioppilastutkintoarvosanoja huomioida tässä valintajonossa lainkaan. (Kauppatieteellisen alan yhteisvalinta 2017)

Todistusvalinnassa pisteitä annetaan ylioppilastutkinnon arvosanojen perusteella ohessa olevan liitetaulukon mukaisesti. Todistuspisteitä annetaan ylioppilastutkintoarvosanojen perusteella enintään 40. Todistuspisteitä laskettaessa huomioon otetaan kaikki samaa tutkintoa varten tehdyt kokeet, myös korotetut arvosanat ja ylimääräiset tutkintokokeet. Voidakseen tulla valituksi ylioppilastutkintoarvosanojensa perusteella, on hakijan saatava vähintään 20 todistuspistettä. Koepistevalinnassa valintapisteen muodostuvat valintakokeen kysymyksistä saatavien pisteiden summasta. Vaihtelevaan valintakoekirjallisuuteen perustuvassa valintakokeessa esitetään monivalintakysymyksiä ja oikein/väärin – kysymyksiä, joissa oikeista vastauksista saa pisteitä ja vääristä vähennetään pisteitä (monivalintakysymyksistä +1 tai -0,5 pistettä, oikein/väärin kysymyksistä +0,5 tai -0,25 pistettä). Vastaamatta jättämistä ei ole rajoitettu eikä vastaamatta jättämisestä vähennetä pisteitä. Valintakokeen pisteytysmalli rohkaisee hakijoita taktikoimaan vastauksissaan: koska vääristä vastauksista vähennetään pisteitä, voi vastaamatta jättäminen olla usein vastauksen arvaamista kannattavampaa. Taktikointi ja koepisteiden optimointi näkyvät selkeästi myös tämän tutkimuksen opiskelija-aineistossa: korkeat lähtöpisteet saaneiden

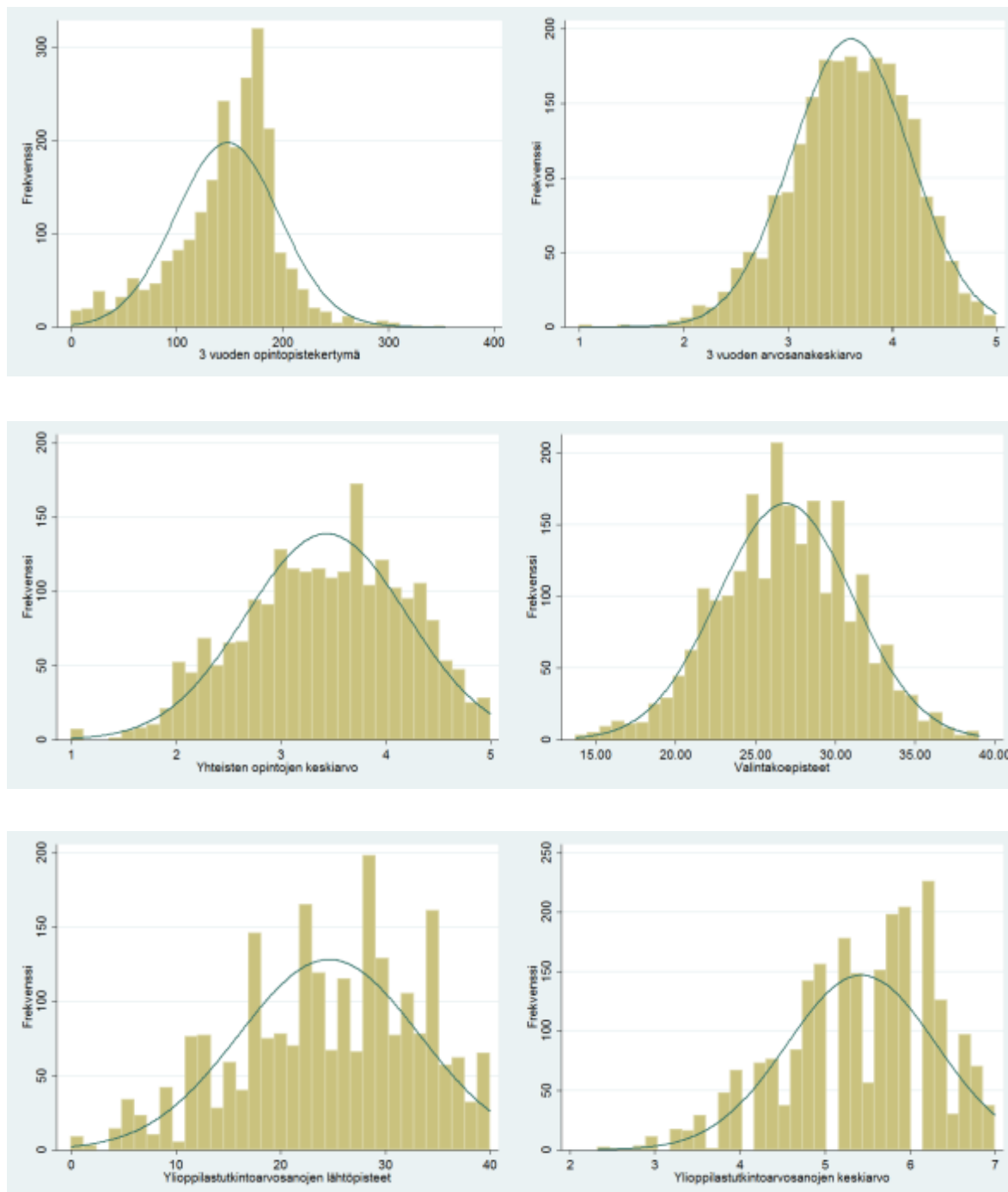
opiskelijoiden valintakoepisteet ovat keskimäärin matalammat kuin matalat lähtöpisteet saaneiden opiskelijoiden (taulukko 4), vaikka hypoteesi ilman taktikointia olisi päinvastainen. Voidakseen tulla hyväksytyksi koepistevalinnassa on hakijan ”*saatava valintakokeesta pisteitä vähintään 60 % valintakokeesta positiivisia pisteitä saaneiden keskiarvosta*”. (Kauppatieteellisen alan yhteisvalinta 2017)

Ylioppilastutkinnon lähtöpisteitys (Kauppatieteellisen alan yhteisvalinta 2017)

	L	E	M	C
Äidinkieli	10	9	7	5
Matematiikka, pitkä oppimäärä	10	9	7	5
Matematiikka, muu oppimäärä	8	6	4	2
A-kieli	8	6	4	2
Kaksi muuta parhaat pisteet antavaa koetta	6	4	2	0

Tässä esiteltyä nykyistä valintamenettelyä sovelletaan tutkielmassa opiskelijavalinnan vaihtoehtoisen valintamenettelyn arvioinnin vertailukohtana luvussa 5.

## Liite 2: Jakaumakuviot



### Liite 3: Persentiilin laskeminen

Persentiilit muodostettiin asettamalla opiskelijat vuosikursseittain lähtöpisteidensä mukaiseen paremmuusjärjestykseen suurimmasta pistemäärästä pienimpään. Yhtä suurille pistemäärille asetettiin sama sijaluku. Lopulta paremmuusjärjestys skaalattiin vuosikursseittain välille 0 – 100 seuraavasti:

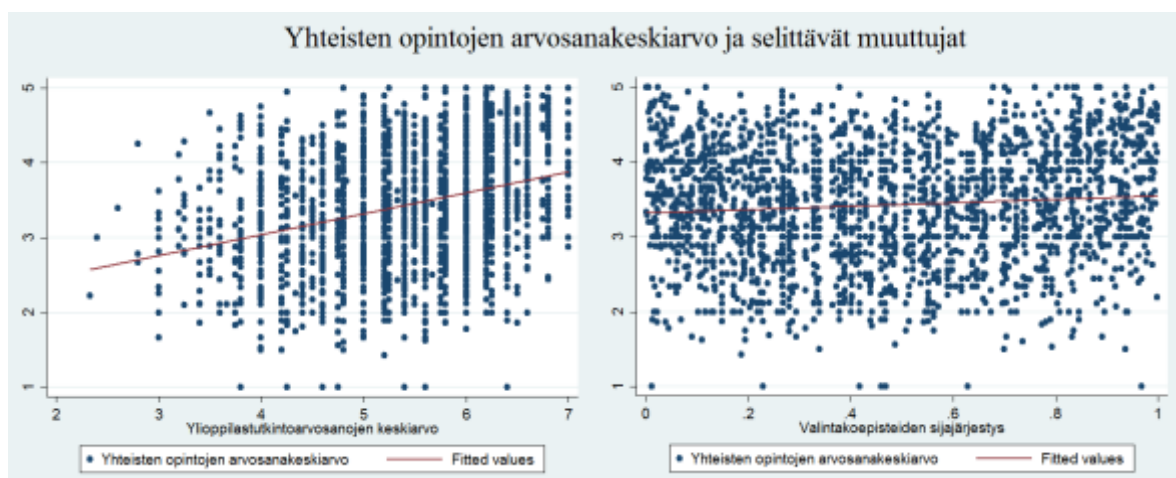
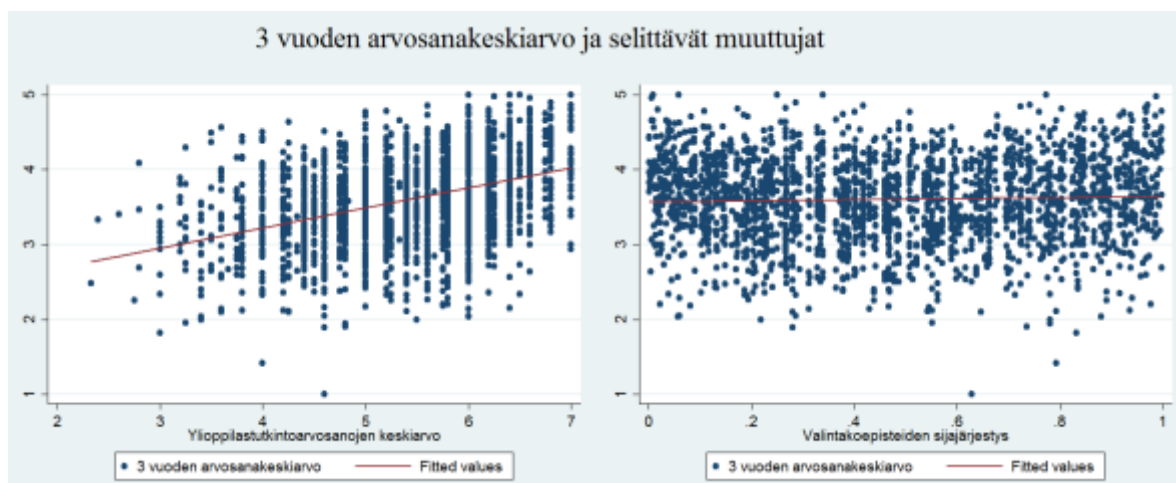
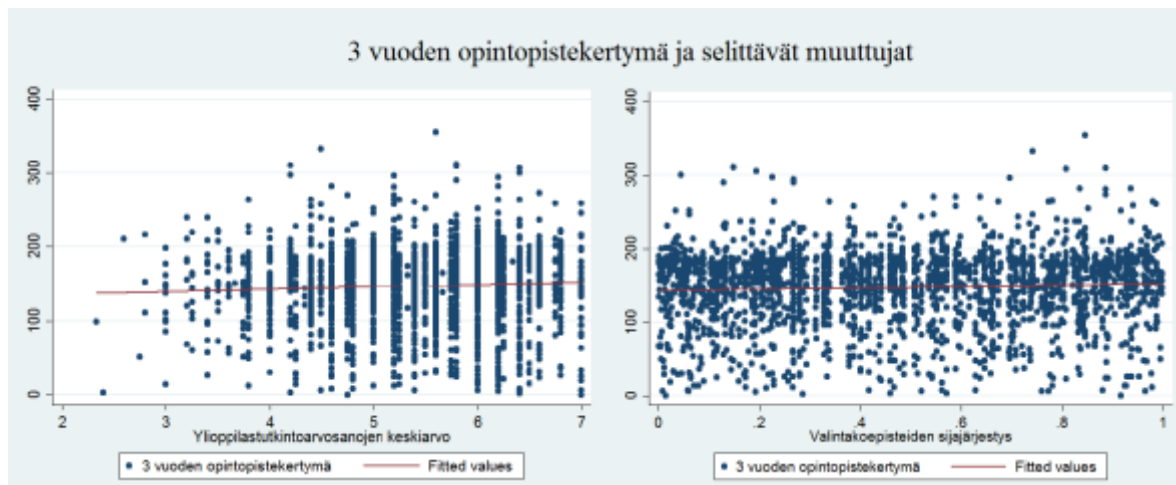
(4)

$$p_s = \frac{s_k - 0.5}{N} * 100$$

Tässä  $p_s$  viittaa skaalattuun persentiiliin,  $s_k$  on todellisten valintapisteiden  $k$  mukainen sijaluku, ja  $N$  valintapisteiden perusteella määritelty suurin sijaluku. Persentiilit laskettiin niin, että saadut persentiilisijat ovat intuitiivisia ja helposti tulkittavissa. Persentiilin laskemisessa alkuperäisestä sijaluvusta vähennetään puolikas sija, jotta parittomalla aineistolla lasketun persentiilin mediaaniksi saadaan 0,5. Ajatellaan esimerkiksi, että valintakokeen on suorittanut 3 opiskelijaa ( $N = 3$ ), joiden valintapisteiden mukaan määräytyvät sijaluvut  $s_k$  ovat vastaavasti 1, 2 ja 3. Yksinkertaisimmalla kaavalla  $s_k/N$  keskimmäisen opiskelijan persentiilisijaksi tulisi  $2/3$ , joka ei kuvaa keskimmäistä sijaa intuitiivisesti ymmärrettävällä tavalla. Kaavalla  $(s_k - 0,5)/N$  keskimmäisen opiskelijan persentiilisijaksi saadaan  $1,5/3 = 0,5$ , joka on intuitiivisesti ymmärrettävämpi. Lopulta saatu persentiilisija on kerrottu sadalla, jotta regressiokertoimien tulkitseminen on suoraviivaisempaa. Kaava käsittelee myös persentiilin päätyarvoja symmetrisesti. Kaavan isä Francis Galton on tunnettu työstään tilastotieteen sekä korrelaatio- ja regressioanalyysin parissa.

Lopputuloksena on valintapisteisiin perustuva persentiili, jossa jokaisen vuosikurssin suurimmat valintapisteet saaneiden opiskelijoiden persentiilisija on 100, ja vastaavasti pienimmät pisteet saaneiden persentiilisija on 0. Muuttuja on jatkuva, standardoitu ja samalla tavoin laskettujen persentiilimuuttujien välillä suoraan vertailukelpoinen. Regressiossa persentiilimuuttujien kertoimia tulkitaan yhden sijan muutoksena: jos opiskelijan persentiilisija kasvaa yhdellä (eli on yhden korkeampi), kasvaa/laskee  $y$  saadun kertoimen verran.

## Liite 4: Hajontakuviot



## Liite 5: Ylioppilastutkintoarvosanojen yhteys arvosanakeskiarvoon

### Ylioppilastutkintoarvosanojen yhteys keskiarvoon

	Äidinkieli	Pitkä matematiikka	Lyhyt matematiikka	Reaaliaineet	Pitkä kieli	Lyhyt kieli
4,0	L					
3,9	L			L		L
3,8	E				L	
3,7	E					E
3,6	M	M		E	EM	
3,5		C	L	M		M
3,4	C	B	EMC	C		C
3,3	BA			B	C	BAI
3,2	I	A		A	BA	
< 3,1		I	BA			

Arvosanakeskiarvo

Taulukko tiivistää ylioppilastutkintoarvosanojen yhteyden opiskelijoiden keskimääräiseen kolmen vuoden arvosanakeskiarvoon. Taulukko on muodostettu suoraan ylioppilaskokeiden arvosanaryhmien arvosanakeskiarvoihin perustuen. Taulukossa ylimpänä olevat arvosanat ovat yhteydessä parhaaseen menestykseen kauppatieteellisissä opinnoissa, ja alimmat arvosanat esiintyvät useimmin heikoiten menestyneillä opiskelijoilla. Keskimäinen, keltainen rivi edustaa Kauppakorkeakoulun keskiarvoisen opiskelijan menestystasoa. Kolmen vuoden arvosanakeskiarvojen keskiarvo on 3,60.



## Liite 6: Keinotekoisien opiskelijavalinnan valintamenettelyt

### **Nykyinen valinta (liite 1):**

Opiskelijoista valitaan ensin 60 % ylioppilastutkinnon lähtöpisteiden paremmuusjärjestyksen perusteella, sitten 40 % koepistevalinnassa. Tasapistetilanteessa valitaan ensin matematiikan arvosanan perusteella enemmän pisteitä saanut, sitten äidinkielen arvosanan perusteella enemmän pisteitä saanut, sitten pitkän kielen (A-kieli) arvosanan perusteella enemmän pisteitä saanut, sitten seuraavaksi parhaan pistemäärän antaneen kokeen perusteella enemmän pisteitä saanut. Jos tasapistesäännön/-sääntöjen jälkeen aloituspaikat ylittyisivät, ketään viimeisen tasapistesäännön mukaisella sijalla olevista ei hyväksytä. (Kauppatieteellisen alan yhteisvalinta 2017)

### **Täysin ylioppilastutkintoon perustuva vaihtoehtoinen valinta:**

100 % opiskelijoista valitaan suorassa todistusvalinnassa tutkielman regressiotulosten perusteella luodun ylioppilastutkinnon vaihtoehtoisen pisteytyksen perusteella (taulukko 13). Tasapistetilanteessa valitaan ensin matematiikan arvosanan perusteella enemmän pisteitä saanut, sitten reaaliaineen, sitten äidinkielen ja lopulta lyhyen kielen perusteella enemmän pisteitä saanut. Tasapistesääntöjen jälkeen tasapisteissä olevien opiskelijoiden valintajärjestyksen valitsee arpa.