JYU DISSERTATIONS 484

Susanne Syrén

Education, Labor Market Outcomes, and Mental Well-Being

JYVÄSKYLÄ UNIVERSITY
SCHOOL OF BUSINESS AND ECONOMICS

JYU DISSERTATIONS 484

Susanne Syrén

Education, Labor Market Outcomes, and Mental Well-Being

Esitetään Jyväskylän yliopiston kauppakorkeakoulun suostumuksella
julkisesti tarkastettavaksi päärakennuksen auditoriossa C2
helmikuun 4. päivänä 2022 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Jyväskylä University School of Business and Economics,
in Main Building, auditorium C2, on February 4, 2022, at 12 o'clock.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2022

Editors
Tuomo Takala
University of Jyväskylä School of Business and Economics
Päivi Vuorio
Open Science Centre, University of Jyväskylä

ISBN 978-951-39-8989-7 (PDF)
URN:ISBN:978-951-39-8989-7
ISSN 2489-9003

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-8989-7

ABSTRACT

Syrén, Susanne
Education, labor market outcomes, and mental well-being
Jyväskylä: University of Jyväskylä, 2022, 117 p.
(JYU Dissertations
ISSN 2489-9003; 484)
ISBN 978-951-39-8989-7 (PDF)

This thesis provides new evidence on education, labor market outcomes, and mental well-being by utilizing various applied microeconometric methods and data sets. The data sets include rarely available measures of discipline problems, personality traits, and mental well-being combined with measures of class size, academic achievement, income, unemployment days, and graduation. The data are drawn from the ProKoulu intervention study, the Jyväskylä Longitudinal Study of Personality and Social Development, application data from centralized admission system in the fields of business and economics, and Finnish administrative data.

The thesis consists of an introduction and three empirical articles. The first article investigates class size effects on discipline problems and academic achievement in primary school. The study shows that larger classes lead to higher levels of discipline problems and that class size seems to negatively associate with academic achievement.

The second article moves on to higher education in early adulthood. It considers whether admission to or enrollment in a more selective institution assures graduation, yields higher income, or results in fewer days of unemployment. The results show that those who gained admission to or enrolled in a more selective business school over a less selective one experience fewer unemployment days and have higher degree completion probability. The effects on income are positive, but at the same time, they do not differ statistically from zero.

The final article of the thesis studies if higher income is associated with higher mental well-being or its dimensions, and if personality traits moderate these associations. The findings indicate that income has a limited role in mental well-being and that personality traits most consistently moderate the relationship between income and a specific dimension of mental well-being, that is, emotional well-being.

Keywords: education, labor market outcomes, mental well-being

TIIVISTELMÄ (ABSTRACT IN FINNISH)

Syrén, Susanne
Koulutus, työmarkkinatulemat, ja henkinen hyvinvointi
Jyväskylä: Jyväskylän yliopisto, 2022, 117 s.
(JYU Dissertations
ISSN 2489-9003; 484)
ISBN 978-951-39-8989-7 (PDF)

Tämä väitöskirja tarjoaa uutta tietoa koulutuksesta, työmarkkinatulemista ja henkisestä hyvinvoinnista hyödyntäen useita soveltavia mikrotaloustieteellisiä menetelmiä ja aineistoja. Aineistot sisältävät muuttujia, kuten työrauha, persoonallisuuden piirteet, ja henkinen hyvinvointi, yhdistettynä tietoihin luokkakoosta, koulumenestyksestä, tuloista, työttömyydestä ja valmistumisesta. Aineistoina hyödynnetään ProKoulu-tutkimuksen aineistoa, Lapsesta aikuiseksi-pitkittäistutkimusaineistoa, kauppatieteellisen alan yhteisvalinta-aineistoa sekä Tilastokeskuksen rekisteriaineistoja.

Väitöskirja koostuu johdantoluvusta sekä kolmesta empiirisestä artikkelista. Ensimmäinen artikkeli tutkii luokkakoon vaikutusta työrauhahäiriöihin ja koulumenestykseen alakoulussa. Tutkimus osoittaa, että suuremmat luokat johtavat työrauhahäiriöiden kasvuun ja näyttäisivät olevan negatiivisesti yhteydessä myös koulumenestykseen.

Toinen artikkeli tarkastelee korkeakoulutusta kauppatieteellisellä alalla. Korkeakoulun laadun vaikutuksia tutkitaan suhteessa työuran alun ansiotuloihin, työttömyyteen sekä valmistumiseen. Tulosten perusteella selektiivisempään kauppakorkeakouluun hyväksytyt ja siellä aloittaneet kokevat vähemmän työttömyyspäiviä ja valmistuvat todennäköisemmin kuin vähemmän selektiiviseen kauppakorkeakouluun valitut tai siellä aloittaneet. Tulojen tapauksessa estimaatit eivät tilastollisesti eroa nollasta, mutta ovat positiivisia.

Väitöskirjan viimeinen artikkeli tutkii, onko korkeampi tulotaso yhteydessä korkeampaan henkiseen hyvinvointiin tai sen eri ulottuvuuksiin. Artikkelissa tarkastellaan myös muokkaavatko persoonallisuuden piirteet tulotason ja henkisen hyvinvoinnin välistä yhteyttä. Tutkimuksessa havaittiin, että tulotason rooli osana henkistä hyvinvointia on rajallinen. Persoonallisuuden piirteet liittyivät lähimmin tulotason ja henkisen hyvinvoinnin väliseen yhteyteen tarkasteltaessa emotionaalisen hyvinvoinnin ulottuvuutta.

Asiasanat: koulutus, työmarkkinatulemat, henkinen hyvinvointi

Author's address

Susanne Syrén
Jyväskylä School of Business and Economics
University of Jyväskylä
susanne.syren@jyu.fi
ORCID 0000-0001-7943-4669

Supervisors

Jaakko Pehkonen
School of Business and Economics
University of Jyväskylä

Roope Uusitalo
University of Helsinki
Helsinki Graduate School of Economics

Reviewers

Kristiina Huttunen
Aalto University
Helsinki Graduate School of Economics

Tuomo Suhonen
Labour Institute for Economic Research

Opponent

Kristiina Huttunen
Aalto University
Helsinki Graduate School of Economics

ACKNOWLEDGEMENTS

I'd heard it so many times. "If I had known what I was getting into, I would have never started". I now understand what that sentence really meant, but at the time of entering the doctoral program, lovely unawareness about what completing a PhD in economics requires, accompanied me to the first year of studies. Eight years later, it feels surreal that this journey is coming to an end. It is time to thank all the people who helped me to reach this point.

First, I want to acknowledge Yrjö Jahnsson Foundation, OP Group Research Foundation and Jyväskylä University School of Business and Economics (JSBE) for the financial support. Without it, completing this dissertation would not have been possible.

I'm deeply grateful to my supervisors Professor Jaakko Pehkonen and Professor Roope Uusitalo for support and guidance during my studies. Jaakko encouraged me to apply to the doctoral program, and during this process, his door has always been open for questions. Jaakko has an extensive knowledge in economics and in supervising the doctoral students, which shows in his work. Similarly, Roope has always devoted time to discuss my research. His encouragement and expertise in economics have without a doubt helped me to increase the quality of my thesis. Finally, I want to thank both of my supervisors for understanding and patience while I've tried to find my place in this world.

I warmly thank Associate Professor Kristiina Huttunen and Research Director Tuomo Suhonen for reviewing my dissertation and for providing helpful and constructive comments. I'm honored to have Kristiina Huttunen as my opponent at the public examination. I further wish to express my gratitude to Professor Kjell Salvanes and Professor Aline Bütikofer for my stay in Norwegian School of Economics (NHH). Thank you also for all the other staff members and graduate students for warmly welcoming me to NHH.

I've been privileged to work with number of researchers from different fields of study. I'm grateful to my co-authors, Katja Kokko, Lea Pulkkinen, Olli-Pekka Malinen and Vesa Närhi for their contributions to the articles. With your encouragement and help, I've become a better writer and researcher. Thank you also for the colleagues in University of Jyväskylä Open University, especially the members of "Supertiimi". The work community in Open University is simply wonderful, and I appreciate the support provided for me when finalizing the thesis.

During these years, I've met many great people. Thank you Salla and Lauri for sharing the ups and downs during the KAVA-courses. Despite the heavy workload, I have many good memories from my first year as a doctoral student. I'm truly thankful for Jaana Kari for the invaluable support and mentoring throughout my studies. I also wish to thank Juho Jokinen, Matthias Strifler, Ville Seppälä, Antti Sieppi, Samu Kärkkäinen, Kristian Koerselman, Elina Riivari, Jussi Huuskonen, Jussi Leskinen, Vu Thanh Nam and all the other doctoral students and faculty members in JSBE. It has been a pleasure to work with you, but also to spend time in all kinds of activities outside the office. I'm further grateful for

the sport pedagogy students (year 2018, group 3) who I've been lucky to study with. You have brought so much lightness and joy to my days, and with you, it has been easy to remember that there is more in life than just research.

Finally, I want to express my gratitude to my parents Kirsi and Ilkka, who strongly believe in education. To my sisters Annmari and Johanna for endless love and care. To my brother Jaakko for insightful discussions and much needed realistic points of view. And to all other friends and family, who shared this journey with me. Never would I have been able to complete this dissertation without you.

Jyväskylä, December 2021
Susanne Syrén

LIST OF ORIGINAL PUBLICATIONS

Article 1. pp. 25–55
Syrén, S. M., Uusitalo, R., Malinen, O., and Närhi, V. Effects of Class Size on Discipline Problems and Academic Achievement. (Unpublished).

Article 2. pp. 56–86
Syrén, S. M. Business school quality and early career outcomes. (Unpublished).

Article 3. pp. 87–117
Syrén, S. M., Kokko, K., Pulkkinen, L., & Pehkonen, J. (2020). Income and Mental Well-Being: Personality Traits as Moderators. *Journal of Happiness Studies*, 21(2), 547-571.

CONTENTS

1 INTRODUCTION

Beginning with primary school, continuing to early adulthood, and finally to middle age, this thesis provides evidence on education, labor market outcomes, and mental well-being. The thesis consists of an introductory chapter (Chapter 1) and three empirical articles (Chapters 2–4). In the articles, a variety of applied econometrics methods, such as ordinary least square (OLS), fixed effects estimation (FE), instrumental variables method (IV), and regression discontinuity design (RDD), are utilized. The employed data sets include both qualitative and quantitative measures and are drawn from the ProKoulu intervention study, the Jyväskylä Longitudinal Study of Personality and Social Development, application data from centralized admission system in the fields of business and economics, and Finnish administrative data.

The first article explores whether class size affects discipline problems and academic achievement, and the second article concentrates on business school quality and labor market outcomes. The final article of the thesis investigates the association between income and mental well-being and discusses whether the marginal utility of income on mental well-being is dependent on personality traits.

The rest of the introductory chapter is organized as follows. The first section presents the institutional setting and earlier literature. The institutional setting discussion concentrates on the education system in Finland, and the earlier literature summarizes literature on education, labor market outcomes, and mental well-being relevant to the thesis. The second section of the introduction provides an overview of the thesis. It presents the research questions, data and methods, main findings, and contributions of the thesis. Limitations and suggestions for future research together with concluding remarks end the introductory chapter. Chapters 2–4, following the introductory chapter, contain the empirical articles.

1.1 Institutional setting and earlier literature

1.1.1 Education in Finland

Education in Finland starts with early childhood education and care and can be continued up to higher education. The present thesis concentrates on comprehensive school and higher education. These parts of the education system are described in more detail, giving less attention to the rest of the education system. Figure 1 summarizes the education system as a whole and gives a starting point for the discussion of the institutional setting.

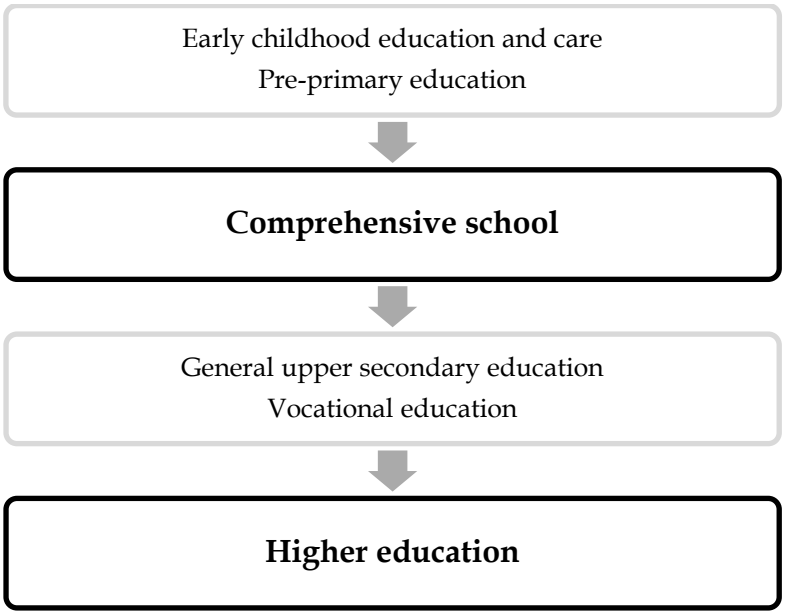

FIGURE 1 The education system in Finland

Early childhood education and care are optional and can also be conducted by parents themselves taking care of the child at home. Early childhood education and care take place after the child turns nine or ten months old and continues until the child turns six. At the age of six years, the child starts one year of compulsory pre-primary education.

Compulsory comprehensive school follows pre-primary education and lasts for nine years. In Chapter 2, data is included from primary education, which covers the first six years of comprehensive school. Primary school is followed by three years of comprehensive middle school. Compulsory schooling is most often organized by municipalities, and all the schools follow a national core curriculum. Compulsory pre-primary education and comprehensive school are free of charge, and unlike educational levels following comprehensive school, include free learning materials, school meals, healthcare services, and transport to school when the route to school is long or unsafe.

Three-year general or vocational upper secondary education (also a combination of general and vocational upper secondary education is possible) follows comprehensive school. Due to the educational reform in 2021, upper secondary education is now part of compulsory education and includes all the benefits of compulsory education. Upper secondary education graduates are eligible to apply for admission to higher education institutions.

Higher education in Finland is provided by universities and universities of applied sciences, and the qualifications of higher education correspond to levels 6, 7, and 8 in the European Qualification Framework (EQF). Higher education is free of charge but does not include free learning materials (except for course book libraries), meals, healthcare, or transport as compulsory education does. Students do, however, receive student aid and loan as well as housing and meal subsidies, and there is an inexpensive, mandatory student health service.

Universities provide the highest possible level of education in Finland and conduct scientific research. University students graduate with bachelor's, master's, and doctoral degrees. Universities of applied sciences concentrate on vocational education and applied research for working life needs. Students from universities of applied sciences graduate with bachelor's and/or master's degrees. Higher education is considered in Chapter 3 from a university-level business education perspective.

1.1.2 Earlier literature

1.1.2.1 Class size effects

The literature on class size has concentrated on academic achievement and labor market outcomes, showing that class size is related to school achievement, cognitive and non-cognitive skills, education completion, and wages (Angrist & Lavy, 1999; Fredriksson et al., 2013; Krueger, 1999). However, more recent studies have suggested that there are insignificant or zero class size effects on academic achievement and labor market outcomes (Angrist et al., 2019; Falch et al., 2017; Hanushek, 1997; Hoxby, 2000; Kupiainen & Hienonen, 2016; Leuven et al., 2008; Leuven & Løkken, 2020). Most of the studies cited above have been conducted using quasi-experimental methods; however, the conclusions about the causal connections between class size and academic achievement or labor market outcomes have remained inconclusive.

Another ambiguous aspect in the class size literature is the question of the mechanisms behind the class size effects. Lazear (2001) proposed in his theoretical model that increasing class size increases discipline problems, and thus offered one possible mechanism behind class size effects on academic achievement. Earlier literature has illustrated descriptive evidence that supports Lazear's (2001) proposition. Bigger classes are associated with distractions from work and tasks (Blatchford et al., 2003) and with more time spent on discipline (Betts & Shkolnik, 1999) and less on teaching (Blatchford et al., 2003). These discipline issues and potentially disruptive peers are further associated with lower academic achievement (Carrell & Hoekstra, 2010; Erätuuli & Puurula, 1992;

Horoi & Ost, 2015; Kristoffersen et al., 2015). The literature investigating the mechanism behind the class size effects is mainly descriptive, and therefore there is a need for causal evidence on this matter.

1.1.2.2 Institution quality and labor market outcomes

The institutional quality literature addresses the question of whether institutional quality affects subsequent labor market outcomes. Typically, the existing literature has studied the gains from elite institution admissions in settings where there are clear differences in institutional quality, measured, for example, by selectivity. Another line of the literature has tried to separate possible gains attained from different fields of study, while controlling for institutional quality. These two lines of the literature provided a basis for the institutional quality study in the present thesis, where the effect of gaining admission to or enrolling in a more selective institution over a less selective one was studied in the fields of business administration and economics.

According to the existing literature, admission to an elite institution or to the most selective state university yields higher earnings (Anelli, 2020; Hoekstra, 2009), and a higher probability of completing a university degree (Anelli, 2020; Kuuppelomäki et al., 2019) and being employed (Saavedra, 2009). Statistically insignificant quality estimates on earnings were reported by Öckert (2010) in Swedish setting. Some studies have suggested that the admission to a more selective institution could benefit certain subgroup of applicants. In a Finnish research setting that involved engineers, Kuuppelomäki et al. (2019) found an increase in earnings for students whose parents were not highly educated, but for others, the income and employment gains did not differ statistically from zero. Further, Dale and Krueger (2002) showed that students from low-income families benefit from attending to a more selective college in terms of higher earnings, and Zimmerman (2019) illustrated how admission to a business-focused elite degree program raised the number of held leadership positions, but only for male students from private high schools. Considering the field of study effects, Kirkeboen et al. (2016) suggested that in Norway, there are clear differences in payoffs from different fields and that the field of study matters more than institution effects in terms of payoffs. Therefore, controlling for the field of study or studying a specific field is important when analyzing the institutional quality effects.

1.1.2.3 Income, personality, and mental well-being

The economics literature concerning the relationship between income and well-being covers a wide range of countries and time periods. Income has often been measured in terms of gross domestic product (GDP), household income, or as individual's personal income, and the most studied well-being concepts in economics literature have been happiness and life satisfaction. Earlier research found a positive relationship between GDP and happiness in the short run (Deaton, 2008; Stevenson & Wolfers, 2008), but not between GDP and life satisfaction in the long run (Easterlin et al., 2010). Income is further positively associated with

happiness, life satisfaction, financial satisfaction, and well-being, and negatively associated with financial stress (Angeles, 2011; Boyce & Wood, 2011; Brown & Gray, 2016; Headey & Wooden, 2004). In addition to income or GDP, earlier literature has shown that wealth matters for subjective well-being and ill-being (Headey & Wooden, 2004). The earlier literature lacks, however, evidence on the relationship between income and other dimensions of mental well-being or mental well-being as a whole.

Big Five personality traits (agreeableness, extraversion, conscientiousness, openness to new experiences, and neuroticism; Costa and McCrae 1985) are directly linked to mental well-being (Kokko et al., 2015; Kokko, Tolvanen, et al., 2013), but they also modify the association between income and mental well-being. Therefore, including personality traits and acknowledging possible changes in personality traits is important in investigations into the relationship between income and mental well-being. Earlier literature has proposed that neuroticism moderates the association between income and life satisfaction (Proto & Rustichini, 2015; Soto & Luhmann, 2013). Boyce and Wood (2011) found that in addition to neuroticism, openness to new experiences and introversion moderated the household income–life satisfaction relationship for women. In their study, conscientiousness was the only Big Five personality trait that moderated the relationship between household income and life satisfaction for both men and women. There is a need for more evidence on the moderating role of personality traits on the income–mental well-being association, especially in settings where personality traits can vary over time. Further, it is important to verify whether personality traits positively or negatively moderate the relationship between income and mental well-being. At present, the evidence on this matter is mixed.

1.2 Overview of the thesis

This section provides an overview of the thesis. The research questions are first presented in order of the empirical articles following the introductory chapter. The next subsection continues with the data and the methods used in answering the research questions. The main findings and contributions, limitations, and suggestions for future research, together and concluding remarks, complete the introductory chapter.

1.2.1 Research questions

1. Does class size affect discipline problems and further associate with academic achievement?

The first article (Chapter 2) concentrates on class size, a much-debated educational component. Reducing class size is expensive, and separating the benefits from the costs is not always straightforward because schools may

decrease class size in hard-to-teach classes or place hard-to-teach students in smaller classes. Further, the mechanisms behind the benefits of the smaller classes are largely unknown. The first article provides evidence on one possible mechanism, discipline problems, behind the class size effects on academic achievement and labor market outcomes illustrated by the earlier literature (Angrist & Lavy, 1999; Angrist et al., 2019; Fredriksson et al., 2013; Hoxby, 2000; Krueger & Whitmore, 2000; Leuven et al., 2008) and further studies the association between class size and academic achievement.

2. Does admission or enrollment in a more selective institution assure graduation, or result in higher income, or fewer days of unemployment?

The second article (Chapter 3) explores another important educational policy aspect, institutional quality. Institutional quality is studied in a business school setting, and the article examines whether there are returns (graduation, higher income, and fewer days of unemployment) arising from admission to or enrollment in a better business school.

3. Does higher income associate with higher mental well-being and do personality traits moderate this association?

The final article (Chapter 4) of the thesis concentrates on income and mental well-being. The association between income and mental well-being is studied by using a broad measure of mental well-being (containing emotional, psychological, and social well-being, and the absence of ill-being) and by taking personality traits into account. Further, the article addresses whether personality traits moderate the association between income and mental well-being.

1.2.2 Data and methods

The first research question concerns class size, discipline problems, and academic achievement. The data used in answering this question were from the ProKoulu project, which was an intervention study conducted in 2013–2016. In Eastern Finland, all primary schools received an offer to participate in the ProKoulu intervention study. The effectiveness of the School Wide Positive Behavior and Support-Model (Sugai & Horner, 2002) on student behavior was tested in schools that took part in the ProKoulu project. The project collected detailed information about classroom behavior from teachers and students as well as other school- and class-specific information, which we utilized in our study.[1]

[1] The intervention provided support in enhancing positive student behavior in all the participating schools. The participating schools were randomized into control and treatment groups, and the positive student behavior intervention was delayed by one year in the control group. We used the information about the classroom behavior together with school- and class-specific information, and we controlled for the intervention delay in the control group but did not study the intervention itself.

The ProKoulu data were combined with student achievement data, which were collected with the help of the ProKoulu research team after the original intervention. In Finland, there are no national registers or standardized tests in primary school, which is why the achievement data collection required on-site visits to municipality school administration offices. We managed to collect student achievement data from 37 out of 60 participating schools and linked this data to the original ProKoulu data.

Class size effects on discipline problems and academic achievement were investigated with the IV method. We followed Angrist and Lavy (1999) in constructing an instrument, which was based on the maximum class size guideline of having a maximum of 25 students in class. The instrument affected class size, but it was unlikely to affect discipline problems or academic achievement through channels other than class size. The IV method overcame the endogeneity of class size and therefore provided causal evidence on class size effects.

The second research question addresses the differences in labor market outcomes arising from institutional quality. To provide causal evidence on the matter, we utilized application data from a centralized admission system in the fields of business and economics from 2005 to 2008 and merged it with Finnish administrative data up to eight years after observed admission. From 2005 to 2008, a total of nine business schools were part of the centralized admission system in the fields of business and economics, meaning that these institutions applied the same admission process and selection criteria in student selection.

The application data included applicant-specific information about entrance examination points, preference order of the institutions applied to, and information about admission and enrollment. The applicant-specific information was combined with early career labor market outcomes (income, unemployment) and graduation information obtained from the administrative data. Admission cutoffs were utilized in a RDD setting, which overcame the selection bias arising from unobserved applicant-specific characteristics. The admission cutoffs between business schools were stacked in estimations, which allowed estimating a single admission effect (describing the effect of being admitted to a more selective business school). In addition to a sharp RDD setting, we also applied fuzzy RDD in estimating the effect of enrolling in a better business school on early career labor market outcomes and graduation.

The final article answers the question of whether income is associated with mental well-being and whether personality traits moderate this association. The data used in the study were obtained from the Jyväskylä Longitudinal Study of Personality and Social Development (JYLS). JYLS has been ongoing since 1968, and we used the data collections from years 2001 and 2009, when the participants were aged 42 and 50, respectively. These two data collection phases were pooled in the analyses. With JYLS, we were able to analyze information about mental well-being and its dimensions (emotional, psychological, and social well-being, and the absence of depression), the Big Five personality traits (agreeableness, extraversion, conscientiousness, openness to new experiences, and neuroticism),

experienced household financial situation, and participants' own gross monthly income.

Our empirical strategy exploited the longitudinal form of the data. The methods used in addressing research question 3 were pooled OLS and FE estimations. OLS reveals the cross-sectional variation in the data, whereas FE utilizes within-individual variation and eliminates the bias arising from unobserved individual-specific characteristics, which are constant over time.

1.2.3 Main findings and contributions

The first empirical article concentrates on class size, discipline problems, and academic achievement in primary school. The main findings of the article were that larger class size leads to higher levels of discipline problems and that class size is negatively associated with academic achievement. By decreasing class size or by using appropriate classroom management strategies (e.g., see Korpershoek et al., 2016) the behavioral outcomes may be improved and learning enhanced. The contributions of the article are that we provided new evidence on largely unknown mechanisms behind class size effects by using a unique data set, which included information about school, classes, students, and teachers, and linked this information to observed discipline problems. It is also rare to have data on academic achievement from primary schools in Finland, which we were able to obtain. Finally, the Finnish setting allowed us the use of a credible quasi-experimental method, enabling causal inference.

The research for the second article moved from primary school to higher education. The results indicated that gaining admission to or enrolling in a more selective business school leads to fewer days of unemployment and increases the degree completion probability. The estimates of the effects of admission or enrollment on income were positive but did not statistically differ from zero. Studying institutional quality can be challenging, as some applicant characteristics, such as ability, are not observable by the researcher. The selection bias arising from such unobserved characteristics was however eliminated by following the earlier literature (e.g., Abdulkadiroglu et al., 2014; Hoekstra, 2009; Kirkeboen et al., 2016; Kuuppelomäki et al., 2019), which used quasi-experimental research settings. Compared with the settings discussed in the earlier literature, the present setting was straightforward. The same entrance exam is used in all studied business schools, and all students graduated with the same Master of Science in Economics and Business Administration degree. Further, concentrating on a specific field allowed us to control for the field of study effects. Finally, this study added to the literature by examining the effect of attending a more selective institution over a less selective one, instead of concentrating only on elite institutions.

The final article of the thesis concerns income and mental well-being in middle age. According to the results, the role of income on mental well-being as a whole is limited; instead, income mainly associated with emotional well-being. Similarly, we found that personality traits most consistently moderate the relationship between income and emotional well-being. The contribution of this

article comes, firstly, from a broad measure of mental well-being. The earlier economics literature concentrated on happiness or life satisfaction, whereas the present study investigated different dimensions of mental well-being (emotional, psychological, and social well-being, absence of ill-being) as well as overall mental well-being. Secondly, we were able to use rarely available measures on personality traits from two measurement points, which allowed us to analyze changes in the Big Five personality traits over time.

1.2.4 Limitations and suggestions for future research

The data sets used in the present thesis are exceptional in the sense that they merge qualitative and quantitative measures. Further, they included rarely available measures, such as discipline problems, mental well-being, and the Big Five personality traits. Obtainment of these unique data sets and the empirical analyzes conducted with them undoubtedly belong to the main contributions of the thesis, but the data also have its limitations.

The data sets were moderately sized, which is understandable when data include qualitative measures and are from a country with a low population density such as Finland. Another possible limitation is the generalizability of the results. Class size effects were studied with the data from the intervention study, which aimed at supporting positive student behavior. As positive student behavior was supported in the studied schools, it is possible that the level of discipline problems was lower in our data than in Finnish classes in general. Similarly, the institutional quality study concentrated on the fields of business and economics, but it would have been equally as important to understand if the institutional quality matters in other fields as well. Finally, the income and mental well-being study concerned middle-aged Finnish people, who on average, have fairly good financial situations. The results in other countries or age groups may differ from the results of the present thesis.

The present thesis suggests interesting avenues for future research. First, it would be beneficial to conduct a cost–benefit analysis related to class size reductions and discipline problems. This analysis could investigate whether the benefits of class size reductions exceed the costs, and whether it would be more beneficial to apply classroom management strategies than reduce class size when aiming for lower levels of discipline problems and enhanced learning. Further, institutional quality research on labor market outcomes could be expanded to fields of study other than business administration and economics, and the mechanisms behind the differences in experienced unemployment days or graduation probabilities among the students from different institutions could be studied. Future research on income and mental well-being could extend the analyses on other countries and age groups. It would also be interesting to investigate if income - mental well-being associations differ by gender. Finally, additional contributions in research concentrating on the moderating role of personality could be achieved by investigating personality profiles instead of separate personality traits, as the personality traits are likely to operate in combination.

1.2.5 Concluding remarks

This thesis provides evidence on education, labor market outcomes, and well-being. From a policy perspective, the results presented in the thesis have the following implications. First, it should be noted that increasing class size increases discipline problems, even in schools where positive student behavior is supported. Higher levels of discipline problems can put a strain on students and teachers. Further, it is expected that academic achievement will be jeopardized in classes with high levels of discipline problems. As decreasing class size is expensive, cost–benefit analyses are an important part of decision-making. In these analyses, it should be kept in mind that discipline problems may also be reduced through revisiting the classroom behavioral expectations, teaching the expected behaviors, and reinforcing appropriate behavior (Epstein et al., 2008), not only through class size reductions.

The second policy implication of the thesis relates to higher education. In Finland, student selection to higher education has been an ongoing debate for quite some time. In the best-case scenario, the student selection procedure channels the most competent and suitable applicants to available slots, without a delay. The delays do occur, however, if students reject the offers from or do not apply to less selective institutions in the hope of gaining admission to a more selective institution. The present study found that there are achievable gains from being admitted to or enrolling in a better institution in terms of fewer unemployment days and increased degree completion probability. When designing the admission systems, these gains that may postpone the beginning of the studies should be noted. The results of the study are relevant for the applicants considering were to apply to, as well as for the policymakers allocating the number and location of the available slots.

Finally, the thesis suggests that income is associated with mental well-being but that the positive relationship is limited. Personality traits moderate this relationship, and together with other factors, such as health or employment, personality traits are likely to contribute to mental well-being. It is therefore not only income that matters in terms of mental well-being. When designing policies that aim to increase well-being, the complexity and dimensionality of well-being is important to keep in mind.

References

Anelli, M. (2020). The returns to elite university education: A quasi-experimental analysis. *Journal of the European Economic Association*, *18*(6), 2824–2868. https://doi.org/10.1093/jeea/jvz070

Angeles, L. (2011). A closer look at the Easterlin Paradox. *Journal of Socio-Economics*, *40*(1), 67–73. https://doi.org/10.1016/j.socec.2010.06.017

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' Rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, *114*(2), 533–575. https://doi.org/10.1162/003355399556061

Angrist, J. D., Lavy, V., Leder-Luis, J., & Shany, A. (2019). Maimonides' Rule redux. *American Economic Review: Insights*, *1*(3), 309–324. https://doi.org/10.1257/aeri.20180120

Betts, J. R., & Shkolnik, J. L. (1999). The behavioral effects of variations in class size: The case of math teachers. *Educational Evaluation and Policy Analysis*, *21*(2), 193–213. https://doi.org/10.3102/01623737021002193

Blatchford, P., Bassett, P., Goldstein, H., & Martin, C. (2003). Are class size differences related to pupils' educational progress and classroom processes? Findings from the Institute of Education Class Size Study of Children Aged 5–7 Years. *British Educational Research Journal*, *29*(5), 709–730. https://doi.org/10.1080/0141192032000133668

Boyce, C. J., & Wood, A. M. (2011). Personality and the marginal utility of income: Personality interacts with increases in household income to determine life satisfaction. *Journal of Economic Behavior and Organization*, *78*(1–2), 183–191. https://doi.org/10.1016/j.jebo.2011.01.004

Brown, S., & Gray, D. (2016). Household finances and well-being in Australia: An empirical analysis of comparison effects. *Journal of Economic Psychology*, *53*, 17–36. https://doi.org/10.1016/j.joep.2015.12.006

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, *2*(1), 211–228. https://doi.org/10.1257/app.2.1.211

Costa, P. T., & McCrae, R. R. (1985). The NEO Personality Inventory. Odessa, FL: Psychological Assessment Resources.

Dale, S. & Krueger, A. (2002). Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *The Quarterly Journal of Economics, 117*(4). 1491-1527. https://doi.org/10.1162/003355302320935089

Deaton, A. (2008). The World: Evidence from the Gallup. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, *22*(2), 53–72. https://doi.org/10.1257/jep.22.2.53

Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness–income paradox revisited. *Proceedings of the National*

Academy of Sciences of the United States of America, *107*(52), 22463–22468. https://doi.org/10.1073/pnas.1015962107

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, K. (2008). *Reducing behavior problems in the elementary school classroom*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education.

Erätuuli, M., & Puurula, A. (1992). *Miksi häiritset minua: Opettajan näkökulma työrauhahäiriöihin yläasteella [Why do you disturb me? Disruptive behavior experienced by lower secondary school teachers]*. Yliopistopaino.

Falch, T., Sandsør, A. M. J., & Strøm, B. (2017). Do smaller classes always improve students' long run outcomes? *Oxford Bulletin of Economics and Statistics 79*(5), 654-688. https://doi.org/10.1111/obes.12161

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *The Quarterly Journal of Economics*, *128*(1), 249–285. https://doi.org/10.1093/qje/qjs048

Hanushek, E. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, *19*(2), 141–164.

Headey, B., & Wooden, M. (2004). The effects of wealth and income on subjective well-being and ill-being. In *Economic Record* (Vol. 80, Issue SPEC. ISS.). https://doi.org/10.1111/j.1475-4932.2004.00181.x

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, *91*(4), 717–724. https://doi.org/10.1162/rest.91.4.717

Horoi, I., & Ost, B. (2015). Disruptive peers and the estimation of teacher value added. *Economics of Education Review*, *49*, 180–192. https://doi.org/10.1016/j.econedurev.2015.10.002

Hoxby, C. M. (2000). The effects of class size on student achievement : New evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239–1285. https://www.jstor.org/stable/2586924

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). Field of study, earnings, and self-selection. *Quarterly Journal of Economics*, *131*(3), 1057–1112. https://doi.org/10.1093/qje/qjw019.Advance

Kokko, K., Korkalainen, A., Lyyra, A. L., & Feldt, T. (2013). Structure and continuity of well-being in mid-adulthood: A longitudinal study. *Journal of Happiness Studies*, *14*(1), 99–114. https://doi.org/10.1007/s10902-011-9318-y

Kokko, K., Rantanen, J., & Pulkkinen, L. (2015). Associations between mental well-being and personality from a life span perspective. In M. Blatný (Ed.), *Personality and well-being across the life-span* (pp. 134–159). Palgrave Macmillan.

Kokko, K., Tolvanen, A., & Pulkkinen, L. (2013). Associations between personality traits and psychological well-being across time in middle adulthood. *Journal of Research in Personality*, *47*(6), 748–756. https://doi.org/10.1016/j.jrp.2013.07.002

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kristoffersen, J. H. G., Krægpøth, M. V., Nielsen, H. S., & Simonsen, M. (2015). Disruptive school peers and student outcomes. *Economics of Education Review*, *45*, 1–13. https://doi.org/10.1016/j.econedurev.2015.01.004

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, *114*(2), 497–532. https://doi.org/10.1162/003355399556052

Krueger, A. B., & Whitmore, D. M. (2000). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR. *NBER Working Paper Series*, *7656*.

Kupiainen, S., & Hienonen, N. (2016). *Luokkakoko [Class size]*. Suomen kasvatustieteellinen seura.

Kuuppelomäki, T., Kortelainen, M., Suhonen, T., & Virtanen, H. (2019). Does admission to elite engineering school make a difference? *VATT Working Papers, 146*. https://www.doria.fi/handle/10024/172697

Lazear, E. P. (2001). Educational production. *Quarterly Journal of Economics*, *116*(3), 777–803. https://doi.org/10.1162/00335530152466232

Leuven, E., & Løkken, S. A. (2020). Long-term impacts of class size in compulsory school. *Journal of Human Resources*, *55*(1), 309–348. https://doi.org/10.3368/jhr.55.2.0217.8574r2

Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *The Scandinavian Journal of Economics*, *110*(4), 663–693. https://doi.org/10.1111/j.1467-9442.2008.00556.x

Proto, E., & Rustichini, A. (2015). Life satisfaction, income and personality. *Journal of Economic Psychology*, *48*, 17–32. https://doi.org/10.1016/j.joep.2015.02.001

Saavedra, J. E. (2009). The learning and early labor market effects of college quality: A regression discontinuity analysis [Unpublished manuscript]. https://economia.uniandes.edu.co/sites/default/files/seminariocede/The-Returns-to-College-Quality.pdf

Soto, C. J., & Luhmann, M. (2013). Who can buy happiness?: Personality traits moderate the effects of stable income differences and income fluctuations on life satisfaction. *Social Psychological and Personality Science*, *4*(1), 46–53. https://doi.org/10.1177/1948550612444139

Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox. *Brookings Papers on Economic Activity*, *2008*(SPRING), 1–87. https://doi.org/10.2139/ssrn.1121237

Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy*, *24*(1-2), 23–50. https://doi.org/10.1300/J019v24n01_03

Zimmerman, S. (2019). Elite Colleges and Upward Mobility to Top Jobs and Top Incomes. *American Economic Review, 109*(1). 1–47. https://doi.org/10.1257/aer.20171019

Öckert, B. (2010). What's the value of an acceptance letter? Using admissions data to estimate the return to college. *Economics of Education Review*, *29*(4), 504-516. https://doi.org/10.1016/j.econedurev.2009.12.003

2 CLASS SIZE EFFECTS ON DISCIPLINE PROBLEMS AND ACADEMIC ACHIEVEMENT*

Abstract**

We examined causal effects of class size on discipline problems in primary school by exploiting exogenous variation in class size due to maximum class size guidelines. The results showed that adding one student in the class increases discipline problems by 0.085 standard deviation units according to the teacher reports, and by 0.041 standard deviation units according to the student reports. Using the same framework we found signs of a negative associations between class size and academic achievement, but no evidence on causal effects of class size on grades in mother tongue, mathematics or first foreign language.

2.1 Introduction

The effect of class size on academic achievement and labor market outcomes has been studied extensively (Angrist et al., 2019; Angrist & Lavy, 1999; Chetty et al.,

* This chapter is a joint article with Roope Uusitalo, Olli-Pekka Malinen and Vesa Närhi.
The authors would like to acknowledge Ville Vauhkonen and Elisa Korhonen for their invaluable work in the collection and management of data used in this study. The authors are further grateful to seminar participants at universities of Jyväskylä, Tampere and Helsinki, Labour Institute for Economic Research, Norwegian School of Economics, as well as to participants at Workshop on Education, Skills, and Labor Market Outcomes (2016), European Conference on Educational Research (2017), and at IWAEE International Workshop on Applied Economics of Education (2019) for the helpful comments.
** Syrén gratefully acknowledges the financial support from the OP Group Research Foundation (project numbers 201600189, 20170093 and 20180102; Helsinki, Finland), and Uusitalo from Strategic Research Council of Academy of Finland (grant 314 314; Helsinki, Finland). Närhi and Malinen would like to express their gratitude for the financial support received from the Finnish Ministry of Education and Culture (grant number 35/529/2012; Helsinki, Finland), through which the Pro-School project was funded. Malinen is also grateful for the financial support provided by the Postdoctoral Research Grant of Academy of Finland (grant 274 483; Helsinki, Finland).

2011; Falch et al., 2017; Fredriksson et al., 2013; Hoxby, 2000; Krueger & Whitmore, 2001; Leuven et al., 2008; Leuven & Løkken, 2020). While most studies have found that smaller classes are beneficial for learning, the mechanisms are still largely a black box. In the present study, we studied one possible mechanism: the effect of class size on discipline problems in the classroom. We further examined how the class size affects academic achievement in the same data.

Our empirical approach was inspired by Lazear (2001), who presented a conceptual model for understanding the effects of class size on student achievement. In the Lazear model, each student has a probability p of behaving well. If any student in the class misbehaves, the entire class suffers, and this affects the learning. The likelihood that all students are behaving well is p^n, where n describes the class size. Disruptions occur $1-p^n$ of the time.

The Lazear model has several interesting implications. In addition to the result that increasing class size increases disruptions, the model also suggests that it is optimal to reduce class size when the students are less-well behaved. Lazear (2001) also showed that under reasonable assumptions, segregation of students and placing less-well behaved students to smaller classes is an optimal response from the schools. This implies that class size is endogenous and that estimating the effects of class size from observational data is likely to lead into biased conclusions with respect to effects of class size on student behavior, and eventually on student achievement.

There is a small empirical literature examining why class size affects academic achievement (see Finn, Pannozzo & Achilles, 2003 for a review), and even less research that identifies causal effects of the potential mechanisms. According to the existing studies, disruptive behavior and discipline issues occur more frequently in larger classes. It has been shown that a reduction in class size is associated with a decline in time spent on discipline (Betts & Shkolnik, 1999), and that in large classes children are more distracted from work and more often off task (Blatchford, Edmonds & Martin, 2003). Existing literature has also illustrated that disruptive behavior is associated with lower academic achievement. Disruptive behavior impedes learning for all students in the class (Erätuuli & Puurula, 1992) and reduces interaction and time spent on teaching (Blatchford et al., 2003). Further, potentially disruptive children (Kristoffersen et al., 2015), children from troubled families (Carrell & Hoekstra, 2010) and students with serious behavioral difficulties (Horoi & Ost, 2015) reduce the academic performance of their peers as well.

While there is only a small empirical literature on the mechanisms behind the class size effect, there is a vast literature concentrating on the effects of class size on academic achievement. Empirical research using experimental or quasi-experimental methods has shown that class size is related to school achievement (Angrist & Lavy, 1999; Krueger, 1999); cognitive and non-cognitive skills, completing education, and wages (Fredriksson et al., 2013); and to the likelihood of taking college entrance examination (Krueger & Whitmore, 2001). Insignificant or zero class size effects on academic achievement have been indicated by both descriptive (Hanushek, 1997; Kupiainen & Hienonen, 2016) and quasi-

experimental (Angrist et al., 2019; Falch et al., 2017; Hoxby, 2000; Leuven et al., 2008; Leuven & Løkken, 2020) studies.

We add to the existing literature by studying causal effects of class size on discipline problems. Our empirical strategy was based on exogenous variation in class size induced by maximum class size guidelines which stated that there should be maximum 25 students in the class. The guidelines did not bind the providers of education[2]. However, the Ministry of Education and Culture has allocated grants for reducing class sizes to the level defined by guidelines, creating monetary incentives to comply with the guidelines. In this paper, we demonstrate that the guidelines affect observed class sizes and that exogenous variation in class size due to maximum class size rules can be used to identify the effects of class size on discipline problems and academic achievement. We found that increasing class size increases discipline problems.

We used data collected in context of an intervention study "ProKoulu"[3] (Savolainen, Närhi & Savolainen, 2017). This project applied School Wide Positive Behavior and Support-Model (Sugai & Horner, 2002) in Finland, and tested its effectiveness on student behavior. During the project, detailed data on classroom behavior was collected from both teachers and students. We had access to seven waves of questionnaires. In total, our sample included 1,377 teacher reports and 15,552 student reports on classroom behavior.

While the emphasis of our study is on class size and discipline problems, we also add to the literature by providing evidence on class size and student achievement in primary schools. Measuring student achievement in primary school in Finland is difficult because Finnish schools do not use standardized tests and teacher-given grades are not collected to national registers. Even the ProKoulu project did not collect data on student achievement, but it provided a unique opportunity to collect the data on student achievement after the ProKoulu project was finished. With the help of ProKoulu research team, we managed to collect data on grades in mother tongue, first foreign language (English), and mathematics from some of the schools that had participated in the original intervention by doing on-site visits to municipality school administration offices. Our results on class size effects on academic achievement were largely insignificant, and in line with the literature supporting small or zero class size effects (e.g. Angrist et al., 2019).

The paper proceeds by discussing the institutional setting and the monetary subsidies granted for reducing class sizes in section 2.2. We continue to section 2.3 where the data and our identification strategy are introduced. In this section, we further discuss about the validity of the instrument used in analyses. Section 2.4 presents the results and is followed by discussion in section 2.5. Finally, we conclude and summarize in section 2.6.

[2] In Finland, providers of education are generally municipalities, but some schools are run by federations of municipalities, private non-profit organizations, or central government.
[3] "ProKoulu" is an acronym for Positiivisesti Ryhmässä Oppiva Koulu, which can be translated as positively in group learning-school.

2.2 Institutional Background

In Finland, compulsory comprehensive school starts in the year a child turns seven. Students spend nine years in comprehensive schools after which they can apply for upper secondary education. The minimum school leaving age is 18 years, which means that students gain an upper secondary qualification during compulsory education. Our data were from primary school, which consists of the first six years of comprehensive school.

Municipalities are responsible for providing education for their residents. The providers of education must follow national curriculum, but otherwise they have substantial autonomy to organize education for their students. The allocation of students to schools is determined in general by residential area. (Basic Education Act 1998/628).

Education in Finland is free of charge at all levels. Responsibility for funding education is divided between the central and the local governments. Funding from the state is channeled through a transfer system between the central and the local governments and through direct funding from the state to the local providers of education. The transfer system is based on lump sum grants that mainly depend on the number of school-aged children. Direct funding is also used for steering purposes and to achieve national goals in education policy, such as a reduction in class size[4].

Until 1999, maximum class size in Finnish comprehensive school was 25 students in first and second grade levels, and 32 students from third grade level onwards. As a part of reforming the Basic Education Act in 1999, the maximum class size regulation regarding the general education was abandoned. However, the providers of education still must follow quality guidelines in forming teaching groups. According to these quality guidelines, the maximum teaching group size varies between 20 and 25 students.

The guidelines concerning class size were not binding. However, the Ministry of Education and Culture has set reducing class size as one of the key priorities in its education policy and has granted more than 275 million euros for reducing class sizes to the level defined in the maximum class size guidelines since 2010. The data utilized in the present study covered years 2013–2017, allowing us to analyze changes in actual class sizes few years after the first class size reduction grants were awarded.

In 2013–2017, all providers of education were eligible to apply for the grants. Grant application process was straightforward and mainly required providers of education to report the number of students per grade level and the number of classes with more than 25 students as well as to present a plan on how they would use the funds to reduce class size. The class sizes were reduced for example by hiring new teachers or by dividing teaching groups. Reporting the use of awarded funds was also required after the grant period. Acceptance rates were

[4] More information about the financing of general education can be found from the Ministry of Education and Culture web site (Ministry of Education and Culture, 2020).

high, in most years well over 90%. The grants were distributed rather evenly across Finland with the largest providers of education receiving the largest grants. These largest education providers also had the highest need for funding: the cities of Helsinki, Turku, Oulu, Pori, Kuopio, and Tampere had the most teaching groups with more than 25 students (See Table A1 in the Appendix). In the Appendix Table A2, we report number of applicants, acceptance rates and granted class size reduction subsidies by year. Figure 1 displays grants awarded in 2013 on a map, rounded to thousands of euros. The circle in the map shows the location of schools participating in the ProKoulu study. The information about the applied and received grants were requested and obtained from the Ministry of Education and Culture.

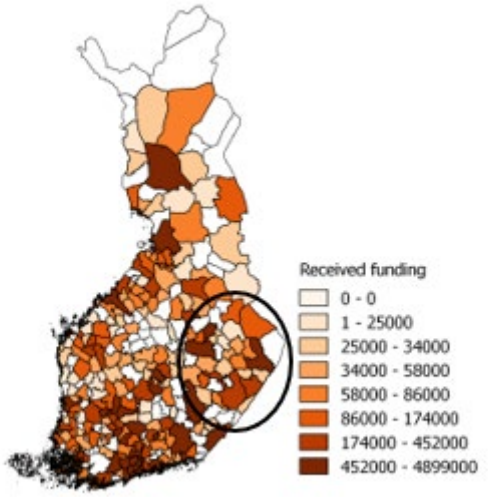

FIGURE 1 Distribution of class-size reduction funding by municipality in 2013. The information about received funding was requested and obtained from the Ministry of Education and Culture.

2.3 Data and Methods

2.3.1 ProKoulu-project

Data used in this study were collected as part of ProKoulu project, an intervention study conducted by the University of Jyväskylä, Niilo Mäki Institute and the University of Eastern Finland. The project was offered to all primary schools in Eastern Finland. Initially, 70 primary schools registered, and 60 schools eventually participated in the study. The data covered approximately 20% of all primary schools in Eastern Finland, and 3% of all primary schools in Finland. In the beginning of the study, the schools were randomized into a treatment group and a control group. In the control group the intervention was delayed by one year. The main aim of the intervention study was to investigate the effectiveness of the

ProKoulu model in supporting positive student behavior. The effectiveness of the ProKoulu model was not analyzed in this study, but we controlled for the intervention effects.

Data consisted of responses to questionnaires by students and teachers. These questionnaires were filled twice during each school year. By the end of the seventh round, 1,531 teachers and 10,092 students had responded to the questionnaires. Of these, 1,288 teachers and 10,088 students responded to the questions concerning discipline problems at least in one of the seven questionnaire rounds (from autumn 2013 to autumn 2016). In the present study, we pooled the data from all seven rounds yielding a total of 4,434 teacher responses and 38,992 student responses. We restricted the data to mainstream education classes and to classes where all students were in the same grade level. As many schools in the data were small, this was the main restriction limiting our sample size.[5] We excluded teachers that taught multiple classes as they reported discipline problems on average in the classes that they taught. We further excluded classes with more than one teacher. After these restrictions, our data included 404 teachers and 4,881 students that responded at least one of the questionnaire rounds. The pooled data included 1,377 responses from the teachers and 15,552 responses from the students from 47 schools.

2.3.2 Measures and Variables

We conceptualized discipline problems following Levin and Nolan (2010) comprising of four components: 1) Students' behavior is not geared towards learning, 2) Students compromise the rights of others to learn, 3) Learning situations are psychologically or physically unsafe, and 4) Students do not take appropriate care of the physical classroom environment. Each of these components of discipline problems has been studied separately (Närhi, Kiiski, Peitso & Savolainen 2015), but for this study, we aggregated the components to a single scale to indicate overall level of observed discipline problems. In the analyses, we used the sum of the items; higher values referring to more discipline problems. We then standardized the measures to have zero mean and unit variance.

Both teachers and students were asked questions related to discipline problems twice a year. Teachers evaluated discipline problems in grades one to six, and students in two to six, as first-grades were deemed not to have necessary reading and writing skills for participating in the questionnaire. The teachers' questionnaire included 14 items (e.g., 'Students concentrate well on teaching') and students' questionnaire 17 items (e.g., 'Classroom behavioral climate is loud and disorganized'). Complete questionnaires are presented in the Appendix (Table A3). Teachers and students evaluated how well the statement described the action during classes. Teachers' response scale ranged from 1 = very poorly to 6 = very well, and students' scale from 1 = never to 4 = in every class.

[5] If the birth cohort was small, and therefore the enrollment count was small in a certain year, it was possible that two or more grade levels were combined into same teaching group.

Crohnbach's alpha reliabilities for the discipline scales were high: 0.89 for teacher responses and 0.88 for student responses.

The information on teacher characteristics and classes that they taught was obtained from the teachers. We categorized teachers' work experience as less than one year, 1–5 years, 6–10 years, 11–15 years, and more than 15 years. We further used teacher reports on number of students in the class who required intensified or special support.[6] Finally, enrollment at grade level and the size of the school were based on official reports given by schools. Class size information was collected from the schools and further inquired from teachers once an academic year. The fraction of boys in each class was based on official school data, complemented by the teacher reports.

We linked all other information to the teacher and the student responses. As one teacher was associated with one class at a time, the observations in the teacher responses file were effectively class level observations. In contrast, the student responses file contained measures at the student level with additional class level information linked to all students in the class.

Table 1 presents the descriptive statistics based on both the teacher responses and the student responses. Mean of teacher evaluated discipline problems before standardization was 34.3 out of highest possible 84. Average class size in the teacher data was 18.6 students, and mean enrollment at grade level was 35.6. The smallest school had 56 students and largest had 533 students. Responses from different grade levels were evenly distributed, and 60% of teachers had more than 15 years of work experience. On average, 13% of students in the class had special or intensified support decision, and on average, 51% of students in the class were boys.

In student response data, the mean of the discipline problems was 31.8 out of the possible maximum 68 before standardizing the measure. Class size, grade level enrollment, school size, and fraction of boys were linked to the student data from school level responses and reflected student-weighted averages of these variables. The average class size in student data was naturally higher than in teacher data, as there were more respondents in larger classes, and the data of the first graders, who typically studied in smaller classes, were not collected from the students. The responses were again evenly distributed throughout the grade levels.

[6] These students were likely to require extra attention from teacher, when included in mainstream education classes. The need for intensified support was based on pedagogical evaluation, and it was made if general support available for all students was inadequate. Intensified support included a learning plan and continuous, individual support. Respectively, students were entitled to special support if intensified support was deemed inadequate. Special support included fulltime remedial education and a personal plan for teaching arrangements (Basic education act, amendments 642/2010).

TABLE 1 Descriptive statistics for pooled ProKoulu data

Teacher responses (*N*=1,377)	*M*	*SD*
Discipline problems (S)	0.00	1.00
Class size	18.60	4.00
Enrollment count	35.55	18.94
School size	212.5	105.71
1st grade	0.21	0.40
2nd grade	0.16	0.37
3rd grade	0.16	0.37
4th grade	0.16	0.36
5th grade	0.16	0.37
6th grade	0.16	0.36
Work experience less than 1 year	0.01	0.10
Work experience 1–5 years	0.10	0.30
Work experience 6–10 years	0.14	0.347
Work experience 11–15 years	0.15	0.36
Work experience more than 15 years	0.60	0.49
Treatment group	0.55	0.50
Fraction special or intensified	0.13	0.11
Fraction boys	0.51	0.12
Student responses (*N*=15,552)		
Discipline problems (S)	0.00	1.00
Class size	19.7	4.1
Enrollment at grade level	35.8	17.7
School size	212	100
2nd grade	0.18	0.39
3rd grade	0.21	0.41
4th grade	0.21	0.41
5th grade	0.20	0.40
6th grade	0.19	0.39
Fraction boys	0.51	0.11

Notes: Pooled ProKoulu data; *M* = mean, *SD* = standard deviation, S = standardized

The class size mode was 18 students in our data. As the class size guidelines were not binding, there were some classes with more than 25 students (5.3%), and we also had small classes with less than 15 students (15.1%). In Appendix Figure A1 shows the distribution of the class size pooled over grade levels and questionnaire rounds in our teacher-response data. We further compared class sizes in our sample in fall 2013 to the class sizes in the entire Finland 2013 (Table

A4). Compared to the entire Finland, there were less students in classes in our data, but the differences were relatively small.

2.3.3 Identification Strategy

The challenge in studying the effects of class size is the endogeneity of class size. Endogeneity may arise from school decisions of decreasing class size in hard-to-teach classes or placing hard-to-teach students to smaller classes. To overcome endogeneity of class size, we used instrumental variable method (IV), which relies on finding an instrument that has an effect on class size, but plausibly no effect on discipline problems. Following Angrist and Lavy (1999), we constructed an instrument based on maximum class size guidelines. Assuming that the schools strictly followed the guidelines setting maximum class size to 25, expected class size f_{sc} in school s and class c depended on number of children at the grade level e_s according to the relationship

$$f_{sc} = \frac{e_s}{[int\left(\frac{e_s-1}{25}\right)+1]}. \quad (1)$$

According to this function, expected class size depends directly on enrollment, but drops discretely whenever the enrollment in grade level exceeds 25 or a multiplier of 25. For example, if 25 students enrolls, the expected class size is 25 students. When enrollment increases to 26, students are divided into two classes with 13 students in each class.

As shown in Figure 2, in our data, expected class sizes and the actual class sizes were strongly related. In the Figure 2, the solid line describes expected class size if the maximum class size guidelines were strictly followed. The dots describe the observed average class sizes by grade level enrolment in school. Deviations from expected class size occurred when schools created classes that were not equal in size and when schools failed to comply with the guidelines.

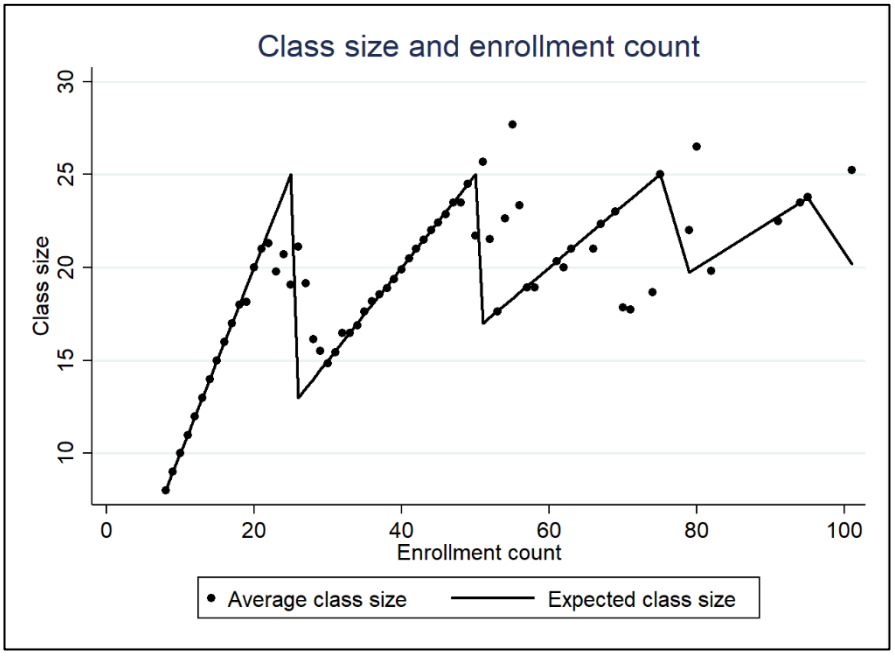

FIGURE 2 Average class size and expected class size by the enrollment count in pooled ProKoulu data. The solid line describes expected class size if the maximum class size guidelines were strictly followed. The dots describe the observed average class sizes by grade level enrollment in school.

This spiky relationship between enrollment at the grade level and class size provided an opportunity to estimate the effects of class size using instrumental variables framework. Specifically, we used expected class size as an instrument for actual class size while controlling for a smooth function of enrollment at the grade level.[7]

The equation of interest (Equation 2) relates discipline problems d reported by individual i to class size cs while controlling other variables X that include a smooth function of enrollment (first, second, and third polynomials), size of the school, dummies for grade level, teachers' work experience and intervention status, percentages of students with special or intensified support decision and percentage of boys in class.

$$d_{isc} = \delta cs_{sc} + \beta X_{sc} + \varepsilon_{sc}. \quad (2)$$

The first stage of this approach (Equation 3) uses expected class size given in Equation 1 as an instrument for observed class size controlling for the same set of X variables

$$cs_{sc} = \gamma f_{sc} + \beta X_{sc} + \mu_{sc}. \quad (3)$$

[7] An alternative approach used by e.g. Fredriksson, Öckert, and Oosterbeek (2013) would have been to focus on observations close to class-size cut-offs and to produce separate estimates at each cut-off. Due to relatively small sample size, we did not have sufficient power to further limit the sample or split the data and therefore we followed the original idea in Angrist and Lavy (1999).

2.3.4 IV Assumptions and the Validity of the Instrument

Validity of a discontinuity-based approach requires that observations are not bunched around cutoffs and that characteristics of the students or teachers are balanced across observations below and above the cutoffs. Discontinuities could arise, for example, if schools assigned experienced teachers to large classes, or if boys or students with learning difficulties were placed into small classes. In our balancing tests, we graphically analyzed whether control variables were continuous at the thresholds. In Figure 3, we show the relation between enrollment and the average teachers' work experience. Visually, we detected very small jumps at the thresholds, especially for the first two thresholds. In the Appendix, we show corresponding figures for percentages of boys and percentages of students with special or intensified support decision (Figures A2 and A3). These figures show the same pattern: there were no jumps or very small jumps at the first two cutoffs.

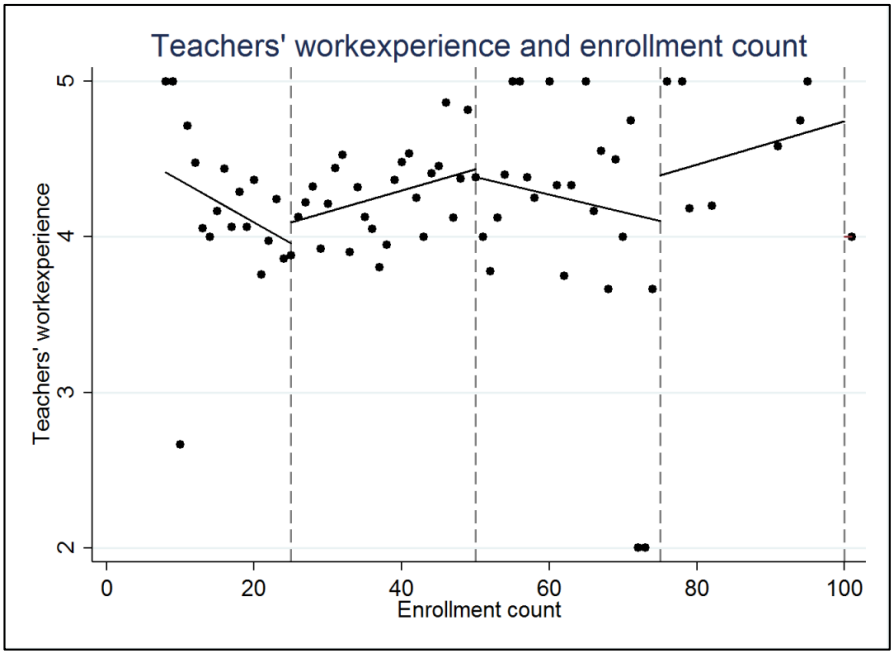

FIGURE 3 Mean of teachers' work experience by the enrollment count in pooled Pro-Koulu data, teacher reports.

As our statistical approach exploited the variation in class size determined by maximum class size guidelines, it was important that there was no bunching at the cutoffs. Bunching would indicate, for example, that municipalities or parents could successfully manipulate the cutoff. Municipalities could have tried to avoid classes with more than 25 students, as having more students would create a need of having more teachers, or parents could have tried to avoid big classes. To see if there was bunching at the cutoffs, we first studied the distribution of the enrollment count in our data. Figure 4 includes all the enrollment counts at the beginning of the school years, and gives no support for such manipulation, as the enrollment counts are not stacked on multiples of 25 students. The McCrary (2008)

density tests confirmed that there were no significant shifts in the discontinuity at the threshold.[8] As the maximum class size guideline was not binding, bunching was less likely than in settings with binding class size rules (Angrist & Lavy, 1999; Fredriksson, Öckert & Oosterbeek, 2016; Gary-Bobo & Mahjoub, 2013; Hoxby, 2000; Leuven & Løkken, 2020; Leuven et al., 2008; Urquiola & Verhoogen, 2009).

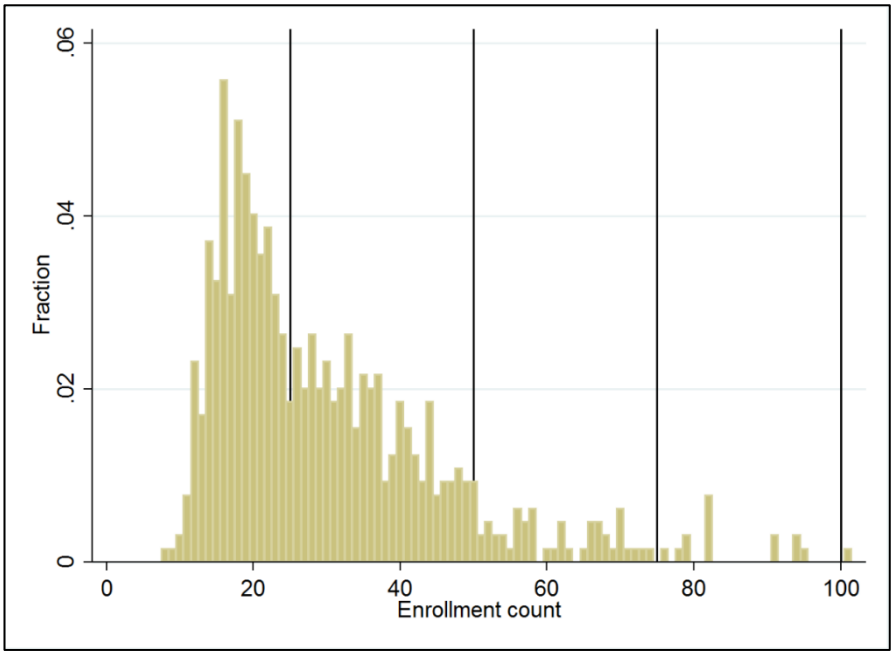

FIGURE 4 Distribution of the enrollment counts at the beginning of the school years in ProKoulu data.

2.4 Results

2.4.1 Descriptive Results

In Table 2, we report correlations between discipline problems, class size, and control variables separately based on the teacher and the student reports. According to both the teacher reports and the student reports, discipline problems positively correlated with class size, enrollment count, grade level, percentage of students with special or intensified support decisions and percentage of boys, whereas the correlations between discipline problems and teachers' work experience were negative. The correlations between discipline problems and class size, work experience, and percentage of students with special or intensified decision were stronger in teachers' data. Correlation between teacher reported discipline

[8] For the density test, we pooled the thresholds (25, 50, 75, and 100). Bin size of one student and bandwidth of five students were used. The log difference in height was - 0.231 with standard error of 0.305.

problems and discipline problems reported by students in the same class was 0.50 ($p < 0.01$; $N = 1{,}377$).

TABLE 2 Correlations of discipline problems (S), class size and control variables in pooled ProKoulu data.

	Discipline problems (S), Teachers ($N = 1{,}377$)	Discipline problems (S), Students ($N = 15{,}552$)
Class size	0.14***	0.07***
Enrollment count	0.06**	–0.03***
Grade level	0.05*	0.02***
Teachers' work experience	–0.19***	–0.11***
Special or intensified support decisions	0.08***	0.05***
Boys	0.11***	0.11***

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$, S = Standardized

Figure 5 displays the relationship between enrollment count and teacher reported discipline problems. While this relationship was clearly less tight than relationship between expected and actual class size displayed in Figure 2, it indicated that discipline problems decrease as enrollment exceeds 25 and expected class size decreases. Changes in the discipline problems at other cutoffs were small. This was expected by the design, as the changes in the recommended class sizes were smaller for cutoffs at higher enrolments.

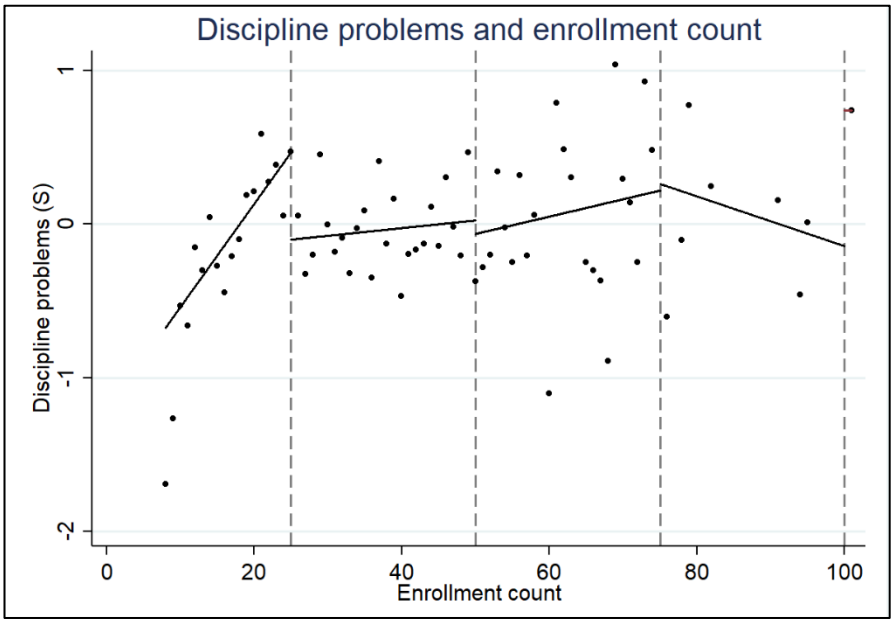

FIGURE 5 Discipline problems (S) by enrollment in pooled ProKoulu data, teacher reports. The dots describe the observed average discipline problems by grade

level enrolment in school. The solid line describes a linear fit between enrollment count and discipline problems using ordinary least squares (OLS). S = standardized.

2.4.2 Class size and Discipline Problems

In Table 3, we report the main empirical results. The top row of the Table 3 presents results from a simple ordinary least squares regression (OLS) explaining discipline problems with actual class size. Estimates in the Columns 1-2 were based on teacher responses and estimates in the Columns 3-4 on the student responses. In the Columns 1 and 3, we controlled only for the questionnaire round and enrollment at the grade level; and in the Columns 2 and 4 the characteristics of the teachers and the students were also controlled for. According to teacher reports, the OLS estimates indicated that adding one extra student in the classroom associates with a 0.037 standard deviation increase in discipline problems. Adding controls in Column 2 slightly increased the estimate. The OLS estimates based on student reports were somewhat smaller, but still statistically significantly different from zero.

The OLS estimates were likely to be biased because of endogeneity of class size. Therefore, we continued with IV method. The first stage of the IV estimation showed that higher expected class size was related to higher class sizes. The relationship was rather strong; F-statistics exceeded 10 in all specifications. According to teacher responses, class size had a positive effect on discipline problems (Columns 1-2). Adding one student into class increased discipline problems by 0.076 standard deviation units. Adding control variables for teacher and student characteristics increased the coefficient slightly, to 0.085 standard deviation units. The student responses (Columns 3-4) showed smaller effects compared to the teacher responses. Adding one student to the class increased the discipline problems by 0.029 standard deviation units when only enrollment and questionnaire rounds were controlled for, and by 0.041 standard deviation units when other controls were added. The estimate was only weakly statistically significant in the latter case.

TABLE 3 OLS and IV-estimates: Class size and discipline problems in Pooled Pro-Koulu data

	Teacher reports		Student reports	
Model	(1) Discipline problems (S)	(2) Discipline problems (S)	(3) Discipline problems (S)	(4) Discipline problems (S)
OLS				
Class size	0.037***	0.044***	0.028***	0.032***
	(0.010)	(0.009)	(0.007)	(0.007)
IV				
Class size	0.076***	0.085***	0.029	0.041*
	(0.027)	(0.026)	(0.027)	(0.023)
FIRST STAGE	Class size	Class size	Class size	Class size
Expected class size	0.530***	0.521***	0.476***	0.497***
	(0.077)	(0.078)	(0.087)	(0.082)
F-statistics	47.29	44.58	29.70	36.72
CONTROLS				
Questionnaire round and functions of enrollment	x	x	x	x
Other controls		x		x
N	1,377	1,377	15,552	15,552
Clusters	47	47	46	46

Notes: S = standardized, N= number of observations. Functions of enrollment included enrollment, enrollment^2 and enrollment^3. Other controls included grade level, teacher's work experience, dummy for treatment group, percentages of students with special or intensified support decision and boys, and size of the school. Standard errors adjusted for clustering by school level are in parenthesis. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

In the Appendix, we show that our results were robust for different choice of the enrollment polynomials, and for using piecewise linear trend (Angrist & Lavy, 1999) instead of functions of enrollment (Table A5). Adding students' socioeconomic status (measured as Family Affluence Scale, FAS) to the control variables increased the point estimates slightly (Table A6). Finally, we followed Fredriksson et al. (2013), and pooled the data from different enrollment thresholds, and used being above the threshold as an instrument for the class size (Tables A7 and A8). We tried bandwidths of ±12.5 and ±10 of the threshold. For smaller window widths, we did not have enough data. The results confirmed that higher class size translates to higher discipline problems for teachers, with point estimates larger than in our main specification (ranging between 0.122 and 0.209 standard deviation units). Using student evaluations, the point estimates were again larger than the estimates presented in Table 3, but in general insignificant.

2.4.3 Class size and Academic Achievement

In Finland, there are no national registers on academic achievement on primary schools. Information on academic achievement of the students was neither collected during the ProKoulu-study, nor were the students assessed with achievement tests. To analyze the effect of class size on student achievement, we collected the data on grades in mother tongue, first foreign language (English), and mathematics after the ProKoulu-study was finished, from the schools that had participated in the intervention. As collecting grade data required on-site visits to municipality school administration offices and in some cases permissions to access their computer terminals, we concentrated on the largest municipalities. We requested data from 54 schools, and eventually managed to collect data from 37 schools. For the students in these schools, we got access to grade histories of each student from the years the experiment was ongoing.

The schools had used various grading scales. For example, some grades were given on a numerical scale from 4 (fail) to 10 (excellent) while some schools used coarser scales and verbal statements ranging from "needs more practice" to "masters the subject". We converted all scales to standardized units, so that in each grading scale achievement had mean zero and variance one.

We linked the grade data to the existing ProKoulu data. As before, we pooled the data from the seven questionnaire rounds and restricted the data analysis to mainstream education classes with one teacher and no combined grade levels. After these restrictions, the pooled data included approximately 4,000 student level observations and 250 class level averages that was linked to teacher responses from 17 to 19 schools. The studies of the first foreign language typically started at third grade at the time of data collection, and therefore there were fewer observations from that subject.

The results on the effect of class size on academic achievement are presented in Table 4. Using the teacher data, we found no support for that class size would affect the academic achievement. For the students, the results implied negative association between the class size and grades of first foreign language and mathematics using OLS estimation. However, the IV estimates were statistically insignificant.

TABLE 4 OLS and IV-estimates: Class size and academic achievement in pooled ProKoulu data

	Teacher data				Student data			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model	Mother tongue (S)	First foreign language (S)	Mathematics (S)	Mean (Mother tongue, English, Mathematics; S)	Mother tongue (S)	First foreign language (S)	Mathematics (S)	Mean (Mother tongue, English, Mathematics; S)
OLS								
Class size	0.007	–0.007	–0.019	–0.010	0.001	–0.017**	–0.026***	–0.016**
	(0.012)	(0.011)	(0.012)	(0.010)	(0.008)	(0.008)	(0.009)	(0.007)
IV								
Class size	0.010	0.016	–0.003	–0.004	0.011	0.011	–0.017	–0.003
	(0.016)	(0.020)	(0.022)	(0.015)	(0.014)	(0.017)	(0.016)	(0.016)
FIRST STAGE	Class size	Class size	Class size	Class size	Class size	Class size	Class size	Class size
Expected class size	0.591***	0.757***	0.593***	0.736***	0.644***	0.773***	0.637***	0.752***
	(0.126)	(0.073)	(0.126)	(0.081)	(0.115)	(0.072)	(0.121)	(0.078)
F-statistics	21.85	107.64	22.28	83.39	31.66	116.58	27.98	92.98
CONTROLS								
Functions of enrollment	x	x	x	x	x	x	x	x
Other controls	x	x	x	x	x	x	x	x
N	286	254	291	241	4,174	3,780	4,205	3,514
Clusters	18	17	18	17	19	18	19	18

Notes: S = standardized, N = number of observations. Functions of enrollment included enrollment, enrollment^2 and enrollment^3. Other controls included grade level, teacher's work experience, dummy for treatment group, percentages of students with special or intensified support decision and boys, and size of the school. Standard errors adjusted for clustering by school level are in parenthesis. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

2.5 Discussion

Based on both teacher and student assessments, increase in class size increased discipline problems. According to the teacher evaluations, adding one student to the class increased discipline problems by 0.085 standard deviation units, whereas in students' case, the corresponding effect was 0.041 standard deviation units. While statistically significant, the magnitudes of these effects were difficult to evaluate. As a point of comparison, Korpershoek, Harms, de Boer, van Kuijk, and Doolaard (2016) reported in their meta-analysis that the average effect size (Hedge's g) of classroom management strategies and classroom management programs on behavior outcomes was 0.24 standard deviation units. According to our teacher evaluation estimates, similar improvements in student behavior could be achieved by reducing class size by 2.8 students (or by 5.9 students using estimates based on student evaluations).

Our results indicated differences in how teachers and students experienced changes in class size in terms of discipline problems: teachers experienced stronger increase in discipline problems than students. Earlier literature has reported similar findings. For example, Korpershoek et al. (2016) found that student reports show less improvement in behavioral outcomes than reports filled in by teachers or observers. We speculated whether the difference resulted from the propositions being difficult to understand for the youngest students. However, excluding the youngest children (2nd graders) from the estimations did not change the results. Another explanation could be that some students may have studied in the same group through primary school and had no point of comparison while assessing the discipline problems. Conversely, the teachers were likely to have multiple points of comparison, as 60% of teachers in our data had at least 15 years of teaching experience (see Table 1). Finally, larger class sizes are often related to bigger workloads for teachers, and therefore teachers may find bigger classes taxing. Thus, compared to students, teachers may have had stronger incentives to report higher levels of discipline problems in bigger classes. For example, Trade Union of Education in Finland (OAJ), which has 95% of Finnish teachers as its members, has reported higher discipline problems in large classes.

Our data were unique in a sense that it linked class size information into classroom behavioral outcomes. It is further rare to have information about academic achievement in primary schools in Finland. However, it is possible, that the level of discipline problems was lower in our data than in classes generally. This is because our data were drawn from an intervention study, in which positive student behavior was supported. More research is needed to confirm the magnitude of the class size effects on discipline problems.

We found limited evidence on the effects of class size on academic achievement. Partially, this may be due to lack of standardized tests in Finnish schools. Students took exams and received grades in most subjects in the end of each term, but grading was done by teachers, and grading standards and grading

scales varied across grade levels and across the schools. It is possible that the teachers graded on a curve, and thus adjusted the grades according to the level of the class. Teachers may also have considered grades as a relative performance measure, and as such, the grades could have been influenced by discipline problems. Hence, we had no way of ensuring the comparability of the grades across schools or classes.

Fredriksson et al. (2013) studied the effects of being placed into smaller classes in the last three years of primary school on cognitive ability and academic achievement. They found that adding one student to the class decreases cognitive ability at the age of 13 by 0.033 (0.015) standard deviation units (standard errors in parenthesis) and academic achievement at the age of 16 by 0.023 (0.01) standard deviation units. Angrist and Lavy (1999) studied the effect of class size on reading comprehension and math scores. Compared to Fredriksson et al. (2013) they reported estimates larger in magnitude for fourth graders, and the estimates were largely insignificant. As for fifth graders, the results were statistically significant. In more recent study, Angrist et al. (2019) reported zero class size effects for fifth graders in Israeli framework. Their estimates were very similar to ours both in magnitude and sign.

Our setting differed from the Swedish and Israeli settings in terms of the age of students and class size. We studied students in primary school (grades from 1 to 6, students aged 7-12), and the drop at the cutoff in our case was from 25 to 13 students. In Swedish and Israeli settings, the students were older and the drops at the cutoffs were larger. Our results are in line with the study done by Angrist et al. (2019), suggesting small class size effects. We found statistically significant results in OLS estimations when we studied the student responses, suggesting a negative association between class size and academic achievement. The IV results were statistically insignificant and positive in the case of mother tongue and English. It is possible that our IV results were not statistically significant because of the lack of statistical power, as the data collected were moderately sized.

2.6 Conclusions

We studied the effect of class size on discipline problems. The participants were drawn from an intervention study, and the data consisted of questionnaire responses from the students and the teachers. The results showed that larger class size increases discipline problems. According to the teacher evaluations, adding one student to a class increased discipline problems by 0.085 standard deviation units, whereas according the student responses, the corresponding effect was 0.041 standard deviation units.

We also analyzed the effect of class size on academic achievement. We found signs of a negative associations between class size and academic achievement, but no evidence on causal effects of class size on grades in mother tongue, mathematics or first foreign language.

Decreasing class size is expensive. Even if the discipline problems increase with the class size, it is not clear that the benefits exceed the costs. In addition to the class size reductions, educational studies have illustrated that appropriate classroom management strategies (see e.g. Korpershoek et al., 2016) can improve behavioral outcomes. For example, revisiting the classroom behavioral expectations, teaching the expected behaviors, and reinforcing appropriate behavior (e.g. Epstein, Atkins, Cullinan, Kutash & Weaver, 2008) may yield similar benefits as class size reductions.

References

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' Rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, *114*(2), 533–575. https://doi.org/10.1162/003355399556061

Angrist, J. D., Lavy, V., Leder-Luis, J., & Shany, A. (2019). Maimonides' Rule redux. *American Economic Review: Insights*, *1*(3), 309–324. https://doi.org/10.1257/aeri.20180120

Betts, J. R., & Shkolnik, J. L. (1999). The behavioral effects of variations in class size: The case of math teachers. *Educational Evaluation and Policy Analysis*, *21*(2), 193–213. https://doi.org/10.3102/01623737021002193

Blatchford, P., Bassett, P., Goldstein, H., & Martin, C. (2003). Are class size differences related to pupils' educational progress and classroom processes? Findings from the Institute of Education Class Size Study of Children Aged 5–7 Years. *British Educational Research Journal*, *29*(5), 709–730. https://doi.org/10.1080/0141192032000133668

Blatchford, P., Edmonds, S., & Martin, C. (2003). Class Size, Pupil Attentiveness and Peer Relations. *British Journal of Educational Psychology* 73(1), 15–36. https://doi.org/10.1348/000709903762869897

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, *2*(1), 211–228. https://doi.org/10.1257/app.2.1.211

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *Quarterly Journal of Economics 126*(4), 1593–1660. https://doi.org/10.1093/qje/qjr041

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, K. (2008). *Reducing behavior problems in the elementary school classroom*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education.

Erätuuli, M., & Puurula, A. (1992). *Miksi häiritset minua: Opettajan näkökulma työrauhahäiriöihin yläasteella [Why do you disturb me? Disruptive behavior experienced by lower secondary school teachers]*. Yliopistopaino.

Falch, T., Sandsør, A. M. J., & Strøm, B. (2017). Do smaller classes always improve students' long run outcomes? *Oxford Bulletin of Economics and Statistics 79*(5), 654-688. https://doi.org/10.1111/obes.12161

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The 'Why's' of Class Size: Student Behavior in Small Classes. *Review of Educational Research* 73(3), 321–368. https://doi.org/10.3102/00346543073003321

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *The Quarterly Journal of Economics*, *128*(1), 249–285. https://doi.org/10.1093/qje/qjs048

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2016). Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule. *Journal of Human Resources 51*(4), 832–868. https://doi.org/10.3368/jhr.51.4.1114-6779R1

Gary-Bobo, R. J., & Mahjoub, M. (2013). Estimation of Class-Size Effects, Using "Maimonides' Rule" and Other Instruments: the Case of French Junior High Schools. *Annals of Economics and Statistics 111/112*, 193–225. https://doi.org/10.2307/23646331

Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis 19*(2), 141–164. https://doi.org/10.2307/1164207

Horoi, I., & Ost, B. (2015). Disruptive Peers and the Estimation of Teacher alue Added. *Economics of Education Review 49,* 180–192. https://doi.org/10.1016/j.econedurev.2015.10.002

Hoxby, C. M. (2000). The effects of class size on student achievement : New evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239–1285. https://www.jstor.org/stable/2586924

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kristoffersen, J. H. G., Krægpøth, M. V., Nielsen, H. S., & Simonsen, M. (2015). Disruptive school peers and student outcomes. *Economics of Education Review*, *45*, 1–13. https://doi.org/10.1016/j.econedurev.2015.01.004

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, *114*(2), 497–532. https://doi.org/10.1162/003355399556052

Krueger, A. B., & Whitmore, D. M. (2000). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR. *NBER Working Paper Series*, *7656*.

Kupiainen, S., & Hienonen, N. (2016). *Luokkakoko [Class size]*. Suomen kasvatustieteellinen seura.

Lazear, E. P. (2001). Educational production. *Quarterly Journal of Economics*, *116*(3), 777–803. https://doi.org/10.1162/00335530152466232

Leuven, E., & Løkken, S. A. (2020). Long-term impacts of class size in compulsory school. *Journal of Human Resources*, *55*(1), 309–348. https://doi.org/10.3368/jhr.55.2.0217.8574r2

Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *The Scandinavian Journal of Economics*, *110*(4), 663–693. https://doi.org/10.1111/j.1467-9442.2008.00556.x

Levin, J., & Nolan, J. (2010). *Principles of Classroom Management: A Professional Decision-Making Model* (6th edition). Boston: Pearson.

Ministry of Education and Culture. (29.6.2020). Financing of General Education. https://minedu.fi/en/financing-of-general-education.

Närhi, V., Kiiski, T., Peitso, S., & Savolainen, H. (2015). Reducing Disruptive Behaviours and Improving Learning Climates with Class-Wide Positive Behaviour Support in Middle Schools. *European Journal of Special Needs Education 30*(2), 274–285. https://doi.org/10.1080/08856257.2014.986913

Savolainen, H., Närhi, V., Malinen, O-P., & Savolainen, P. (2017). ProKoulu model in Finland. *Remediaal : tijdschrift voor leer- en gedragsproblemen in het vo/bve, 17*(4-5), 16-18. https://tuhat.helsinki.fi/ws/portalfiles/portal/159645334/Savolainen_N_rhi_Malinen_Savolainen_2017_ProKoulu_model_in_Finland.pdf

Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy*, *24*(1-2), 23–50. https://doi.org/10.1300/J019v24n01_03

Urquiola, M., & Verhoogen, E. (2009). Class-Size Caps, Sorting, and the Regression-Discontinuity Design. *American Economic Review 99*(1), 179–215. https://doi.org/10.1257/aer.99.1.179

Appendix

TABLE A1 Ten municipalities with highest need of funding in terms of class sizes, and ten municipalities with highest grants in 2013

Highest grant		Highest need	
Municipality	Grant	Municipality	Amount of teaching groups with >25 students
Helsinki	4,898,700	Helsinki	523
Tampere	3,753,000	Turku	257
Vantaa	3,621,000	Oulu	228
Espoo	3,513,000	Pori	204
Turku	2,550,600	Kuopio	200
Oulu	2,404,900	Tampere	173
Kuopio	1,721,300	Joensuu	158
Lahti	1,547,700	Jyväskylä	158
Nurmijärvi	1,216,800	Espoo	145
Pori	1,119,500	Kerava	144

Notes: The grant information was requested and obtained from the Ministry of Education and Culture. The amount of teaching groups with more than 25 students were calculated from the collection of teacher data in 2013.

TABLE A2 Monetary subsidies on reducing class size: year, number of applicants, acceptance rate, and granted subsidies in millions of euros

Year	Number of applicants	Acceptance rate %	Granted subsidies m€
2010	236	100	31
2011	299	72	30
2012	257	96	50
2013	249	96	60
2014	417	97	75*
2015	236	94	30

Notes: The information was requested and obtained from the Ministry of Education and Culture. * In 2014, there were two application rounds. In the first round, 230 applied and received funding. In total, 60 million euros was granted. In the second round, among the 187 applied, 173 received funding, yielding 15 million euros in total.

TABLE A3 Discipline problem-questionnaires

Teachers	Students
Classroom behavioral climate is good during classes (R)	Teacher has to wait for a long time until the students quiet down
Students concentrate well on teaching (R)	Classroom behavioral climate is loud and disorganized
Students perform exercises calmly during classes (R)	In the beginning of the class, it takes more than five minutes before anything happens
It is too loud during classes	Classroom behavioral climate is good (R)
During the classes, there is movement unrelated to teaching	Students concentrate well on teaching (R)
Students discuss topics other than the theme of the class	Students perform exercises calmly during classes (R)
Students disrupt each other's studying	It is too loud
Students interrupt teacher's speech	There is movement unrelated to teaching
Students call each other names maliciously	Students discuss topics other than the theme of the class
Students mock each other for incorrect answers	Students disrupt each other's studying
Students wade or threat to wade into each other	Students interrupt teacher's speech
Students take good care of items in the classroom (R)	Students call each other names maliciously
Students intentionally break items in the classroom	Students mock each other for incorrect answers
Students leave the classroom in neat condition after the classes (R)	Students wade or threat to wade into each other
	Students take good care of items in the classroom (R)
	Students intentionally break items in the classroom
	Students leave the classroom in neat condition after the classes (R)

Note: R = scale reversed for the analyses

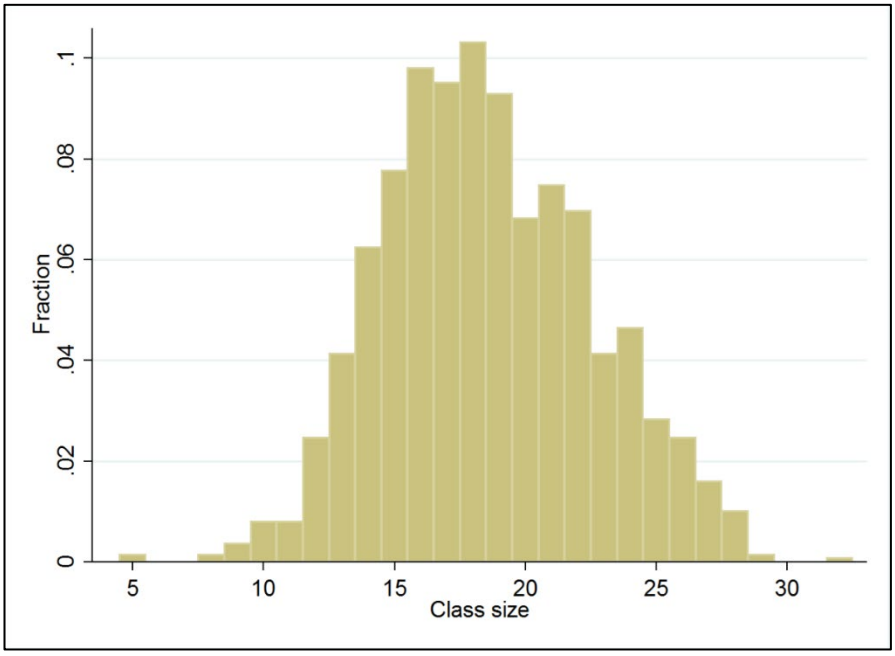

FIGURE A1 Distribution of class size in pooled ProKoulu data, teacher responses.

TABLE A4 Class sizes in the ProKoulu data and in entire Finland in 2013

	ProKoulu data 2013					Whole Finland 2013					
Grade	*N*	*M*	*SD*	*Min*	*Max*	*N*	*M*	*SD*	*Min*	*Max*	*p-value*
1st grade	87	17.45	3.05	11	27	2,223	18.55	4.05	2	33	0.002
2nd grade	81	19.01	4.59	9	31	2,167	18.84	4.05	2	35	0.743
3rd grade	83	18.60	5.02	4	34	2,012	19.96	4.29	1	34	0.017
4th grade	79	19.57	5.46	4	36	1,941	19.84	4.19	3	32	0.665
5th grade	77	18.86	4.67	4	28	1,904	20.39	4.33	4	36	0.006
6th grade	80	19.90	4.35	7	30	1,888	20.66	4.37	4	35	0.130
Combined grades	110	13.25	6.40	2	30	2,359	15.82	5.22	1	36	0.001
1st-6th grades	487	18.87	4.61	4	36	12137	19.66	4.28	1	36	0.000

Notes: N =Number of classes, M = mean, *SD* = Standard deviation, *Min*: smallest class size, *Max* = largest class size; Class sizes for the whole Finland were obtained from the collection of teacher data in 2013.

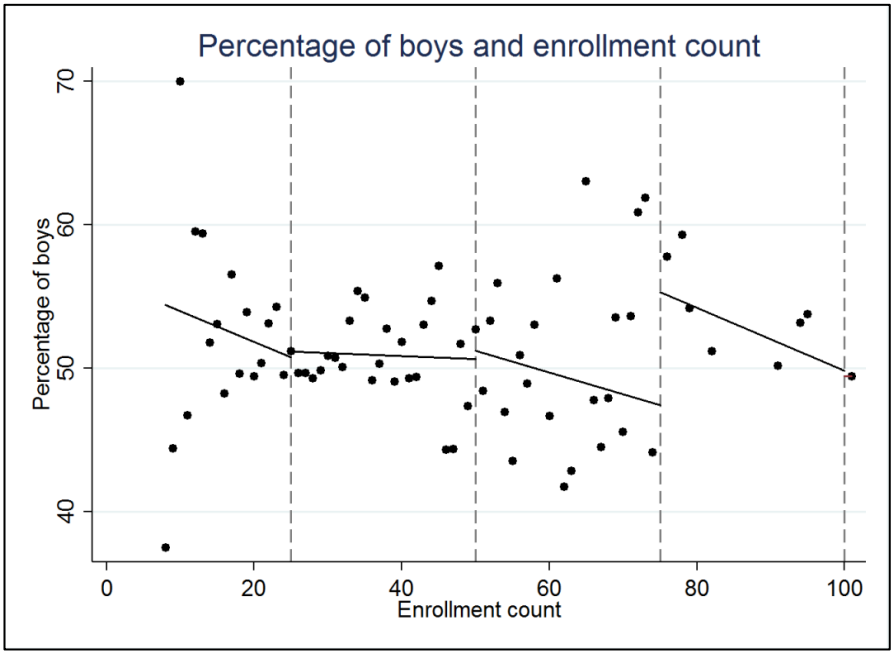

FIGURE A2 Percentage of boys and enrollment count in pooled ProKoulu data, teacher reports. The dots describe the observed average percentage of boys by grade level enrolment in school. The solid line describes a linear fit between enrollment count and percentage of boys using OLS.

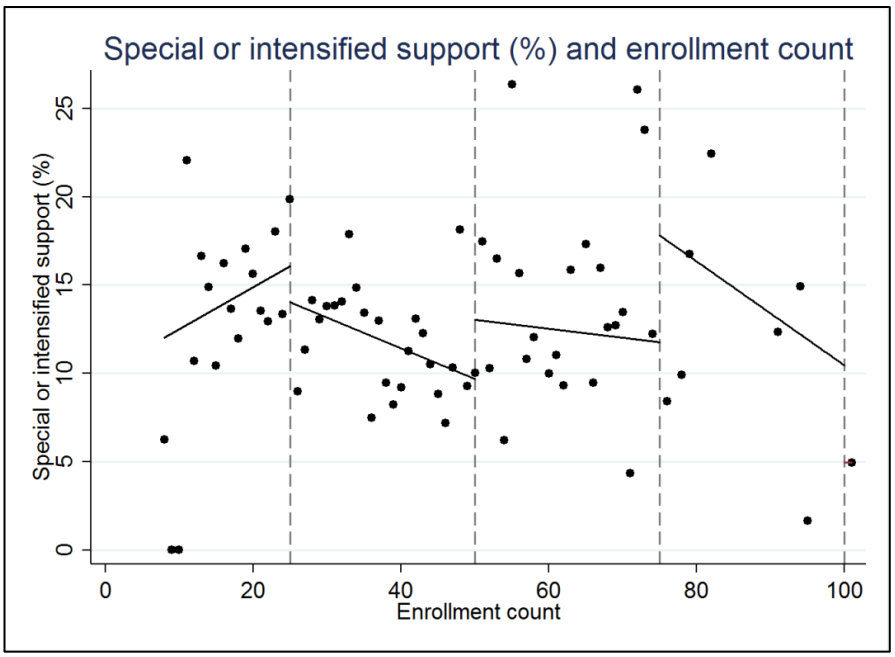

FIGURE A3 Percentage of students with special or intensified support and enrollment count in pooled ProKoulu data, teacher reports. The dots describe the observed average percentage of students with special or intensified support by grade level enrolment in school. The solid line describes a linear fit between enrollment count and percentage of students with special or intensified support using OLS.

TABLE A5 OLS, Reduced form and IV estimates: Class size and discipline problems, choice of enrollment functions in pooled ProKoulu data.

	Teacher reports		Student reports	
	(1)	(2)	(3)	(4)
Model	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)
OLS				
Class size	0.045***	0.044***	0.032***	0.031***
	(0.009)	(0.009)	(0.007)	(0.007)
REDUCED FORM				
Expected class size	0.044***	0.042***	0.020*	0.018
	(0.013)	(0.013)	(0.012)	(0.012)
IV				
2SLS				
Class size	0.085***	0.083***	0.041*	0.038
	(0.026)	(0.027)	(0.023)	(0.024)
FIRST STAGE	Class size	Class size	Class size	Class size
Expected class size	0.520***	0.508***	0.498***	0.475***
	(0.081)	(0.080)	(0.083)	(0.083)
F-statistics	41.22	40.37	36.18	32.68
CONTROLS				
Questionnaire round and functions of enrollment	x		x	
Questionnaire round and piecewise linear trend		x		x
Other controls	x	x	x	x
N	1,377	1,377	15,552	15,552
Clusters	47	47	46	46

Notes: *S* = standardized, *N* = number of observations. Functions of enrollment included enrollment and enrollment^2. Other controls included grade level, teacher's work experience, dummy for treatment group, percentages of students with special or intensified support decision and boys, and size of the school. Standard errors adjusted for clustering by school level are in parenthesis. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

TABLE A6 OLS, Reduced form and IV-estimates: Class size and discipline problems, socioeconomic status included in pooled ProKoulu data.

	Teacher reports		Student reports	
	(1)	(2)	(3)	(4)
Model	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)
OLS				
Class size	0.058***	0.058***	0.034***	0.033***
	(0.011)	(0.011)	(0.007)	(0.007)
REDUCED FORM				
Expected class size	0.047***	0.046***	0.024*	0.022*
	(0.017)	(0.017)	(0.012)	(0.012)
IV				
2SLS				
Class size	0.087***	0.087***	0.049**	0.046*
	(0.031)	(0.032)	(0.024)	(0.025)
FIRST STAGE	Class size	Class size	Class size	Class size
Expected class size	0.540***	0.530***	0.494***	0.472***
	(0.094)	(0.095)	(0.087)	(0.088)
F-statistics	32.90	31.31	33.15	29.70
CONTROLS				
Questionnaire round and functions of enrollment	x		x	
Questionnaire round and piecewise linear trend		x		x
Other controls	x	x	x	x
N	787	787	13,507	13,507
Clusters	46	46	46	46

Notes: *S* = standardized, *N* = number of observations. Functions of enrollment included enrollment, enrollment^2 and enrollment^3. Other controls included grade level, teacher's work experience, dummy for treatment group, percentages of students with special or intensified support decision and boys, size of the school, and socioeconomic status of the students. Standard errors adjusted for clustering by school level are in parenthesis. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

TABLE A7 OLS, Reduced form and IV-estimates: Class size and discipline problems, approach from Fredriksson et al. (2013), bandwidth ±12. Pooled ProKoulu data.

	Teacher reports		Student reports	
	(1)	(2)	(3)	(4)
Model	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)
OLS				
Class size	0.047***	0.044***	0.035***	0.028***
	(0.009)	(0.011)	(0.008)	(0.008)
REDUCED FORM				
Above threshold	-0.522***	-0.590***	-0.270*	-0.184
	(0.151)	(0.158)	(0.152)	(0.156)
IV				
2SLS				
Class size	0.122***	0.145***	0.065*	0.076
	(0.042)	(0.055)	(0.037)	(0.071)
FIRST STAGE	Class size	Class size	Class size	Class size
Above threshold	-4.274***	-4.079***	-4.146***	-2.417***
	(1.192)	(1.059)	(1.144)	(1.066)
F-statistics	12.85	14.83	13.13	5.14
CONTROLS				
Questionnaire round and Functions of enrollment				
1st order polynomial	x	x	x	x
Interacted with segments		x		x
Interacted with threshold		x		x
Other controls	x	x	x	x
Bandwidth	12.5	12.5	12.5	12.5
N	1,342	1,342	15,298	15,298
Clusters	47	47	45	45

Notes: *S* = standardized, *N* = number of observations. Other controls included grade level, teacher's work experience, dummy for treatment group, and enrollment segments, percentages of students with special or intensified support decision and boys, and size of the school. Standard errors adjusted for clustering by school level are in parenthesis. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

TABLE A8 OLS, Reduced form and IV estimates: Class size and discipline problems, approach from Fredriksson et al. (2013), bandwidth ±10. Pooled ProKoulu data

	Teacher reports		Student reports	
	(1)	(2)	(3)	(4)
Model	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)	Discipline problems (S)
OLS				
Class size	0.047***	0.040***	0.037***	0.025***
	(0.009)	(0.011)	(0.007)	(0.008)
REDUCED FORM				
Above threshold	–0.556***	–0.589***	–0.291	–0.113
	(0.195)	(0.175)	(0.183)	(0.172)
IV				
2SLS				
Class size	0.173**	0.209*	0.086	0.081
	(0.077)	(0.108)	(0.053)	(0.138)
First stage	Class size	Class size	Class size	Class size
Above threshold	–3.223***	–2.815**	–3.402***	–1.408
	(1.272)	(1.205)	(1.204)	(1.195)
F-statistics	6.42	5.45	7.99	1.39
CONTROLS				
Questionnaire round and Functions of enrollment				
1st order polynomial	x	x	x	x
Interacted with segments		x		x
Interacted with threshold		x		x
Other controls	x	x	x	x
Bandwidth	10	10	10	10
N	1,083	1,083	12,580	12,580
Clusters	46	46	44	44

Notes: *S* = standardized, *N* = number of observations. Other controls included grade level, teacher's work experience, dummy for treatment group, and enrollment segments, percentages of students with special or intensified support decision and boys, and size of the school. Standard errors adjusted for clustering by school level are in parenthesis. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

3 BUSINESS SCHOOL QUALITY AND EARLY CAREER OUTCOMES

Abstract[*]

This study examined the returns of admission to or enrolling in a more selective business school over a less selective one. The returns were measured in terms of higher income and fewer days of unemployment in the early stages of a career and in terms of a higher probability of completing a master's degree at a business school. For this purpose, administrative data were merged with university-level application data from the centralized admission system in the fields of business and economics. We used the regression discontinuity design, which exploited admission cutoffs generated by the admission process. The results indicated that admission to a more selective business school translates to fewer days of unemployment and a higher probability of degree completion. Regarding income, the estimates were positive but statistically insignificant.

3.1 Introduction

The connections between institutional quality and economic returns have been studied extensively. Providing causal evidence on this question is challenging, however, as attendance at a more selective institution is likely to be correlated with unobserved differences in applicant characteristics (e.g., ability) that also affect economic returns. Earlier research has attempted to solve this selection bias arising from unobserved applicant characteristics with quasi-experimental designs, exploiting, for example, discontinuities created by admission cutoffs. The

[*] The author gratefully acknowledges the OP Group Research Foundation (project numbers 20170093 and 20180102) for the financial support. She is also grateful to Roope Uusitalo, Tiina Kuuppelomäki, and Kristian Koerselman, as well as to the seminar participants at the Labour Institute for Economic Research and the Labor/Public PhD workshop for the invaluable comments regarding this paper.

discontinuity-based research design has been applied, for example, by Abdulkadiroglu et al. (2014), Hoekstra (2009), Kirkeboen et al. (2016), Kuuppelomäki et al. (2019), and Zimmerman (2019). In this study, a similar approach was exploited to investigate whether admission to or enrollment in a more selective institution assures graduation, yields higher income, or results in a lower number of unemployment days.

This study makes three primary contributions to the existing literature. First, the centralized admission system and educational processes in the fields of business and economics provide an exceptionally clear research setting. Under the centralized admission system, all applicants apply with the same entrance exam, as the institutions use the same admission criteria. The applicants further graduate with the same degree. Second, the research setting allowed for an analysis of the causal effects of being admitted to or enrolling in a more selective business school over a less selective one by using a quasi-experimental research design. By concentrating on a specific field, it was further possible to control for the field of study effects when investigating institutional effects, which was crucial because different fields have been shown to yield different payoffs (see, e.g., Kirkeboen et al., 2016). Finally, in the fields of business and economics, business schools are oversubscribed and therefore provide a more competitive fields of study to analyze than, for example, science, technology, engineering, and mathematics, which Kuuppelomäki et al. (2019) studied earlier. Because the number of applicants greatly exceeds the number of available slots, a large number of applicants are not accepted by their first-choice business school. Application data from the centralized admission system also included information about the second and third preferred business schools, allowing us to observe what happened to the applicants who were not accepted to the most competitive institution.

The existing literature has illustrated a wide range of returns arising from institutional quality. For example, Hoekstra (2009) determined an approximately 20% increase in earnings due to attending the most selective state university, and Anelli (2020) reported a premium of almost 250,000 euros earned over 15 years for those admitted to an elite institution. Admission to an elite school has shown to be beneficial especially for students from low-income (Dale & Krueger, 2002) or low-education families (Kuuppelomäki et al., 2019). On the other hand, admission to a business-focused elite degree program considerably raised the number of held leadership positions and the probability of attaining income in the top 0.1% of the distribution only for male students from private high schools (Zimmerman, 2019). Öckert (2010) reported statistically insignificant earning differences between admitted and rejected applicants in high- and low-quality institutions, whereas Kirkeboen et al. (2016) suggested that institutional effects are small compared with field of study effects. Anelli (2020) showed that students just above the admission cutoff are more likely to complete a university degree and graduate on time. As for employment, Kuuppelomäki et al. (2019) found no gains for those admitted to an elite institution, whereas Saavedra (2009) reported

a 16% higher likelihood of being employed for applicants just above the top-ranked college's admission cutoff relative to applicants just below the cutoff.

The remainder of the article is organized as follows. Section 3.2 presents the institutional background, including higher business education in Finland and the centralized admission system in the fields of business and economics. Section 3.3 describes the data and the methodology and discusses the validity of the research design. Section 3.4 presents the results, and Section 3.5 discusses and concludes the paper.

3.2 Institutional background

3.2.1 Higher business education in Finland

Finnish higher education is mainly government funded and free of charge for citizens of the European Union (EU) and the European Economic Area (EEA) countries. Most students receive financial support, including a housing supplement, a study grant, and a government-guaranteed student loan. Higher education is provided by universities of applied sciences and universities. This study concentrated on education provided in universities leading to bachelor's and master's degrees in economic sciences, leaving universities of applied sciences out of consideration. Because the universities providing business education are often called business schools, the terms universities and business schools are used interchangeably in this study.

Applicants have the opportunity to complete both bachelor's and master's degrees when admitted to the university.[1] The Master of Economic Sciences degree (MSc in Economics and Business Administration, from here on abbreviated as MSc) requires completing 120 European Credit Transfer and Accumulation System (ECTS) points (two academic years) in addition to a bachelor's degree (180 ECTS points, three academic years).

Although business school students all graduate with the same degree, they can have different majors during their studies (for example, marketing, accounting, entrepreneurship, leadership and management, and economics). The selection of majors is business school specific. Applicants either apply to their major programs during the application process or choose their major during their studies (for example, after one year of studies). Regardless of the major selection procedure, an applicant must first be admitted to a business school. Only after passing the admission cutoff for a business school can an applicant be admitted to or choose a major.

[1] It is further possible to apply only and directly to a master's program if an applicant has completed the appropriate prerequisites, but such applicants were excluded from this study.

3.2.2 Centralized admission system in business education

In Finland, students are selected for higher education through a joint application system. The centralized admission system in the fields of business and economics was established within the larger joint application system in 2005. It was implemented by the Helsinki School of Economics, the University of Joensuu, the University of Kuopio, the University of Lapland, the Lappeenranta University of Technology, the University of Oulu, the University of Tampere, the University of Turku, and the University of Vaasa. These institutions agreed to apply the same selection criteria in their admission process. The University of Jyväskylä, the Hanken School of Economics, and Åbo Akademi University continued to use their own admission systems instead of joining the centralized admission system. The centralized admission system is now used among all institutions offering university-level bachelor's and master's programs in economics and business administration. The data utilized in this study cover the first four years (2005–2008) of the admission system.

All applicants applying through the centralized admission system in 2005–2008 submitted their applications in March. The applicants could apply to a maximum of three business schools, which they ranked according to their personal preferences. The submitted preference order was binding and could not be changed during the application process. Later in spring, all the applicants took the same entrance exam. The applicants were assigned to business schools through a centrally run business school-proposing deferred acceptance algorithm. Applicants with the highest application points received an offer for their first university preference. Applicants with the second highest points received an offer for their first university preference if there was a place left for another student. If there were no slots available at the first preferred university, the second preferred university was considered, and so on. By early July, each applicant had received a single offer, which they accepted or rejected, or no offer at all.

The students were admitted through two quotas, for which applicants automatically applied. The total points quota was based on success on entrance and matriculation exams, whereas the test points quota was based on the entrance exam only. The matriculation exam was taken at the end of general upper secondary education. In this national examination, the grades were standardized, and they described the students' position in a national distribution. The entrance exam included multiple-choice questions concerning business-related exam material, which was assigned beforehand. Overall, 60% of the applicants were admitted through the total points quota, and the rest were admitted through the test points quota. In both quotas, additional points were awarded for recent high school graduates and for the two highest ranked universities in the preference order.

Admission cutoffs were determined according to the points of the applicants who received and accepted the offers for the last available slots of the business schools. Therefore, the cutoffs depended on the quality of the applicants,

number of slots, and number of applicants. The admission cutoffs differed across business schools, as shown in Figure 1. In Figure 1, the average university admission cutoff is illustrated by the light gray bar in the total points quota and by the dark gray bar in the test points quota. The highest admission points were required in the Helsinki School of Economics and lowest in the University of Lapland for both quotas. These admission cutoffs are employed as a measure for the quality of an institution, allowing for comparisons between marginally accepted and rejected applicants and enabling causal inference about applicants' subsequent early career outcomes.

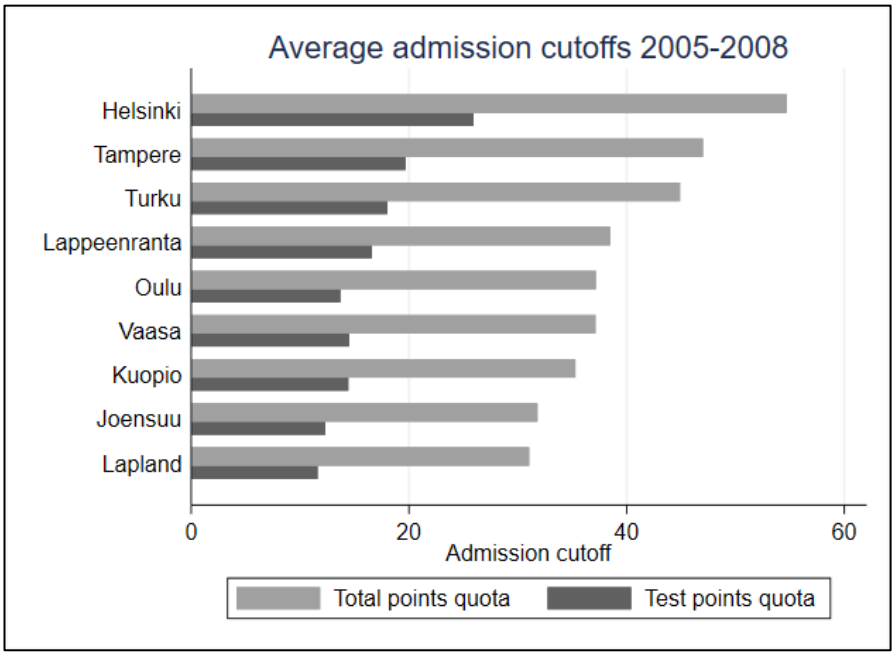

FIGURE 1 Average admission cutoffs for the years 2005–2008 by business school and quota. Total points quota is based on entrance and matriculation exams and test points quota on the entrance exam. Extra points were granted to recent high school graduates and for the two highest ranked universities in preference order in both quotas. Admission cutoffs were determined according to the points of the applicants who received and accepted the offers for the last available slots of the business schools.

Selectivity is not the only possible measure for institutional quality. Well-known rankings, such as the Shanghai Academic Ranking of World Universities or QS World University Rankings, represent other possible estimates for institution quality. Table 1 shows the recent national placements of the studied institutions in these rankings. Recent rankings were selected over the rankings in 2005–2008, as very few institutions were ranked at the time. The last two columns show subject-specific national rankings for business administration and economics. Note that the Helsinki School of Business and Economics is currently called Aalto University and that Kuopio and Joensuu universities are part of the University of Eastern Finland.

The obvious challenge with the academic rankings is that not all the business schools are ranked. The rankings that were available are in line with the institutional quality measure employed in the present study. Helsinki School of Economics is the highest ranked university in three out of four studied rankings and top ranked in both subject-specific rankings. Similarly, the universities of Turku and Oulu as well as Lappeenranta University of Technology are in the top five of the most selective institutions measured by average admission cutoffs, and these institutions also appear in almost all four rankings.

TABLE 1 National rankings of the business schools in Finland in academic rankings

Rank	Shanghai Academic Ranking of World Universities 2020	QS World University Ranking 2020	Shanghai Global Ranking of Academic Subjects, Business Administration 2021	Shanghai Global Ranking of Academic Subjects, Economics 2021
1.	University of Turku	Helsinki School of Economics	Helsinki School of Economics	Helsinki School of Economics
2.	Helsinki School of Economics	University of Turku	University of Turku	Universities of Kuopio and Joensuu
3.	Universities of Kuopio and Joensuu	University of Oulu	University of Vaasa	University of Oulu
4.	University of Oulu	Universities of Kuopio and Joensuu	Lappeenranta University of Technology	University of Turku
5.	University of Tampere	Lappeenranta University of Technology		
6.	Lappeenranta University of Technology	University of Tampere		

Note: Helsinki School of Economics is now Aalto University, and Joensuu and Kuopio universities are part of the University of Eastern Finland.

3.3 Data and methodology

3.3.1 Data, variables, and the estimation sample

In this study, application data from the centralized admission system in business and economics were merged with Finnish administrative data. Application data included information about business school applicants from 2005–2008. The data covered entrance and matriculation examination points, preference order of the institutions applied to, possible admission and enrollment for each year, and quota. In the present setting, admission refers to a received offer and enrollment to an accepted offer.

The administrative data included information about the outcome variables, which were income, unemployment days, and completion of the MSc. The income measure was a mean of yearly earnings six, seven, and eight years after admission to a business school (hereafter referred to as early career income), and it was adjusted for inflation. Similarly, unemployment was measured as a mean amount of unemployment days six, seven, and eight years after admission (hereafter referred to as early career unemployment). Degree completion was an indicator variable and was assigned a value of 1 if an admitted applicant completed an MSc by eight years after admission and 0 otherwise. Applicants' background characteristics (age, gender, parental income and education, mean matriculation examination grade of mathematics and mother tongue, level of mathematics,[2] and county one year before application) were also drawn from the administrative data.

In the analyses, those admitted to a more selective business school were compared with those admitted to a less selective business school. Therefore, each applicant included in the estimation sample had applied to at least two business schools in a given year and was then admitted to one of the schools.

The application data included a total of 5,284 admissions from 2005–2008.[3] Of the admitted applicants, 855 applied to one business school only and were therefore excluded from the sample. For 2,200 admissions, the applicant listed at least two business schools on the preference order, and the preference order corresponded to the actual selectivity order of the business schools. For these admissions, there further were at least two observations each side of the cutoff in the given year. Out of these 2,200 admissions, outcome variables and background characteristics were available for 1,849 admitted applicants. The admissions were pooled over years and business schools, meaning that university cutoffs were stacked to one single cutoff. The pooled sample meeting all the requirements presented above included 2,642 observations from 1,849 individuals. The between-university pairs included in the estimation sample are presented in Table 2.

[2] In the matriculation examination, the mathematics test has two levels of difficulty, advanced and basic, from which the candidates can choose.
[3] The ones who applied only to the master's program were excluded.

TABLE 2 Between-university pairs included in the estimation sample

Universities above the cutoff	**Universities below the cutoff**
Helsinki School of Economics	University of Joensuu
Helsinki School of Economics	University of Kuopio
Helsinki School of Economics	University of Lapland
Helsinki School of Economics	University of Lappeenranta University of Technology
Helsinki School of Economics	University of Oulu
Helsinki School of Economics	University of Tampere
Helsinki School of Economics	University of Turku
Helsinki School of Economics	University of Vaasa
University of Tampere	University of Kuopio
University of Tampere	University of Lapland
University of Tampere	University of Oulu
University of Tampere	University of Turku
University of Tampere	University of Vaasa
University of Tampere	University of Lappeenranta University of Technology
University of Turku	University of Joensuu
University of Turku	University of Oulu
University of Turku	University of Vaasa
University of Oulu	University of Lapland
University of Vaasa	University of Lapland
Lappeenranta University of Technology	University of Joensuu

In the pooled sample, applicants could appear more than once. The following example clarifies which applicants were included in the estimation sample and in which cases they entered the data more than once.

Consider three example admission cutoffs: Helsinki, Tampere, and Turku. Helsinki had the highest admission cutoff and Turku the lowest. If an applicant applied to these three institutions in that order, the possible admission outcomes and appearances in the estimation sample would be as follows:

1. Outcome: The applicant's admission points were below all three cutoffs, and therefore, they were not admitted to any business schools and did not appear in the sample.
2. Outcome: The applicant passed Turku's admission cutoff. This applicant was compared with those who applied to Turku and Tampere and were just above the Tampere cutoff, as well as with those who applied to Turku and Helsinki and were just above the Helsinki cutoff. Thus, this applicant appeared twice in the sample.
3. Outcome: The applicant passed the Tampere cutoff. There were then two observations of this applicant in the sample, one below the Helsinki cutoff (compared with those who applied to Tampere and Helsinki and were just above the Helsinki cutoff) and one above the Tampere cutoff (compared with those who applied to Turku and Tampere and were just below the Tampere cutoff).

4. Outcome: The applicant passed the Helsinki cutoff. They then appeared twice in the sample, above the Helsinki cutoff (compared with those who applied to Tampere and Helsinki and were just below the Helsinki cutoff as well as compared with those who applied to Turku and Helsinki and were just below the Helsinki cutoff).

3.3.2 Descriptive statistics

The summary statistics for the pooled estimation sample are presented in Table 3. Mean early career income was 34,004 euros per year. Early career unemployment was, on average, nine days per year, and 67% of the admitted applicants in the estimation sample completed an MSc within eight years after admission. Almost 60% of the applicants were male and, on average, 21 years old. In the matriculation examination, 67% of the applicants had participated in the advanced-level mathematics test. Their mean mathematics and mother tongue grade was 4.85 on a scale ranging from 1 to 7, with 1 describing failed and 7 representing the highest possible grade. The sum of parental income was on average 118,229 euros per year, and at least one of the applicant's parents had completed a bachelor's degree or higher in 65% of the cases. Of the applicants in the estimation sample, 63% lived in Uusimaa one year before admission.

TABLE 3 Summary statistics for the pooled estimation sample (N = 2,642)

Variable	*M*	*SD*
Early career income, euros	34,004	19,516
Early career unemployment, days	9.16	31.19
MSc completion, share	0.67	0.47
Male, share	0.57	0.49
Age, years	21.03	2.78
Advanced mathematics, share	0.67	0.47
Mean grade, mother tongue and mathematics	4.85	1.10
Sum of parental income, euros	118,229	94,742
Parental education	0.65	0.48
From Uusimaa, share	0.63	0.48

Notes: Early career income and unemployment days are means from corresponding yearly values six, seven, and eight years after admission. The mean grade of mother tongue and mathematics is from the matriculation examination and varies from 1 to 7, with 1 describing failed and 7 representing the highest possible grade. Parental education is an indicator variable, which was assigned a value of 1 if at least one of the applicant's parents had completed a bachelor's degree or higher and 0 otherwise. Uusimaa is a county on the south coast of Finland where approximately 30% of the Finnish population lives. Finland's capital, Helsinki, is also situated in Uusimaa.

It is common to start working during studies in Finland. According to Official Statistics of Finland (2021), half of the students aged over 18 years were working in 2019, which could be one of the explanations for why only 70% of the students completed higher education in seven years (Ministry of Education and Culture, 2019). In the present sample, almost 70% of the applicants had completed their MSc by eight years after admission.

Graduates from the fields of business and economics obtain good employment situations. In 2017–2018, 81% of graduates in the fields of business, administration, and law were employed within one year after graduation (Official Statistics of Finland, 2018). Business and education graduates also belong to the top third-earning fields among the university-level degrees (Suhonen & Jokinen, 2018). The median of total earnings among highly educated Finns aged 25–29 was 3,389 euros per month in 2019 (Official Statistics of Finland, 2021). The corresponding median in the present sample was 3,616 euros per month after eight years of admission among those who completed an MSc.

3.3.3 Methodology

The selection bias related to unobserved applicant-specific characteristics can be eliminated with a discontinuity-based research design. Therefore, a regression discontinuity design was adopted for the present study investigating the causal effect of attending a more selective business school over a less selective one. Forming the model and the running variable, r_{ikqt}, followed Abdulkadiroglu et al. (2014). In forming the running variable, each applicant *i* was given a rank c_{ikqt} based on the admission points in given year *t,* business school *k,* and quota *q*. This rank was from the quota that was the least competitive from the applicant's perspective (how many applicants there were to pass in the quota). The running variable was centered, as the admission cutoffs were year-, business school-, and quota-specific, and because the aim was to compare and pool the cutoffs. Centering was done by demeaning the rank of the last admitted applicant, τ_{kqt}, in the same business school *k*, quota *q*, and application year *t* and from c_{ikqt}.

In Equation 1, the running variable, r_{ikqt}, describes an applicant's centered distance to the threshold ($c_{ikqt} - \tau_{kqt}$) in percentages of the number of applicants (N_{kt}) to a business school *k* in year *t*. The running variable receives a positive value when an applicant's admission points exceed the admission cutoff and a negative value otherwise. The positive value of the running variable also means that an applicant received an offer from a more selective institution. This is because the present study compares those admitted to a more selective institution with those admitted to a less selective institution. The running variable is presented as follows:

$$r_{ikqt} = \frac{100}{N_{kt}}\left(c_{ikqt} - \tau_{kqt}\right). \quad (1)$$

After forming the running variable, the discontinuities utilized in the regression discontinuity design were pooled. In the pooled sample, the cutoffs were stacked,

and the coefficient of interest was a single admission effect, describing average cutoff-specific estimates. The causal effect of being admitted to a better business school on outcome variables (income, unemployment, or graduation) was identified by using the following reduced form model specification:

$$y_{ikqt} = \beta_1 d_{ikqt} + \beta_2 r_{ikqt} + \beta_3 d_{ik} r_{ikq} + \delta_{kqt} + \varepsilon_{ikqt}, \quad (2)$$

where y_{ikqt} is the outcome variable for applicant i that was admitted to business school k in quota q in year t; d_{ikqt} is a dummy variable of being admitted to a more selective business school ($r_{ikqt} \geq 0$), and β_1 is the coefficient of interest; r_{ikqt} is the centered running variable; the $d_{ikqt} r_{ikqt}$ interaction term allows for the slope of the running variable to change on each side of the cutoff; δ_{kqt}controls for the year and cutoff fixed effects (school-application-cohort fixed effects); and standard errors ε_{ikqt} are clustered at the applicant level, as applicants can appear more than once in the data. All the analyses further include dummies for the university pairs the applicants applied for. In the main analyses, the estimations also included interactions between application year and cutoffs and interactions between the running variable and the cutoffs, which allowed the slope of the running variable to differ by the cutoff.

In addition to the reduced form estimates, the results using the fuzzy regression discontinuity design are presented in Section 3.4. In this fuzzy regression discontinuity setting, admission to a more selective business school was used as an instrument for enrollment. Enrollment was defined as an applicant accepting the offer, and receiving an offer was assumed to affect outcomes only through enrollment. The optimal bandwidths were defined following Calonico et al. (2014) in all models. The sample was weighted using the triangular kernel function:

$$K_h(r_{ikqt}) = 1\left\{\left[\frac{r_{ikqt}}{h}\right] \leq 1\right\} * \left(1 - \left|\frac{r_{ikqt}}{h}\right|\right), \quad (3)$$

where h describes the optimal bandwidth and r_{ijkqt} the centered running variable.

The discontinuities applied in the main specification (Equation 2) and in the fuzzy regression discontinuity setting are presented graphically in Figures 2 and 3. Figure 2 illustrates the discontinuity in the probability of being admitted near the pooled cutoffs, exploited in both sharp and fuzzy regression discontinuity settings. In the present setting, we compare the applicants admitted to a more selective business school over a less selective one. Therefore, all the applicants above the cutoff were admitted, and no one below the cutoff was admitted. The ones accepted from the queue were also among the admitted applicants.

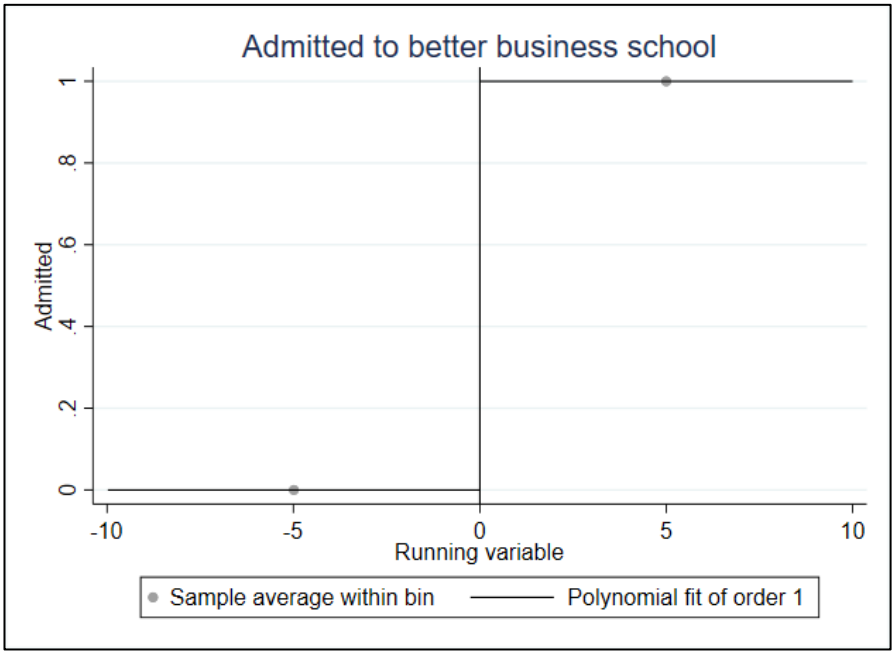

FIGURE 2 Admission to a better business school and the centralized running variable in the pooled sample. The figure shows the shares of admitted applicants below and above the pooled admission cutoffs against the centralized running variable (applicant's distance to the admission cutoff in percentages out of the number of applicants). The dots describe bin-specific means of the admitted applicants (1 bin on the left and 1 on the right; bin unit widths 9.98 and 10.00, respectively). The functions were estimated using first-order polynomials weighted with a triangular kernel.

In the fuzzy regression discontinuity design, enrollment was instrumented by admission. Figure 3 shows that just above the cutoff, approximately 80% of the admitted applicants enrolled in a better business school. The applicants below the threshold were not able to enroll because they were not admitted.

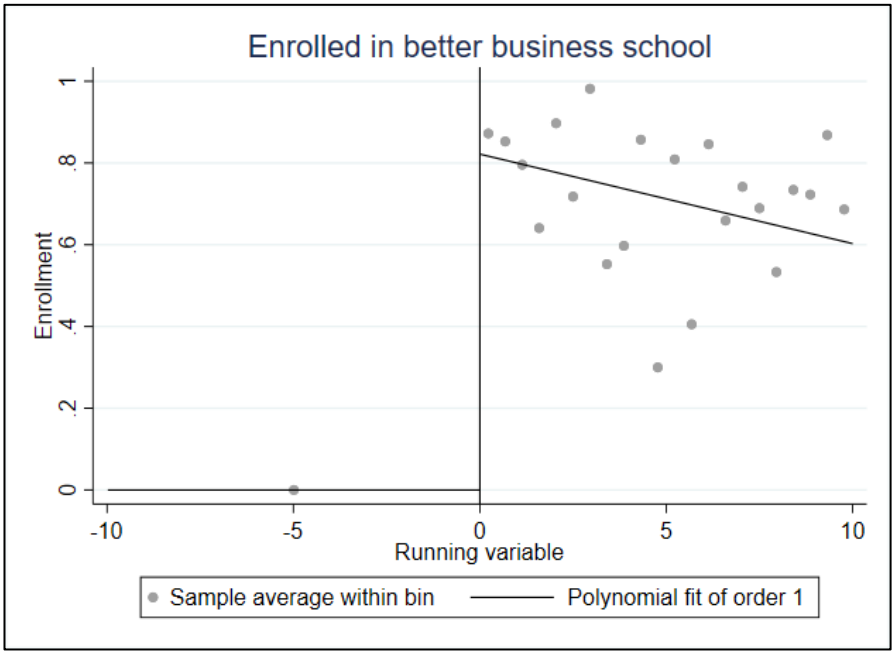

FIGURE 3 Enrollment in a better business school and the centralized running variable in the pooled sample. The figure represents the shares of applicants that enrolled in a better business school below and above the pooled admission cutoffs against the centralized running variable (applicant's distance to the admission cutoff in percentages out of the number of applicants). The dots describe bin-specific means of the enrolled applicants (1 bin on the left and 24 on the right; bin unit widths 9.98 and 0.42, respectively). The functions were estimated using first-order polynomials weighted with a triangular kernel.

3.3.3.1 Validity of the design

In the regression discontinuity setup, the validity of the design requires that the running variable cannot be manipulated. In the present setting, perfect manipulation is unlikely because neither the admission cutoffs nor success in the admission process can precisely be predicted beforehand. Additionally, recent high school graduates apply before the results of the matriculation examinations are published.

To support the reasoning above, potential manipulation of the running variable was analyzed in the estimation sample. McCrary's (2008) density test suggested no bunching in the sample (discontinuity estimate -0.09, standard error 0.12) when the observations defining the cutoff were excluded. Including applicants defining the cutoffs would naturally create a jump at the cutoff, as there was an observation with a zero value for each year, institution, and quota (these observations defined the cutoffs). The smoothness of the density of the running variable is presented graphically in the appendix (Figure A1), both with and without applicants defining the cutoffs.

Another criterion for the validity of the design was that the applicants' background characteristics were smooth across the cutoff. The balance check for the covariates was done by replacing the outcome variable in Equation 2 with each of the background characteristics. The results from the balance checks are

presented in Table 4. Both rows included dummies for year and cutoff, interactions between year and cutoff, and interactions between cutoff and running variable, as well as dummies describing the university pair to which the applicant applied. In the lower row, background characteristics other than the present outcome were included in the model as additional controls. The estimates are similar across rows one and two, showing that adding controls changed the estimates very little. The results suggest discontinuity in the mean grade (Column 3) and in parental income (Column 5). To account for these possible discontinuities, the main results in the next section are also presented with the full set of controls.

TABLE 4 Discontinuities in background characteristics for the pooled sample

	Male (1)	Age (2)	Mean grade (3)	Advanced mathematics (4)	Parental income (euros) (5)	Originated from Uusimaa (6)	Parental education (7)
Discontinuity	–0.06	0.34	–0.74***	–0.02	33,164*	0.12	-0.05
	(0.10)	(0.46)	(0.25)	(0.09)	(17,791)	(0.09)	(0.07)
Discontinuity with additional controls	–0.09	0.38	–0.73***	–0.03	32,286*	0.02	-0.04
	(0.10)	(0.44)	(0.26)	(0.09)	(18,099)	(0.09)	(0.06)
Potential outcome estimate with additional controls	0.62***	21.28***	5.48***	0.64***	112,908***	0.69***	0.65***
	(0.07)	(0.28)	(0.20)	(0.06)	(11,487)	(0.07)	(0.05)
Optimal bandwidth	5.40	7.48	3.67	6.02	4.22	4.34	9.02
N	911	1,265	579	994	708	739	1,546

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The table presents the estimates from Specification 2, using applicants' background characteristics as outcomes in the pooled sample. The optimal bandwidths were selected following Calonico et al. (2014), as shown in Equation 3. All estimations included dummies for year and cutoff, interactions between year and cutoff, interactions between cutoff and running variable, and dummies describing the university pair to which the applicant applied. In Row 2, background characteristics other than the present outcome were included in the model as additional controls. Standard errors were clustered at the applicant level and are presented in parenthesis.

3.4 Business school quality and its effects on early career outcomes

How the business school quality affects early career outcomes was first studied graphically (Figures 4, 5, and 6). In Figures 4-6, the vertical line shows the pooled admission cutoff. Applicants below the cutoff were admitted to a less selective business school than the applicants above the cutoff. The dots represent the bin-specific averages of the outcome measures, and the functions were estimated using first-order polynomials, weighted with a triangular kernel. Graphical analyses suggest that admission to a better business school increases income (Figure 4), decreases unemployment days (Figure 5), and increases the probability of completing an MSc (Figure 6).

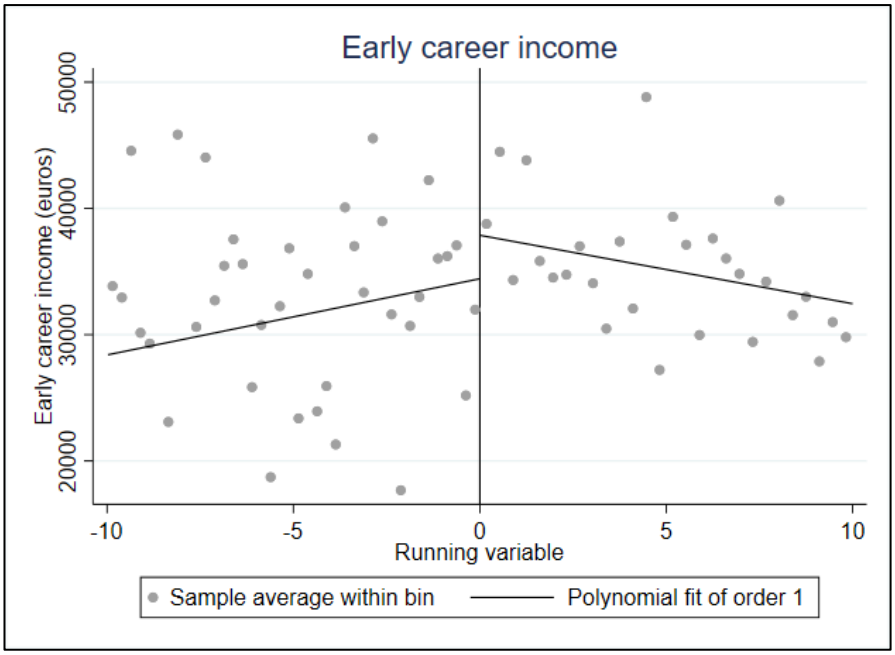

FIGURE 4 Early career income and the running variable. The vertical line shows the admission cutoff for a better business school. Applicants below the cutoff were accepted to a less selective business school than the applicants above the cutoff. The dots represent the bin-specific averages of the income measure (40 bins on the left and 28 on the right side of the cutoff; bin unit widths 0.25 and 0.36, respectively). The functions were estimated using first-order polynomials weighted with a triangular kernel.

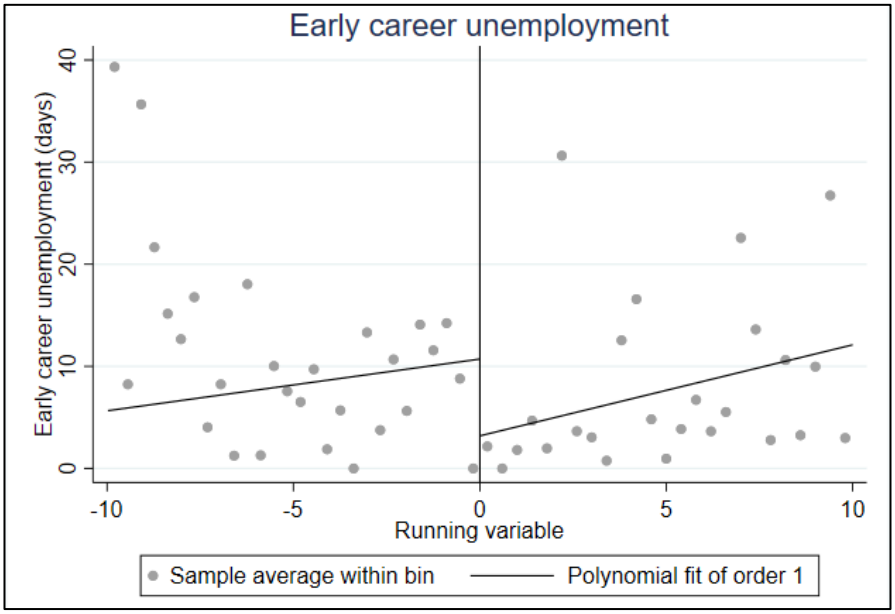

FIGURE 5 Unemployment and the running variable. The vertical line shows the admission cutoff to a better business school. Applicants below the cutoff were accepted to a less selective business school than the applicants above the cutoff. The dots represent the bin-specific averages of the income measure (28 bins on the left and 25 on the right side of the cutoff; bin unit widths 0.36 and 0.40, respectively). The functions were estimated using first-order polynomials weighted with a triangular kernel.

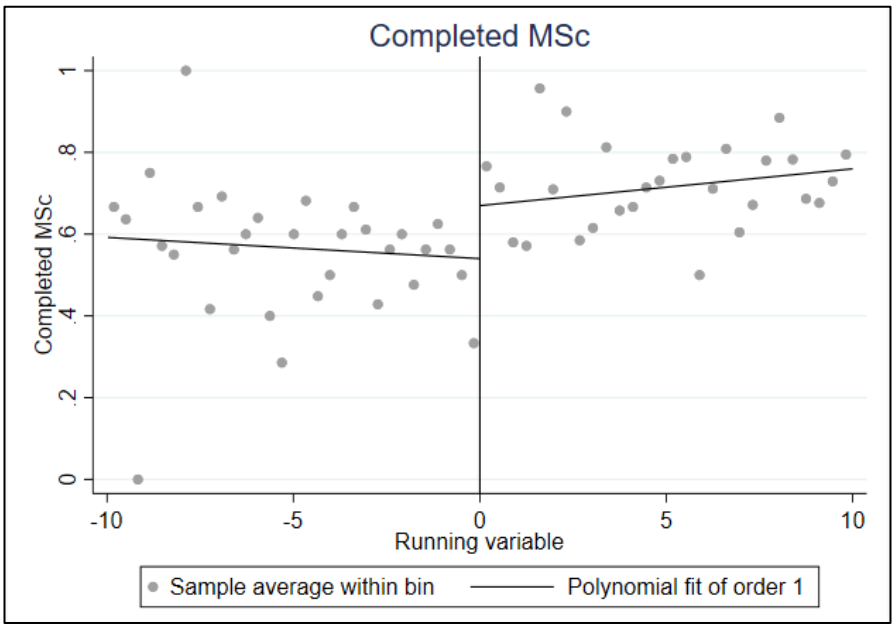

FIGURE 6 MSc completion and the running variable. The vertical line shows the admission cutoff for a better business school. Applicants below the cutoff were accepted to a less selective business school than the applicants above the cutoff. The dots represent the bin-specific averages of the income measure (31 bins on the left and 28 on the right side of the cutoff; bin unit widths 0.32 and 0.36, respectively). The functions were estimated using first-order polynomials weighted with a triangular kernel.

The main results and the robustness checks are presented in Tables 5–7. In Table 5, the outcome is early career income; in Table 6, early career unemployment; and in Table 7, MSc completion. Before presenting the estimates, the main specification and the models used in robustness checks are reviewed.

The main specification includes dummies for year and cutoff, interactions between year and cutoff, and interactions between cutoff and running variable, as well as dummies describing the university pair to which the applicant applied. Additional controls (gender, age, mean mathematics and mother tongue matriculation examination grade, dummy for advanced mathematics, county one year before admission, parental income and education) are included in Columns 2–7. To gain more statistical power, a less flexible model (Column 3) was used to exclude interactions between cutoff and running variable. Donut (Column 4) refers to a specification where the applicants defining the cutoffs (i.e., zeros) are excluded, accounting for possible endogeneity of the admission cutoff definition. The nearest threshold (Column 5; see Section 3.1 for details) uses only the two nearest thresholds applied by an applicant. In Column 6, the required number of observations on each side of the cutoff is increased from two to four. In Column 7, the cutoff-specific fixed effects and interactions were excluded.

The optimal bandwidths were selected following Calonico et al. (2014) in Columns 1–7. The final column uses a fixed bandwidth for all the outcomes, calculated as the mean of the optimal bandwidths used in the main specifications (Column 1). Additional robustness checks for bandwidth selection are presented graphically in the appendix (Figures A3–A5). According to the figures, the results were robust for a wide range of bandwidths.

Table 5 reports the results using early career income as an outcome variable. The reduced form estimates illustrating the admission effects vary from 1,439 to 3,738 euros per year, and the local average treatment estimates (LATEs) demonstrating the enrollment effects vary from 1,188 to 5,412 euros per year. Standard errors of these estimates were too high for causal inference.

For early career unemployment (Table 6), the admission to a more selective business school over a less selective one seems to matter. The reduced form estimates in the first row suggest that crossing the cutoff decreases unemployment by approximately 8 to 14 days per year. The next two rows present the results from the fuzzy regression discontinuity model. First-stage estimates show that admission to a better business school increases the enrollment rate. LATE estimates suggest that enrollment in a better business school decreases unemployment by 11 to 22 days per year for those applicants whose enrollment decision is determined by the admission offer.

Table 7 displays the MSc completion effects. The results imply that admission to a more selective business school increases the degree completion probability by 10–17 percentage points. For those whose enrollment decision was defined by the admission offer, enrollment in a more selective business school further increased the degree completion probability by 15–25 percentage points.

TABLE 5 The effect of being admitted to or enrolling in a better business school on early career income. Regression discontinuity results.

	Main model (1)	Main model with controls (2)	Less flexible model (3)	Donut (4)	Nearest thresholds (5)	At least 4 observations/side (6)	Without cut-off-specific fixed effects and interactions (7)	Fixed bandwidth (8)
REDUCED FORM Admitted to a better business school	1,439 (4,099)	2,205 (3,900)	2,070 (3,881)	3,738 (4,271)	847 (3,676)	2,290 (3,902)	1,970 (3,815)	1,986 (3,408)
IV	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
FIRST STAGE Admitted to a better business school	0.68*** (0.05)	0.69*** (0.05)	0.71*** (0.05)	0.69*** (0.05)	0.71*** (0.04)	0.69*** (0.05)	0.86*** (0.04)	0.70*** (0.04)
LATE Enrolled in a better business school	2,109 (5,906)	3,186 (5,486)	2,936 (5,370)	5,412 (6,085)	1,188 (4,950)	3,304 (5,511)	3,562 (3,999)	2,822 (4,734)
Potential outcome estimate	37,066*** (3,189)	37,566*** (3,129)	37,647*** (3,115)	36,980*** (3,125)	37,898*** (3,073)	37,543*** (3,132)	37,849*** (3,027)	37,017*** (2,736)
Additional controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Optimal bandwidth	5.70	5.70	5.70	5.63	6.55	5.83	5.70	7.066
N	957	957	957	910	763	932	957	1,167

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The table shows the estimates from the specification presented in Equation 2 in the pooled sample with early career income as an outcome. Early career income is a mean of yearly earnings six, seven, and eight years after admission to a business school. Po-

tential outcome estimate predicts the outcome at the cutoff when $r_{ikq} \to 0-$; that is, it describes the value of the outcome at the cutoff without admission. The main specification includes dummies for year and cutoff, interactions between year and cutoff, and interactions between cutoff and running variable, as well as dummies describing the university pair to which the applicant applied. Additional controls (gender, age, mean mathematics and mother tongue matriculation examination grade, dummy for advanced mathematics, county one year before admission, parental income and education) are included in Columns 2–7. A less flexible model excludes interactions between cutoff and running variable. Donut specification excludes the applicants defining the cutoffs (i.e., zeros). The nearest threshold uses only the two nearest thresholds applied by an applicant. In Column 6, the required number of observations on each side of the cutoff is increased to four. In Column 7, the cutoff-specific fixed effects and interactions were excluded. The final column uses a fixed bandwidth for all the outcomes, calculated as a mean of the optimal bandwidths used in the main specifications (Column 1). The optimal bandwidths were selected following Calonico et al. (2014), as shown in Equation 3, and standard errors were clustered at the applicant level (in parentheses).

TABLE 6 The effect of being admitted to or enrolling in a better business school on early career unemployment. Regression discontinuity results.

	Main model (1)	Main model with controls (2)	Less flexible model (3)	Donut (4)	Nearest thresholds (5)	At least 4 observations/side (6)	Without cut-off-specific fixed effects and interactions (7)	Fixed bandwidth (8)
REDUCED FORM Admitted to a better business school	–8.59** (3.74)	–8.95** (3.80)	–8.49*** (2.96)	–14.02** (6.38)	–12.02** (5.64)	–10.44** (4.64)	–8.88*** (3.08)	–8.95** (3.93)
IV	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
FIRST STAGE Admitted to a better business school	0.70*** (0.04)	0.70*** (0.04)	0.72*** (0.04)	0.66*** (0.07)	0.69*** (0.05)	0.69*** (0.05)	0.85*** (0.04)	0.70*** (0.04)
LATE Enrolled in a better business school	–11.39** (5.08)	–12.37** (5.08)	–12.61** (4.96)	–21.72** (9.12)	–17.45** (7.46)	–14.35** (6.32)	–11.57*** (4.06)	–12.39** (5.27)
Potential outcome estimate	11.22*** (3.14)	12.14*** (3.44)	12.11*** (2.53)	13.40** (6.15)	12.81** (5.42)	12.69*** (4.30)	12.46*** (2.65)	12.22*** (3,71)
Additional controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Optimal bandwidth	7.49	7.49	7.49	4.39	4.88	5.78	7.49	7.066
N	1,265	1,265	1,265	703	583	921	1,265	1,167

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The table shows the estimates from the specification presented in Equation 2 in the pooled sample with early career unemployment as an outcome. Early career unemployment is a mean of yearly unemployment days six, seven, and eight years after admission to a business school. The potential outcome estimate predicts the outcome at the cutoff when $r_{ikq} \rightarrow 0-$; that is, it describes the value of the

outcome at the cutoff without admission. The main specification includes dummies for year and cutoff, interactions between year and cutoff, and interactions between cutoff and running variable, as well as dummies describing the university pair to which the applicant applied. Additional controls (gender, age, mean mathematics and mother tongue matriculation examination grade, dummy for advanced mathematics, county one year before admission, parental income and education) are included in Columns 2–7. A less flexible model excludes interactions between cutoff and running variable. Donut specification excludes the applicants defining the cutoffs (i.e., zeros). The nearest threshold uses only the two nearest thresholds applied by an applicant. In Column 6, the required number of observations on each side of the cutoff is increased to four. In Column 7, the cutoff-specific fixed effects and interactions were excluded. The final column uses a fixed bandwidth for all the outcomes, calculated as a mean of the optimal bandwidths used in the main specifications (Column 1). The optimal bandwidths were selected following Calonico et al. (2014), as shown in Equation 3, and standard errors were clustered at the applicant level (in parentheses).

TABLE 7 The effect of being admitted to or enrolling in a better business school on MSc completion. Regression discontinuity results.

	Main model (1)	Main model with controls (2)	Less flexible model (3)	Donut (4)	Nearest thresholds (5)	At least 4 observations/side (6)	Without cut-off-specific fixed effects and interactions (7)	Fixed bandwidth (8)
REDUCED FORM Admitted to a better business school	0.17** (0.07)	0.16** (0.07)	0.17** (0.07)	0.10 (0.09)	0.15** (0.07)	0.15* (0.08)	0.15** (0.07)	0.15* (0.08)
IV	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
FIRST STAGE Admitted to a better business school	0.70*** (0.04)	0.70*** (0.04)	0.72*** (0.04)	0.70*** (0.04)	0.72*** (0.04)	0.71*** (0.04)	0.85*** (0.04)	0.70*** (0.04)
LATE Enrolled to a better business school	0.24** (0.11)	0.25** (0.10)	0.23** (0.10)	0.15 (0.12)	0.21** (0.10)	0.21* (0.11)	0.15* (0.08)	0.23** (0.11)
Potential outcome estimate	0.51*** (0.05)	0.50*** (0.05)	0.50*** (0.05)	0.49*** (0.06)	0.49*** (0.06)	0.50*** (0.06)	0.52*** (0.05)	0.50*** (0.06)
Additional controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Optimal bandwidth	8.01	8.01	8.01	6.72	7.48	6.76	8.01	7.066
N	1,356	1,356	1,356	1,081	897	1,073	1,365	1,167

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The table shows the estimates from the specification presented in Equation 2 in the pooled sample with MSc completion within eight years after admission as an outcome. Potential outcome estimate predicts the outcome at the cutoff when $r_{ikq} \to 0-$; that is, it describes the value of the outcome at the cutoff without admission. The main specification includes dummies for year and cutoff, interactions between year and cutoff, and interactions between cutoff and running variable, as well as dummies describing the university pair to which the

applicant applied. Additional controls (gender, age, mean mathematics and mother tongue matriculation examination grade, dummy for advanced mathematics, county one year before admission, parental income and education) are included in Columns 2–7. A less flexible model excludes interactions between cutoff and running variable. Donut specification excludes the applicants defining the cutoffs (i.e., zeros). The nearest threshold uses only the two nearest thresholds applied by an applicant. In Column 6, the required number of observations on each side of the cutoff is increased to four. In Column 7, the cutoff-specific fixed effects and interactions were excluded. The final column uses a fixed bandwidth for all the outcomes, calculated as a mean of the optimal bandwidths used in the main specifications (Column 1). The optimal bandwidths were selected following Calonico et al. (2014), as shown in Equation 3, and standard errors were clustered at the applicant level (in parentheses).

3.5 Discussion and conclusions

In this paper, institutional quality and early career outcomes were studied in a business school setting using the regression discontinuity method. Concentrating on the fields of business and economics allowed for controlling for the field of study when investigating institutional effects. The business schools in Finland provided a straightforward and credible quasi-experimental research setting, in which the effect of attending a more selective institution over a less selective one could be studied instead of concentrating only on elite institutions. The results of the present study are informative for future applicants as well as for those designing admission systems or allocation of the available slots.

The results show that admission to a better business school decreases unemployment days experienced during the early career stages and increases the probability of completing an MSc. The income estimates suggest an increase at the cutoff, but more research must be done to confirm these effects, as causal conclusions cannot be confirmed with the present data.

The results are in line with the literature suggesting insignificant (Öckert, 2010; Kuuppelomäki et al., 2019) or modest (Kirkeboen et al., 2016) institutional effects on income, contrary to the literature suggesting high elite institution premiums (e.g., Anelli, 2020; Hoekstra, 2009). The results of the present paper further support the finding that admission to a more selective institution decreases unemployment. This finding is in line with the results of Saavedra (2009), who showed that in Colombia, crossing the top-ranked institution's admission cutoff increases the likelihood of being employed one year after college. On the other hand, Kuuppelomäki et al. (2019) found no employment effects for engineers in a Finnish setting. Finally, the present study results suggest that admission and enrollment pay off in terms of a higher probability of MSc completion. Similar results have been reported, for example, by Anelli (2020).

The more selective business schools in Finland are located in bigger cities than the less selective ones. Therefore, the results of the present study could reflect regional differences in, for example, employment opportunities in different business school locations. An avenue for future research would be to study the mechanisms behind the institutional quality effects. The study could further be replicated under different admission systems or by concentrating on fields of study other than business and economics. It would also be valuable to use measures of quality other than selectivity and expand the discussion to include other labor market and educational outcomes.

References

Abdulkadiroglu, A., Angrist, J., & Pathak, P. A. (2014). The elite illusion: Achievement effects at Boston and New York exam schools. *Econometrica*, *82*(1), 137–196. https://doi.org/10.3982/ecta10266

Anelli, M. (2020). The returns to elite university education: A quasi-experimental analysis. *Journal of the European Economic Association*, *18*(6), 2824–2868. https://doi.org/10.1093/jeea/jvz070

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, *14*(4), 909–946. https://doi.org/10.1177/1536867x1401400413

Dale, S., & Krueger, A. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics, 117*(4), 1491–1527. https://doi.org/10.1162/003355302320935089

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, *91*(4), 717–724. https://doi.org/10.1162/rest.91.4.717

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). Field of study, earnings, and self-selection. *Quarterly Journal of Economics*, *131*(3), 1057–1112. https://doi.org/10.1093/qje/qjw019.Advance

Kuuppelomäki, T., Kortelainen, M., Suhonen, T., & Virtanen, H. (2019). Does admission to elite engineering school make a difference? *VATT Working Papers, 146*. https://www.doria.fi/handle/10024/172697

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005

Ministry of Education and Culture. (2019). Tilannekuvaa korkeakoulutuksesta ja tutkimuksesta [Situation picture on higher education and research]. *Opetus- ja kulttuuriministeriö, korkeakoulu- ja tiedepolitiikan osasto, 34*. http://urn.fi/URN:ISBN:978-952-263-662-1

Official Statistics of Finland. (2018). *Transition from school to further education and work*. https://www.stat.fi/til/sijk/2018/sijk_2018_2020-01-23_tie_001_en.html

Official Statistics of Finland. (2021a). *Employment of students*. https://www.stat.fi/til/opty/2019/opty_2019_2021-03-12_tie_001_en.html

Official Statistics of Finland. (2021b). *Wages, salaries and labour costs*. https://www.tilastokeskus.fi/tup/suoluk/suoluk_palkat_en.html

Saavedra, J. E. (2009). The learning and early labor market effects of college quality: A regression discontinuity analysis [Unpublished manuscript]. https://economia.uniandes.edu.co/sites/default/files/seminariocede/The-Returns-to-College-Quality.pdf

Suhonen, T., & Jokinen, J. (2018). Mikä on tutkintotodistuksesi tuotto? [What is the return of your diploma?]. *Talous Ja Yhteiskunta, 2*, 30–37. https://vatt.fi/documents/2956369/7171534/Mik%C3%A4+on+tutkintotodistuksesi+tuotto/f43a98e1-6cd9-4764-89c5-aaed821ca30d/Mik%C3%A4+on+tutkintotodistuksesi+tuotto.pdf

Zimmerman, S. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review, 109*(1), 1–47. https://doi.org/10.1257/aer.20171019

Öckert, B. (2010). What's the value of an acceptance letter? Using admissions data to estimate the return to college. *Economics of Education Review, 29*(4), 504-516. https://doi.org/10.1016/j.econedurev.2009.12.003

Appendix

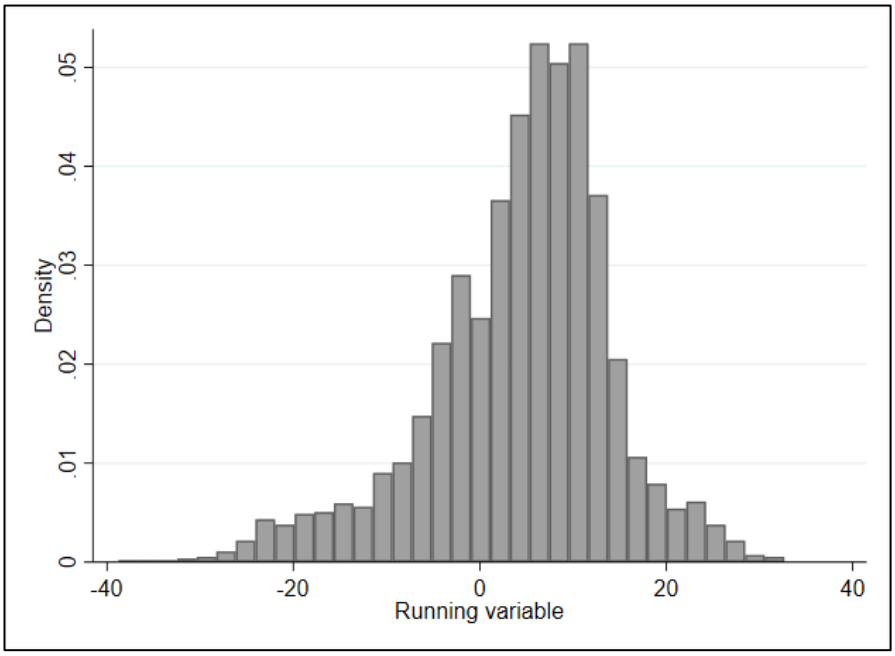

FIGURE A1 Density of the observations by the running variable. The graph includes the applicants that define the cutoffs (i.e., zeros).

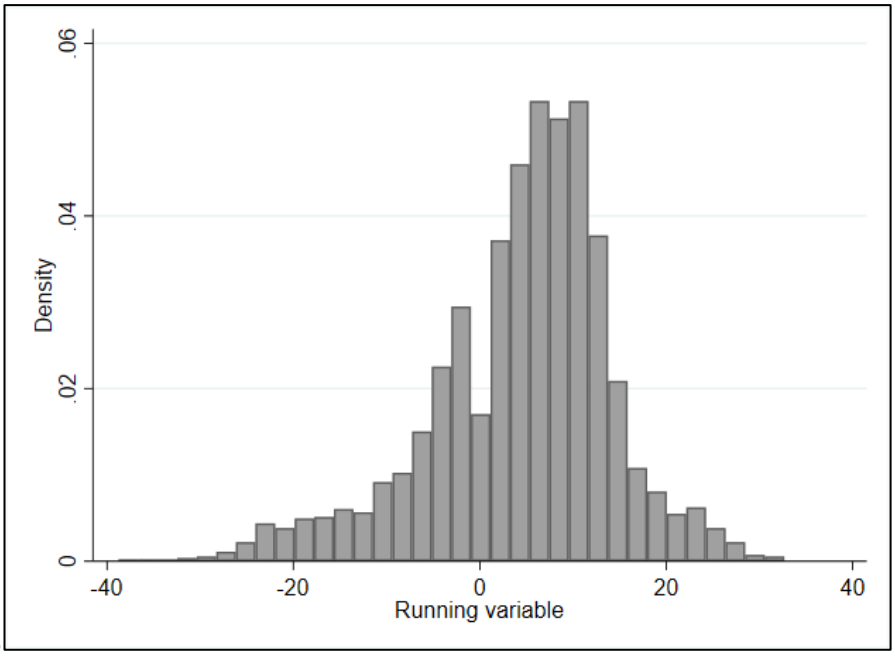

FIGURE A2 Density of the observations by the running variable. The graph excludes the applicants that define the cutoffs (i.e., zeros).

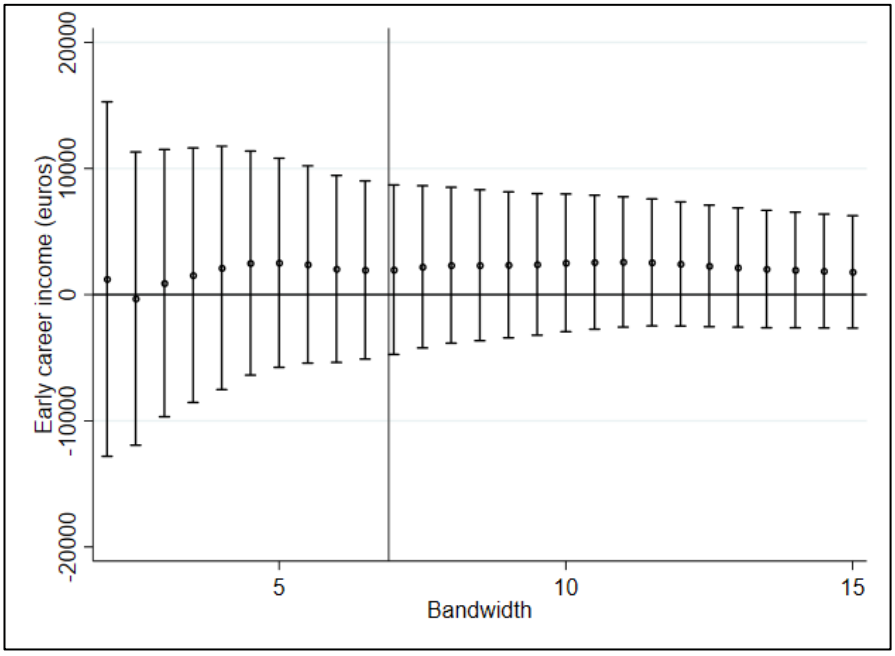

FIGURE A3 Early career income estimates and bandwidth choice. The figure shows reduced form estimates for early career income estimated across bandwidths from 2–15 at intervals of 0.5. The graph further shows the 95% confidence intervals (CIs) for the point estimates. The optimal bandwidth estimated following Calonico et al. (2014) is marked as a vertical line.

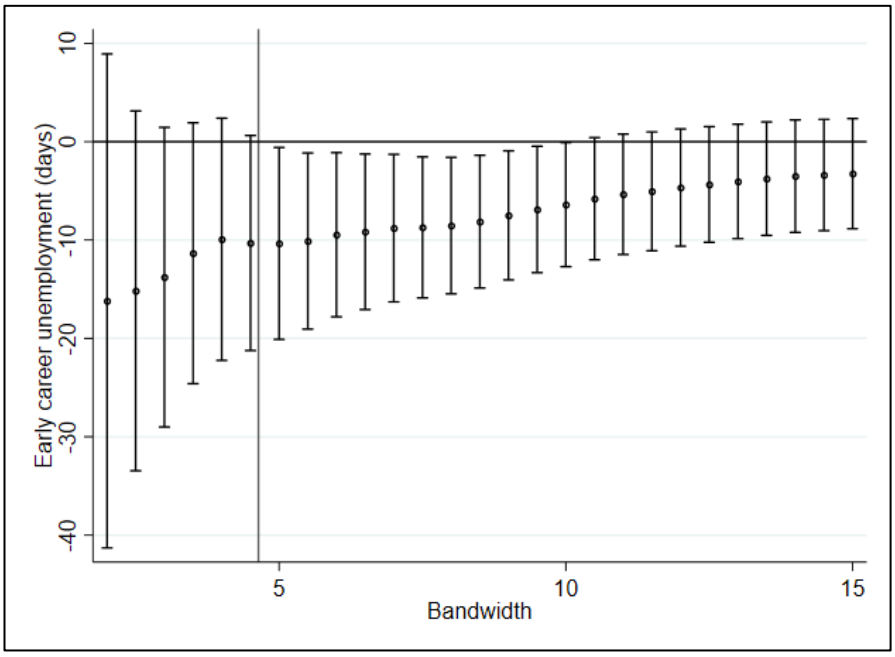

FIGURE A4 Early career unemployment estimates and bandwidth choice. The figure shows reduced form estimates for early career unemployment estimated across bandwidths from 2–15 at intervals of 0.5. The graph further shows the 95% confidence intervals (Cis) for the point estimates. The optimal bandwidth estimated following Calonico et al. (2014) is marked as a vertical line.

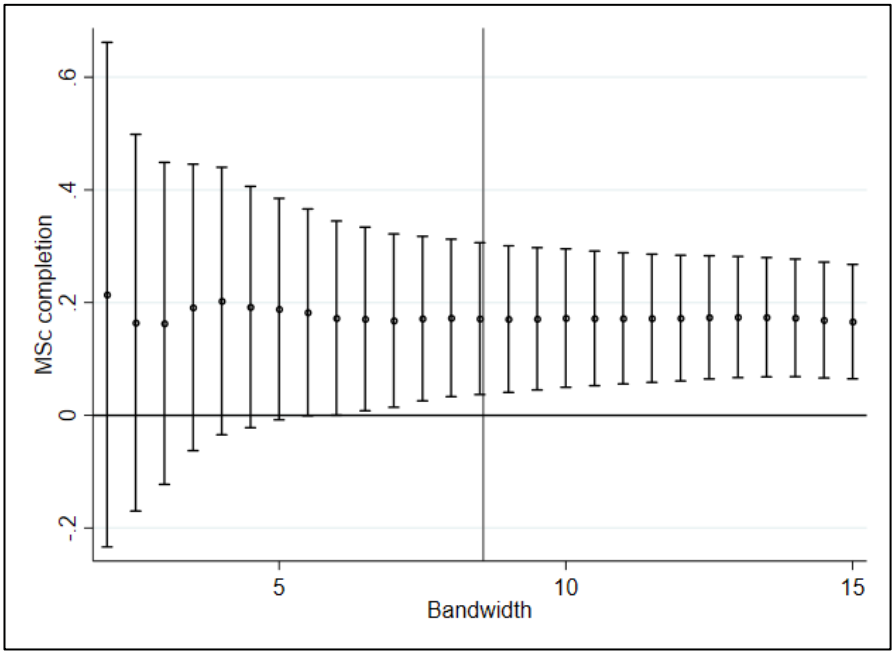

FIGURE A5 Master of Economic Sciences degree (MSc) completion estimates and bandwidth choice. The figure shows reduced form estimates for MSc completion estimated across bandwidths from 2–15 at intervals of 0.5. The graph further shows the 95% confidence intervals (Cis) for the point estimates. The optimal bandwidth estimated following Calonico et al. (2014) is marked as a vertical line.

4 INCOME AND MENTAL WELL-BEING: PERSONALITY TRAITS AS MODERATORS*

Abstract**

Using data from the participants of the Jyväskylä Longitudinal Study of Personality and Social Development (JYLS) at ages 42 and 50 (N = 326), this study provides empirical evidence of the relation between income and mental well-being and of the possible role of personality traits in modifying this relation. The relationships were analyzed using pooled ordinary least squares (OLS; bi- and multivariate settings) and fixed effects estimations (FE; multivariate settings). Positive bivariate associations were found between gross monthly income and the sum score of mental well-being and its separate dimensions (emotional, psychological, and social well-being and the absence of depression) as well as between experienced household finances and the sum score of mental well-being and its separate dimensions (except for social well-being). The multivariate OLS analyses detected positive relationships between gross monthly income and the absence of depression and between experienced household finances and mental well-being, along with one of its dimensions, i.e., emotional well-being. Further, the marginal utility of income appeared to depend on personality traits (FE): agreeableness and extraversion negatively moderated the gross monthly income–emotional well-being relationship, while openness positively moderated this relationship. In addition to emotional well-being, extraversion negatively moderated the relationship between gross monthly income and general mental

* This chapter has been published as "Syrén, S. M., Kokko, K., Pulkkinen, L., & Pehkonen, J. (2020). Income and Mental Well-Being: Personality Traits as Moderators. *Journal of Happiness Studies*, 21(2), 547-571. The numbering of the sections and the citation method were adjusted for the techniques and styles used in the thesis.

** The most recent data collection in the Jyväskylä Longitudinal Study of Personality and Social Development (JYLS) in 2009 was funded through Academy of Finland: award number 118316 was granted to Kokko and 127125 to Pulkkinen. Syrén also gratefully acknowledges financial support from the OP Group Research Foundation (project numbers 201500090, 201600189 and 20170093).

well-being, and neuroticism negatively moderated the association between gross monthly income and social well-being.

4.1 Introduction

The relationship between income and well-being has been extensively analyzed over the last four decades. Empirical research has focused on numerous developed and developing countries over various time periods. For example, Stevenson and Wolfers (2008) found a positive correlation between gross domestic product (GDP) and both happiness and life satisfaction in developed and developing countries. Regarding different time periods, a strong positive short-term relationship has been reported between GDP and happiness (Deaton, 2008; Stevenson & Wolfers, 2008), whereas Easterlin et al. (2010) suggested that there was no long-run relationship between GDP growth and life satisfaction.

The existing economics literature has paid limited attention to defining well-being, typically describing this in terms of happiness or life satisfaction, with these concepts being treated synonymously (e.g., Frey, 2008; Veenhoven, 1991) or as different notions of well-being (e.g., Deaton, 2008; Inglehart, 2008). In the psychological literature, happiness and life satisfaction constitute components of emotional well-being, also called subjective well-being, together with high positive and low negative affectivity (Diener, 1984). Psychological research has also examined psychological well-being, manifested as one's attempt at self-actualization and personal growth (Ryff, 1989), and social well-being, indicative of one's resolution in social tasks and encounters (Keyes, 1998). Emotional, psychological (Ryff, 1989), and social well-being (Keyes, 1998) are part of Keyes' (2002, 2005) tripartite model of well-being, which has been empirically corroborated (Gallagher et al., 2009; Keyes, 2005; Kokko et al., 2013b; Robitschek & Keyes, 2009). Further, using the same data from the Jyväskylä Longitudinal Study of Personality and Social Development (JYLS) as the present study, these well-being indicators have been shown to correlate negatively with ill-being, such as depression, in mid-adulthood (Kokko et al., 2013b). Emotional, psychological, and social well-being, together with the absence of ill-being, capture a latent factor of well-being (Kokko et al., 2013b), referred to as mental well-being (Kokko et al., 2015). Our aim in the present study was to shed further light on the relationship between income and both mental well-being and its various dimensions. Studying the different dimensions of mental well-being is important, because different aspects of well-being differ in what influences them and what they influence (Diener et al., 2017). It is further important for learning about and understanding the income–mental well-being relationship and for designing policies related to, for example, income taxation and welfare benefits.

Based on the JYLS and other empirical research, personality is closely related to well-being (e.g., Diener & Lucas, 1999; Steel et al., 2008; Kokko et al., 2013a). Personality traits are further associated with income (e.g., Mueller & Plug, 2006; Viinikainen et al., 2010) and with individuals' economic and financial

decision-making (Brown & Taylor, 2014). Moreover, personality has received increasing attention in research regarding the associations between income and well-being. Theoretically, as individuals have heterogeneous preferences (Sen, 1973), particular preference types may extract greater utility from a given income increase (Boyce & Wood, 2011). Soto and Luhmann (2012) further argue that effects of life circumstances on subjective well-being may vary depending on personality traits. Therefore, the marginal utility from an income increase on well-being may depend on personality traits. Empirically, Boyce and Wood (2011) showed that an increase in monthly income induced higher levels of life satisfaction for individuals with higher conscientiousness scores than for those with lower scores. Further, Proto and Rustichini (2015) as well as Soto and Luhmann (2012) reported that an increase in income was associated with higher levels of life satisfaction for individuals with high neuroticism scores than for those with low scores. Nevertheless, the empirical literature lacks evidence on whether personality moderates the link – i.e., affects well-being reactions arising from changes in income – between income and mental well-being and its dimensions.[1] In the analysis, the existing economic literature has assumed complete stability of personality traits (Boyce & Wood, 2011; Soto & Luhmann, 2012; Proto & Rustichini, 2015) due to convenience or unavailability of personality measures. We add to the literature by allowing individual variability in the Big Five personality traits.

Our analyses are based on the Finnish age-cohort group drawn from the JYLS (Pulkkinen, 2017).[2] The study aimed to contribute to the literature by exploring: 1) the associations between income and mental well-being, including its dimensions (emotional, psychological, and social well-being and the absence of depression; Keyes, 2002, 2005; Kokko et al., 2013b) and 2) the possible role of the Big Five personality traits (agreeableness, extraversion, conscientiousness, openness to new experiences, and neuroticism; Costa & McCrae, 1985) in modifying this relation. To answer these questions, we quantified income by using measures of individual income (gross monthly income) and the experienced financial situation of the household (household finances). We employed a longitudinal approach, which allowed us to observe the same individuals and changes in their income, mental well-being, and personality traits between the ages of 42 and 50. This approach also allowed us to control for unobserved time-invariant characteristics and to increase the efficiency of the estimations.

[1] The moderating role of the Big Five personality traits is studied using fixed effects estimation which utilizes within-individual variation. For example, the estimate of income is identified from changes in within-individual income. In this paper we refer to this within-individual variation with "changes in income".

[2] The JYLS data has been extensively utilized, but the present focus on the combined effects of personality and income on mental well-being has not previously been examined.

4.1.1 The concept of mental well-being and its relation to income

Mental well-being is comprised of emotional, psychological, and social well-being, together with the absence of ill-being (Keyes, 2005; Kokko et al., 2015). Emotional well-being describes how and why individuals experience their lives in a positive way (Diener, 1984). Individual income and the household's financial situation have been positively linked to emotional well-being, both theoretically (e.g., Sen, 1999) and empirically (e.g., Angeles, 2011; Boyce & Wood, 2011; Brown & Gray, 2016; Headey & Wooden, 2004). The components of emotional well-being usually include happiness, life satisfaction, and positive and negative moods (e.g., Diener, 1984; Russell & Carroll, 1999).

Psychological well-being emphasizes personal growth and living out one's possibilities (Keyes, 2006; Ryff, 1989; Waterman, 1993) and consists of six dimensions: self-acceptance, positive relationships with others, environmental mastery, autonomy, purpose in life, and personal growth (Ryff, 1989). The economics literature has shown that debt, particularly unsecured debt, is negatively related to psychological well-being (Brown et al., 2005). Income is further associated with a sense of mastery and control (Lachman & Weaver, 1998), and it is positively related to emotional (Kokko et al., 2013a; Ryff & Keyes, 1995) and social (Keyes & Waterman, 2003; Kokko et al., 2013a) well-being.

Social well-being relates to the surrounding society and describes how an individual evaluates his/her relationship with other people, residential area, and community (Keyes, 1998). The components of social well-being are social acceptance, social coherence, social integration, social contribution, and social actualization. Social well-being is further related to the relative income hypothesis, which postulates that higher levels of happiness require higher levels of income relative to a reference group (Clark et al., 2008). The relevant reference group could consist of a circle of acquaintances, neighbors, or the whole world – through globalization (Clark et al., 2008; Deaton, 2008). The existing literature has substantiated the relative income hypothesis by using both individual (Ferrer-i-Carbonell, 2005) and household income (Brown & Gray, 2016; Luttmer, 2005).

We measured the absence of ill-being as low depression. According to Zimmerman and Katon (2005), high income may reduce financial distress and provide greater levels of resources for treating depression. Empirical studies have shown that financial strain is associated with symptoms of depression (Zimmerman & Katon, 2005), that sudden loss of wealth is associated with feelings of depression (McInerney et al., 2013), and that debt is positively related to anxiety (Drentea, 2000).

4.1.2 Description and time variability of the Big Five personality traits

The Big Five personality traits include agreeableness, extraversion, conscientiousness, neuroticism, and openness to new experiences. A highly agreeable individual is trustful, straightforward, altruistic, compliable, modest, and tender-minded. Regarding extraversion, a high score is characterized by warmness, gre-

gariousness, assertiveness, activeness, and excitement seeking. Highly conscientious individuals can be characterized as competent, dutiful, achievement striven, self-disciplined, and deliberate. Conversely, high neuroticism is related, for example, to anxiety, hostility, and vulnerability to stress. Finally, openness to new experiences relates to fantasy, aesthetics, feelings, actions, ideas, and values (Costa & McCrae, 1989).

Although psychological studies have shown the absolute (mean level) and relative (correlative) stability of personality traits at the population level over time (e.g., Kokko et al., 2015), there are absolute and relative changes in these traits at the individual level. The JYLS (Pulkkinen, 2017, pp. 90, 95) showed that the mean of neuroticism decreased and that of agreeableness increased until age 42. From age 42 to 50, the means remained on the same level. There was, however, individual variation within the mean scores. For instance, agreeableness did not increase over time in some individuals or in all sub-groups, although it generally increased. The correlations between ages 42 and 50 varied from 0.70 to 0.80 for all the traits. A stability coefficient of 0.80 indicates that only 64% of the variance across the two time points was shared and that the rest was explained by true individual variability and measurement error. Since the personality tests measuring the Big Five personality traits are generally highly reliable, part of the variance was explained by true individual variability.

In addition to the present JYLS data, the empirical literature supports the existence of both individual- and mean-level changes throughout the life span. Roberts and Mroczek (2008) illustrated that individuals have unique patterns of development, which are affected by life experiences. In a meta-analysis conducted using longitudinal studies, Roberts et al. (2006b) also showed statistically significant mean-level changes in Big Five personality traits in middle (40–60) and old (>60) age. Personality changes can result from environmental changes in social roles or cultural milieu (Helson et al., 2002a; Scollon and Diener 2006) or from life and work experiences (e.g., Roberts et al., 2003; Roberts et al., 2006a; Mroczek & Spiro, 2003; Elkins et al., 2016; Anger et al., 2017; Golsteyn & Schildberg-Hörisch, 2017).

The question of whether personality changes represent temporary fluctuations or measurement error has also been addressed in the empirical literature through the use of the Reliable Change Index (Roberts et al., 2001) and growth models (e.g., Helson et al., 2002b; Mroczek & Spiro, 2003; Small et al., 2003). This literature has established that variability across individuals, both in the direction and rate of personality change, can be demonstrated by the Big Five personality traits (Roberts & Mroczek, 2008). Recent literature has further implied that economic models ignoring the personality change may be incorrectly specified (Boyce et al., 2013). All things considered, we believe that it is reasonable to treat personality traits as time-variant and to make statistical inferences based on personality changes. In this study, we examined what happens to the marginal utility of income on mental well-being and its dimensions when within-individual personality-trait changes are taken into account.

4.2 Method

4.2.1 Data collection and the population of the study

The JYLS began in 1968, since which six data collection phases have been conducted (Pulkkinen, 2017). The initial sample consisted of 12 randomly selected second-grade school classes in the town of Jyväskylä, Finland. These classes comprised 369 eight-year-old pupils (173 girls and 196 boys), with an initial participation rate of 100%. The participants were mailed a Life Situation Questionnaire (LSQ), inviting them to participate in a semi-structured psychological interview, with self-report inventories and medical examinations using laboratory tests. More information about the data collection can be found in Pulkkinen (2017) reference.

4.2.2 Present sample and the representativeness

We utilized JYLS data collected at ages 42 and 50 (in 2001 and 2009, respectively). By ages 42 and 50, six and twelve participants had died, yielding available sample sizes of 363 and 357, respectively. At age 42, 77% of the available sample returned the LSQ, and 71% participated in the interview. At age 50, the LSQ was returned by 76% of the available sample, and 64% took part in the interview. At ages 42 and 50, 163 participants had no missing data regarding any of the studied variables. In the present analysis, we pooled information about these participants, yielding a total of 326 observations.

At ages 42 and 50, the participants were representative of the initial sample in terms of socioemotional behavior at age eight and school achievement at age 14 (Pulkkinen, 2017). Furthermore, compared against the statistics provided by Statistics Finland, these participants represented the Finnish age-cohort group born in 1959 with respect to marital status, number of children, and employment. To examine attrition, we compared, at age eight, the present sample (N = 163) with those who were excluded due to missing data (N = 206). Regarding socioemotional behavior in childhood (Pulkkinen et al., 2012), the t-tests for independent groups revealed no statistically significant differences between the groups in terms of behavioral activity (p = 0.61). However, the excluded participants scored higher on negative emotionality (p = 0.028) and lower on well-controlled behavior (p = 0.046) than those who were included. No between-group differences were observed in school success (p = 0.28) or parental occupational status (p = 0.91). We concluded that the present sample represents the initial sample reasonably well.

4.2.3 Measures and variables

In the LSQ, the participants were asked to rate their gross monthly income (including all taxable income, pensions, and social benefits, but not capital income) using a pre-refined response scale (Pulkkinen & Kokko, 2010). At age 42, the scale

ranged from 1 to 12 (FIM4,000 or less to over FIM32,000) and at age 50 from 1 to 14 (€1,000 or less to over €7,000). For the statistical analyses, we utilized the averages of the income classes, converted Finnish marks into euros, and adjusted for inflation using the Consumer Price Index (Official Statistics of Finland). For the top-coded groups, we utilized the lower bound of the income classes. At age 50, the annual income information from the tax authority registers was available for 158 of our sample of 163. For these 158 participants, the Spearman correlation between the self-evaluated gross monthly income and the register-based annual income was 0.87, with a pairwise correlation of 0.83 ($p < 0.01$ in both cases). Such high correlations suggest that the participants accurately evaluated their gross monthly income. In addition to individual income, we evaluated experienced household finances, which implicitly account for factors that may tighten a participant's financial situation, such as liabilities. The participants evaluated their experienced household finances at ages 42 and 50 based on the following question presented in the LSQ: "How do you consider your current personal financial situation or that of the family you have set up?" The scale included 1 = extremely tight, 2 = fairly tight, 3 = fairly good, and 4 = extremely good financial situation.

Emotional well-being was measured at ages 42 and 50 using the sum of the standardized scores[3] of four subcomponents: happiness, life satisfaction, positive mood, and reversed negative mood. Happiness was assessed with the question: "How happy or satisfied have you been at the different stages in your life (Perho & Korhonen, 1993)?" The response scale ranged from -3 to +3 (very unhappy or dissatisfied to very happy or satisfied). At age 42, the most recent time point referred to ages 40–42 years, whereas at age 50, the participants were asked to estimate their current happiness and satisfaction. General life satisfaction was based on seven life domains (for which an average score was calculated): housing, financial situation, choice of occupation, present occupational situation, present intimate relationship or lack thereof, content of leisure time, and present state of friendships (Kokko et al., 2013b), with the response scale ranging from 1 to 4 (very dissatisfied to very satisfied). Positive and negative moods were measured using the Brief Mood Introspection Scale (Feldman, 1995; Mayer et al., 1988). Positive mood was calculated as an average score of two items (e.g., "My present mood is satisfied") and negative mood as an average score of five items (e.g., "My present mood is frightened"; Kokkonen, 2001; Kokko et al., 2013b). The response scale ranged from 1 to 4 (does not describe my mood at all to describes my mood very well).

Psychological well-being was based on the Scales of Psychological Well-Being (Ryff, 1989) at ages 42 and 50, which consisted of 18 items (e.g., "In general, I feel I am in charge of the situation in which I live"). Social well-being was measured using the Scales of Social Well-Being (Keyes, 1998) at the same ages, which consisted of 15 items (e.g., "I have something valuable to give the world").

[3] The standardization of the scores for general well-being and its dimensions and the personality traits was conducted in Stata as follows: for each observation, the mean of the variable was subtracted from the value of the observation. This difference was then divided by the standard deviation of the variable.

For both psychological and social well-being, the response scale ranged from 1 to 4 (strongly disagree to strongly agree). Depression was assessed at ages 42 and 50 using the Depression Scale of General Behavior Inventory (Depue, 1987), which consisted of 16 items (e.g., "Have you become sad, depressed or irritable for several days or more without really understanding why?"). The response scale varied from 1 to 4 (never to very often). Average scores were calculated, and Cronbach's alpha reliability was higher than 0.63 for all the above cases (Kokko et al., 2015). Finally, we constructed mental well-being at ages 42 and 50 by summing the standardized scores for emotional, psychological, and social well-being and reversed depression.

Personality traits were measured at age 33 using the Big Five Personality Inventory (Pulver et al., 1995). This is an authorized adaptation of the NEO Personality Inventory (NEO-PI), which contains 180 items (Costa & McCrae, 1985), about one-fourth of which are substitutes for the original American items (Rantanen et al., 2007). A shortened version was formed in order to correspond with the shortened 60-item NEO-Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1989). In the Finnish NEO-FFI, which was administered to the participants at ages 42 and 50 (Pulkkinen, 2017), three items served as substitutes for the original items. All five subscales – neuroticism (e.g., "I often feel tense and jittery"), extraversion (e.g., "I like to have a lot of people around me"), conscientiousness (e.g., "I'm pretty good about pacing myself so as to get things done on time"), openness (e.g., "I have a lot of intellectual curiosity"), and agreeableness (e.g., "I would rather cooperate with others than compete with them") – were assessed using 12 items, and the mean score of the items was calculated for each trait. The response scale ranged from 1 to 5 (strongly disagree to completely agree). Cronbach's alpha reliability was above 0.75 for each personality trait (Kokko et al., 2015).

To alleviate possible omitted variable bias, we controlled factors that could be linked to income and mental well-being. Categories of the control variables were included in the estimations as dummies (reference categories described in parentheses). The predetermined control variables included gender (0 = male) and the follow-up year 2009 (0 = 2001). Other controls were relationship status, state of health, and size of household. The categories for relationship status were married, lives in cohabitation without marriage, single, and divorced or widowed (0 = married). State of health was evaluated based on the following question: "During the past year, how would you describe your health as a whole?" The categories used were very good, fairly good, moderate, fairly bad, and very bad health (0 = fairly good health). Lastly, we controlled for the size of the household that the participants reported as part of the LSQ.

In the appendix, we show that our results are robust for an additional set of labor market and education controls. Specifically, we added controls for employment situation, stability of career line, occupational status, and education. Approximately 90% of workers in Finland are trade unions members. Therefore, job loss does not result in a dramatic fall in income; instead, income declines gradually with the duration of unemployment. However, losing a job might

contribute to mental well-being, despite income stability, as one's daily work routine and colleagues are lost. Similarly, the stability of the career line and occupational status are likely to have separate associations with stress level and mental well-being. Therefore, it is important to show the results with these controls. Employment status was categorized as unemployed, part-time, or full-time employee (0 = full-time employee). The stability of the career line was based on information collected about work history and was categorized as unstable, changeable, or stable working career (0 = stable working career). Occupational status was based on a question related to one's latest professional title and was classified into upper white-collar, lower white-collar, and blue-collar occupation (0 = lower white-collar), and education was categorized as course, vocational school, vocational college, and university degree (0 = vocational school). Further information about the control variables can be found in Pulkkinen and Kokko (2010).

4.2.4 Data analysis

The statistical analyses were conducted using Stata/SE 14, employing pooled ordinary least squares (OLS; Equation 1) and fixed effects (FE; Equations 2 and 3) estimations. The baseline specifications were:

$$mwe_{it} = \alpha + \beta \log i_{it} + \lambda_\pi P_{it} + \gamma_k X_{it} + \varepsilon_{it} \quad (1)$$
$$mwe_{it} = \alpha + \beta \log i_{it} + \lambda_\pi P_{it} + \gamma_k X_{it} + \mu_i + \varepsilon_{it} \quad (2)$$
$$mwe_{it} = \alpha + \beta \log i_{it} + \lambda_\pi P_{it} + \gamma_k X_{it} + \delta_\rho P_{it} * \log i_{it} + \mu_i + \varepsilon_{it}, \quad (3)$$

where an individual's mental well-being (mwe_{it}) was regressed on income (i_{it}), personality traits (P_{it}), and observable covariates (X_{it}). In models 2 and 3, unobserved time-invariant effects (μ_i) were removed, and model 3 further examined whether personality traits moderated the associations between income and mental well-being, with $P_{it} * \log i_{it}$ capturing the interaction of each personality trait with income.

OLS and FE estimations utilize different types of variation: cross-sectional variation and within-individual variation, respectively. As we pooled the data, i.e., combined the observations from ages 42 and 50, there were two observations from each individual therein. OLS regression considers each observation separately, regardless of age. Standard errors are likely to be correlated over time at the individual level, which we corrected by clustering the standard errors at the individual level. Conversely, FE examines how the change in an individual's income associates with changes in his/her well-being between ages 42 and 50, taking the controls into account. The FE estimate gives the average of these individual-level changes.

4.3 Results

4.3.1 Data description

Table 1 presents the descriptive statistics for the key variables at ages 42 and 50. Mental well-being and its dimensions—emotional (and its sub-dimensions happiness and positive mood), psychological, and social well-being—gross monthly income, and agreeableness increased, whereas extraversion, neuroticism, and openness decreased significantly ($p < 0.05$) from age 42 to 50. Descriptive statistics for the control variables are presented in the appendix (Table 6). Table 2 presents the correlations in the pooled sample, i.e., including observations from ages 42 and 50. The dimensions of mental well-being (emotional, psychological, and social well-being and low depression) were strongly positively correlated (ranging from 0.28 to 0.85), as shown in previous JYLS analyses (Kokko et al., 2015). We found further moderately positive correlations between income and the well-being measures, with the coefficients ranging from 0.12 to 0.34. The correlations were more consistent between gross monthly income and the well-being measures (0.24–0.34) than between household finances and the well-being measures (0.12–0.24). The correlation coefficient between gross monthly income and household finances was 0.34, describing a moderate correlation.

TABLE 1 Means and standard deviations of mental well-being and its dimensions, income, and personality traits (N = 163); p-values from paired t-tests.

	Age 42 (2001)		Age 50 (2009)		
	Mean	*SD*	*Mean*	*SD*	*p-value*
Well-being					
Mental well-being (S)	-0.14	1.00	0.14	0.98	.000
Emotional well-being (S)	-0.09	1.05	0.09	0.94	.047
Happiness	1.75	1.16	2.04	0.69	.002
Life satisfaction	3.13	0.35	3.13	0.36	.998
Positive mood	2.88	0.64	3.00	0.56	.038
Negative mood (R)	3.84	0.26	3.84	0.26	.920
Psychological well-being	3.17	0.34	3.22	0.31	.015
Social well-being	2.85	0.36	2.99	0.39	.000
Depression (R)	3.54	0.35	3.59	0.37	.057
Income					
Gross monthly income (thousands)	2.23	1.06	2.84	1.34	.000
Household finances	2.77	0.70	2.87	0.65	.077
Personality traits					
Agreeableness	3.65	0.52	3.73	0.44	.005
Conscientiousness	3.70	0.53	3.68	0.50	.484
Extraversion	3.31	0.55	3.23	0.55	.006
Neuroticism	2.34	0.64	2.24	0.59	.005
Openness to new experiences	3.37	0.58	3.27	0.54	.001

Notes: S = standardized score; R = reversed score; mental and emotional well-being are sum variables constructed using standardized scores.

TABLE 2 Correlations between mental well-being, its dimensions, and income variables – pooled sample (N = 326).

	1	2	3	4	5	6	7
1. Mental well-being	-						
2. Emotional well-being	.78**	-					
3. Psychological well-being	.85**	.60**	-				
4. Social well-being	.72**	.39**	.54**	-			
5. Depression (R)	.70**	.41**	.44**	.28**	-		
6. Gross monthly income	.34**	.24**	.25**	.28**	.30**	-	
7. Household finances	.24**	.24**	.19**	.12*	.19**	.34**	

Notes: R = reversed score; ** $p < 0.01$; * $p < 0.05$.

4.3.2 Income and mental well-being

We first investigated the relationship between income and well-being using gross monthly income (Table 4, upper part). Three specifications (1, 2, and 3) were separately estimated for the standardized scores for mental well-being and its dimensions: emotional, psychological, and social well-being and reversed depression. The bivariate OLS model (specification 1) indicated positive associations between income and all the well-being measures. Including the standardized scores for the personality traits and control variables (specification 2) and accounting for unobserved heterogeneity (specification 3) yielded a statistically significant coefficient for income in one case: reversed depression.

We replicated the analysis using household finances as an income variable (Table 4, lower part). In the OLS models, household finances were positively associated with mental well-being and with the dimension of emotional well-being, indicating more robust relations between income and well-being. The associations between household finances and other well-being measures were similar to those of our previous results: in the bivariate OLS models, household finances were positively associated with psychological well-being and reversed depression.[4]

The appendix (Table 7) reports on the relationships between income and the well-being variables when an additional set of education and labor market controls are included in the models. The results are robust for the inclusion of the additional control variables. The only difference was that the estimate of gross monthly income was no longer statistically significant in the case of reversed depression. The standard errors are similar between Tables 4 and 7, confirming that our results do not suffer from multicollinearity. All the results were further robust for the exclusion of individuals with top-coded income values.[5]

[4] The bivariate FE model (not reported) indicated a positive association between household finances and emotional well-being (point estimate 0.294, standard deviation 0.136).
[5] The bivariate OLS associations between income and the well-being measures were positive and statistically significant, except for social well-being. In the multivariate OLS setting, gross monthly income was positively associated with reversed depression, and household finances were positively associated with mental well-being and, particularly, one of its dimensions, i.e., emotional well-being.

TABLE 4: Pooled OLS and FE regressions: Gross monthly income, household finances, and mental well-being and its different dimensions

	Mental well-being (S)			Emotional well-being (S)			Psychological well-being (S)			Social well-being (S)			Depression (S,R)		
	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)
Log gross monthly income	.516**	.044	.019	.339**	-.081	-.057	.363**	-.053	-.191	.307*	.027	.059	.564**	.240*	.247
	(.124)	(.078)	(.089)	(.122)	(.105)	(.154)	(.115)	(.084)	(.121)	(.132)	(.128)	(.110)	(.134)	(.093)	(.164)
R^2	.096	.658	.424	.041	.408	.297	.046	.562	.280	.061	.402	.312	.096	.469	.182
Household finances	.343**	.125*	-.011	.345**	.170*	.221	.268*	.063	-.116	.156	.056	-.085	.279**	.091	-.052
	(.103)	(.058)	(.073)	(.103)	(.079)	(.119)	(.104)	(.069)	(.104)	(.084)	(.069)	(.079)	(.099)	(.070)	(.085)
R^2	.072	.664	.424	.061	.418	.315	.040	.563	.279	.045	.403	.317	.039	.423	.166
Personality traits and controls	-	√	√	-	√	√	-	√	√	-	√	√	-	√	√
N	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326

Notes: $^{**}p < 0.01$; $^{*}p < 0.05$. S = standardized score; R = reversed score. Controls: gender, relationship status, state of health, household size, and a dummy for the second survey year. Specifications 1 and 2 were analyzed using OLS. Specification 1 controlled for the second survey year, and specification 2 added all other controls. Specification 3 was analyzed using FE and included controls other than gender, which was excluded because of the lack of within variation between the given time periods. Standard errors are in parentheses.

4.3.3 The moderating role of personality traits in the association between income and mental well-being

To explore the moderating role of personality traits in the income–mental well-being associations, we augmented our FE models with income–personality trait interaction terms (Table 5), and thus utilized only the within-individual variation. The income variable interacting with the standardized scores for the personality traits is gross monthly income in specification 1, followed by household finances in specification 2. Our analyses indicated, first, that extraversion negatively moderated the monthly gross income–mental well-being relationship (specification 1), implying that a higher score in extraversion is associated with a more negative income–mental well-being relationship. Second, when the dimensions of mental well-being were separately analyzed, agreeableness and extraversion negatively moderated the association between gross monthly income and emotional well-being, while openness positively moderated this association (specification 1). Contrary to the negative moderators, the result for openness suggests that the higher the score in this personality trait, the more positive the association between gross monthly income and emotional well-being. Finally, neuroticism negatively moderated the association between gross monthly income and social well-being (specification 1). The inclusion of the labor market controls (Appendix, Table 8) yielded coefficients that were more often statistically significant and consistent in magnitude when compared with those presented in Table 5. The results were further robust for the exclusion of the top-coded income values.[6]

[6] Extraversion negatively moderated the association between gross monthly income and mental well-being. Agreeableness and extraversion negatively moderated the relationship between gross monthly income and emotional well-being, while openness positively moderated this relationship. Further, extraversion negatively moderated the gross monthly income–psychological well-being association, and neuroticism negatively moderated the gross monthly income–social well-being relationship. Using household finances as an independent variable, neuroticism was found to be a negative moderator of reversed depression.

TABLE 5 FE regressions: Gross monthly income, household finances, and personality trait interaction effects on mental well-being and its different dimensions

	Mental well-being (S)		Emotional well-being (S)		Psychological well-being (S)		Social well-being (S)		Depression (S, R)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Income										
Log gross monthly income	0.009		0.023		-0.171		0.030		0.145	
	(0.106)		(0.167)		(0.131)		(0.121)		(0.172)	
Household finances		-0.034		0.210*		-0.170		-0.101		-0.045
		(0.069)		(0.105)		(0.087)		(0.083)		(0.090)
Personality trait interactions with income										
Agreeableness (S)	-0.214	-0.154	-0.496*	-0.214	-0.157	-0.151	-0.076	-0.040	0.075	-0.065
	(0.133)	(0.080)	(0.195)	(0.122)	(0.142)	(0.077)	(0.140)	(0.059)	(0.126)	(0.097)
Conscientiousness (S)	0.075	0.019	0.235	0.007	0.132	0.0503	-0.011	0.094	-0.127	-0.094
	(0.092)	(0.076)	(0.160)	(0.118)	(0.097)	(0.086)	(0.088)	(0.063)	(0.141)	(0.082)
Extraversion (S)	-0.275*	-0.162	-0.421*	-0.110	-0.261	-0.171	-0.160	-0.096	0.004	-0.118
	(0.113)	(0.099)	(0.180)	(0.148)	(0.142)	(0.102)	(0.131)	(0.081)	(0.164)	(0.112)
Neuroticism (S)	-0.171	-0.128	-0.103	-0.015	-0.190	-0.080	-0.254*	-0.111	0.024	-0.185
	(0.105)	(0.070)	(0.191)	(0.104)	(0.116)	(0.088)	(0.099)	(0.076)	(0.125)	(0.097)
Openness to new experiences (S)	0.184	0.000	0.273*	0.159	0.165	-0.122	0.078	-0.067	0.045	0.031
	(0.096)	(0.071)	(0.134)	(0.097)	(0.106)	(0.089)	(0.099)	(0.079)	(0.174)	(0.099)
Personality traits										
Agreeableness (S)	1.723	0.527*	3.760**	0.523	1.343	0.619**	0.732	0.290*	-0.580	0.177
	(1.050)	(0.273)	(1.526)	(0.394)	(1.113)	(0.259)	(1.061)	(0.168)	(1.000)	(0.307)
Conscientiousness (S)	-0.398	0.0953	-1.538	0.198	-0.923	-0.0949	0.0635	-0.295*	1.184	0.482*
	(0.718)	(0.217)	(1.241)	(0.339)	(0.745)	(0.240)	(0.662)	(0.177)	(1.098)	(0.265)
Extraversion (S)	2.285**	0.643*	3.326**	0.377	2.225**	0.713**	1.346	0.408	0.0731	0.462
	(0.907)	(0.333)	(1.422)	(0.485)	(1.121)	(0.348)	(1.013)	(0.249)	(1.289)	(0.369)

continues

TABLE 5 continues

	Mental well-being (S)		Emotional well-being (S)		Psychological well-being (S)		Social well-being (S)		Depression (S, R)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Neuroticism (S)	1.012	0.0466	0.418	-0.324	1.211	-0.0511	1.815**	0.180	-0.359	0.338
	(0.838)	(0.202)	(1.489)	(0.298)	(0.914)	(0.266)	(0.756)	(0.209)	(0.961)	(0.323)
Openness to new experiences (S)	-1.184	0.200	-1.527	0.119	-1.099	0.482**	-0.626	0.133	-0.359	-0.124
	(0.719)	(0.204)	(0.994)	(0.272)	(0.788)	(0.241)	(0.741)	(0.236)	(1.330)	(0.292)
Personality and Controls	√	√	√	√	√	√	√	√	√	√
R^2	0.452	0.466	0.355	0.343	0.304	0.343	0.334	0.344	0.196	0.197
N	326	326	326	326	326	326	326	326	326	326

Notes: $^{**}p < 0.01$; $^{*}p < 0.05$. S = standardized score; R = reversed score. Specification 1 was analyzed using gross monthly income as an independent variable and specification 2 with household finances as an independent variable. Both specifications were estimated using FE, and the controls (relationship status, state of health, household size, and a dummy for the second survey year) were included in all the regressions. Gender was excluded because of the lack of within variation between the given time periods. Standard errors are in parentheses.

Figure 1 provides a graphical illustration of the results, using emotional well-being as a dependent variable (Table 5, specification 1), as the statistically significant interaction effects were mainly found for emotional well-being. Specifically, Figure 1 graphs average emotional well-being for the different values of gross monthly income and the different values of income and agreeableness, extraversion, or openness, adjusted for the other covariates in our FE model. The predictive margins are presented with 95% confidence intervals for the means. Figure 1a shows the slightly positive relationship between gross monthly income and emotional well-being. In Figure 1b, the association between gross monthly income and emotional well-being is graphed for individuals with very high and very low agreeableness scores: a value of -2 describes agreeableness at two standard deviation units below the mean (marked as a hollow square), while a value of 2 describes agreeableness at two standard deviation units above the mean (marked as a circle). Figure 1b confirms the negative moderating role of agreeableness: at very high agreeableness scores (agreeableness = 2), the association between income and emotional well-being seems to be negative, whereas at very low agreeableness scores (agreeableness = -2), the association appears to be positive. While individuals with high agreeableness scores start with higher emotional well-being at low income levels, those with low agreeableness surpass them at higher income levels. The difference in emotional well-being seems largest at low income levels.

Figure 1c clarifies the negative moderating role of extraversion. Compared to the income–emotional well-being association shown in Figure 1a, individuals with very high extraversion scores (extraversion = 2) demonstrate a more negative association between income and emotional well-being, while those with very low extraversion scores (extraversion = -2) show a more positive association. Figure 1d turns to openness and its moderating role in the gross monthly income–emotional well-being association. The figure illustrates that individuals with high openness scores (openness = 2) seem to have a positive association between income and emotional well-being, whereas for individuals with low openness scores (openness = -2), the association is negative. Unlike for agreeableness and extraversion (Figures 1b and 1c), the differences in emotional well-being seem to be greatest at high income levels in the case of openness. For example, individuals with high and low openness scores have very similar emotional well-being when gross monthly income is low, but individuals with high openness scores seem to be better off when gross monthly income is high.

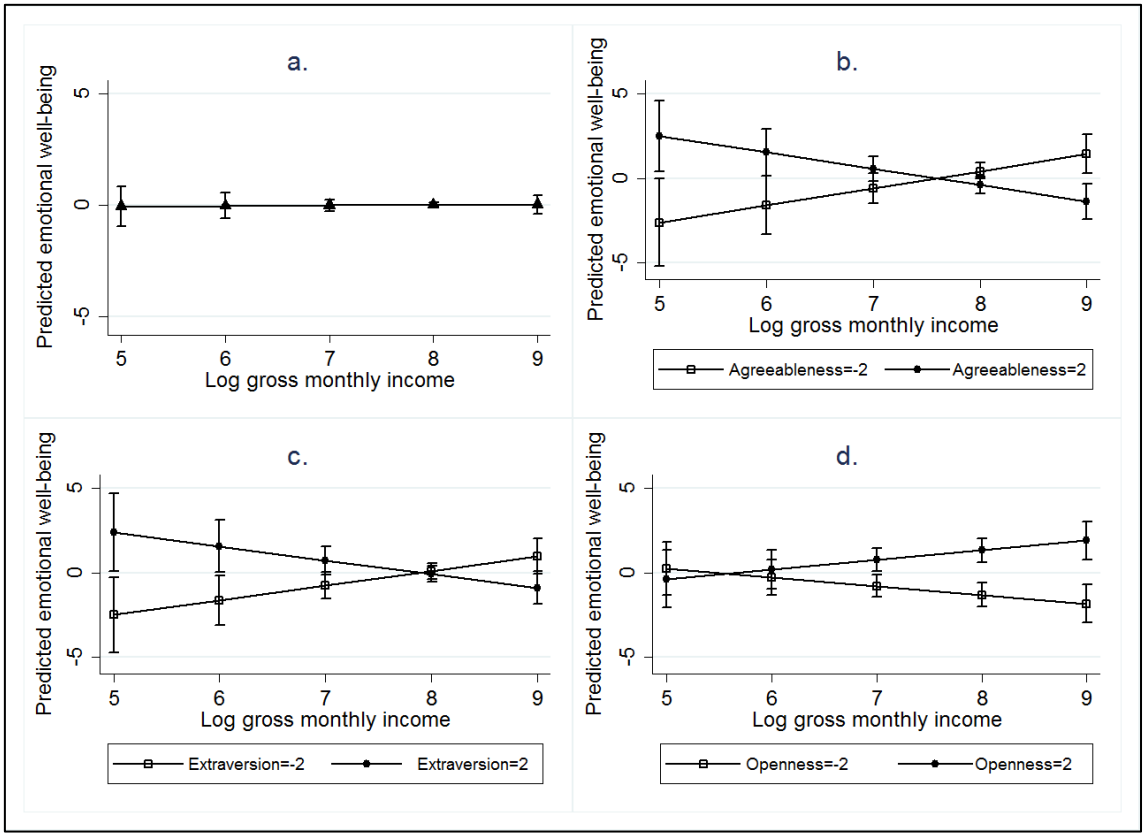

FIGURE 1 Predicted emotional well-being (S) by (a) gross monthly income, (b) gross monthly income and agreeableness (S), (c) gross monthly income and extraversion (S): and (d) gross monthly income and openness (S). A value of -2 describes a personality trait value of two standard deviation units below the mean, and a value of 2 describes two standard deviation units above the mean. S = standardized.

4.4 Discussion

The existing economics literature has shown a positive short-term relationship between GDP or income and the dimensions of emotional well-being (typically happiness or life satisfaction; see, e.g., Boyce & Wood, 2011; Deaton, 2008; Stevenson & Wolfers, 2008). We found income to be positively associated with the well-being measures in the bivariate OLS setting (except for the household finances–social well-being association). Further, the relationships between gross monthly income and reversed depression and between household finances and mental well-being and one of its dimensions (emotional well-being) were positive in the multivariate OLS setting. After the inclusion of the labor market and education controls, the estimate of income in the case of reversed depression was no longer statistically significant. This result suggests that households share assets and liabilities, and therefore, the financial situation of the household may be

more crucial than individual income for mental well-being. A household's financial situation may be further tightly related to emotional well-being, especially because emotional well-being is composed of happiness, life satisfaction, and affectivity. One component of life satisfaction is satisfaction with one's financial situation. Therefore, an individual evaluating his/her household financial situation as tight might also report dissatisfaction with his/her current financial situation. Further, the income variables used were moderately correlated (0.34) and describe different kinds of income. Gross monthly income illustrates the taxable income, pensions, and social benefits received by an individual, whereas household finances illustrate how an individual experiences his/her personal financial situation or that of the family that he/she has set up.

The majority of our income–mental well-being estimates, however, described insignificant associations between income and the mental well-being measures. This supports the existing economics literature, which has illustrated very small effect sizes (Angeles, 2011), and has shown that, in the long run, the positive relationship between GDP and life satisfaction vanishes (Easterlin et al., 2010). The limited role of income in mental well-being is also supported by the psychological literature, which has illustrated that factors such as personality traits are highly important contributors to mental well-being. According to a meta-analysis by Steel, Schmidt, and Shultz (2008), personality traits explain 40–60% of the variation in the mental well-being indices. Using the present JYLS data, it has been shown that the role of the personality traits is smallest for happiness (approximately 20%; Korkalainen, 2007) and smaller for life satisfaction than for psychological well-being (Kokko et al., 2013a). Interestingly, the economics literature has concentrated on happiness and life satisfaction.

Another possible explanation for these insignificant relationships is that a higher income level might be associated with financial resources being exceeded, such as high debt. Tay et al. (2017) showed that debt is linked to subjective well-being through satisfaction. This reasoning is further supported by the JYLS data: at low gross monthly income levels (from €1,000 to €2,200 per month), the household financial situation improved when income increased. Between €2,200 and €3,400 per month, the household financial situation remained stable, and from €4,600 per month onwards, household finances did not improve; rather, they weakened (Pulkkinen, 2017). It could also be that not only are financial resources exceeded when income increases, but higher income may relate to larger workloads, higher stress levels, or less free time, which could further translate into lower well-being.

The personality traits moderated – i.e., affected well-being reactions arising from changes in income – the relationship between income and well-being. Extraversion negatively moderated the relationships between gross monthly income and mental well-being. In addition to general mental well-being, the dimensions of emotional and social well-being were moderated by the personality traits. Agreeableness and extraversion negatively moderated the gross monthly income–emotional well-being association, while openness

positively moderated this association. The gross monthly income–social well-being relationship was moderated by neuroticism.

The negative moderating effect of agreeableness is intuitive, as highly agreeable individuals can be characterized as compliant and altruistic (Costa & McCrae, 1989). These individuals reported high scores for questions such as "I would rather cooperate with others than compete with them." Therefore, such individuals may not enjoy high income as much as less agreeable individuals. Compared to our results, the existing literature has reported insignificant and, in general, smaller coefficients regarding the moderating role of agreeableness on the income–life satisfaction association (Boyce & Wood, 2011; Soto & Luhmann, 2012; Proto & Rustichini, 2015).

The existing literature on the moderating role of extraversion and neuroticism has mixed results. For example, for life satisfaction, Boyce and Wood (2011) reported a positive interaction effect between income and extraversion for women, but found no significant effects for men. We speculated whether highly extraverted individuals on high incomes were satisfied with their financial situation and, therefore, gained nothing from income changes. For neuroticism, Proto and Rustichini (2015) as well as Soto and Luhmann (2012), found that higher levels of life satisfaction due to income increases for individuals with high neuroticism scores. However, the results of Boyce and Wood (2011) showed inconsistencies between the different models used. For openness, Boyce and Wood (2011) found a negative interaction for women, whereas Soto and Luhmann (2012) reported inconsistencies between the different data sets analyzed. Proto and Rustichini (2015) illustrated that openness has no effect on how income affects life satisfaction. Our results suggest the opposite moderating effect. However, the well-being variables assessed here differed, as the emotional well-being variable used in this study consisted of several variables (i.e., life satisfaction, happiness, and affectivity) instead of only life satisfaction.

We studied each personality trait separately. However, human beings comprise a combination of several personality traits that likely operate together (Pulkkinen, 2017). A possible avenue for future research would be to examine the marginal utility of income and the moderating role of personality profiles, i.e., homogenous subgroups with distinct Big Five personality traits, instead of separate personality traits. For example, Kinnunen et al. (2012; see Pulkkinen, 2017 for updated titles of the profiles) illustrated the existence and continuity of the following personality profiles in the JYLS data: Resilient (high in extraversion and conscientiousness, low in neuroticism); Brittle (high in neuroticism, low in extraversion, and lower than average in openness, conscientiousness, and agreeableness); Overcontrolled (low in extraversion and openness, but differed from the Brittle in higher conscientiousness and agreeableness and lower neuroticism); Undercontrolled (high in openness and extraversion, low in conscientiousness), and Ordinary (mean value in all personality traits).

Our main aim in the present study was to explain mental well-being on the basis of income variables and how personality traits moderate the relation between income and mental well-being. However, it is also likely that high

mental well-being contributes to an individual's work career and, consequently, income level. We would like to emphasize that our results do not provide evidence of a causal relation between income and mental well-being, as significant relationships were only found using OLS. Evidence of the moderating role of personality traits in the income–mental well-being associations may be interpreted as support for causal relations, as the FE estimates describe, at best, average causal effects. Further research is needed to confirm our results, as the analyses were based on a moderately-sized dataset, and the insignificant results might have been due to a lack of statistical power. It would be interesting to further examine the causal relations over a longer period. Further, an avenue for future research would be to investigate whether the moderating effect of income on mental well-being differs between positive and negative income shocks. We believe that personality traits moderate both types of income shocks, but an assessment of how the moderating effects differ between these shocks is beyond the scope of the present paper.

Finally, the data collection at age 50 was undertaken in 2009, that is, after the 2008 US financial crisis and at the time of the financial crisis in Europe. In Finland, the crisis led to an 8% decrease in GDP, a 20% decrease in exports, and a 17% decrease in private investments in 2009. However, the decrease in private consumption was small due to fiscal policy actions and low interest rates. Further, unemployment increased only by three percentage points between 2008 and 2009 and started to decrease by 2010. Overall, the labor markets survived the crisis years well (Freystätter & Mattila, 2011). In our data, unemployment decreased, gross monthly income increased, and the household financial situation improved between the ages of 42 and 50. Therefore, we believe that even though the public sector suffered significantly during the crisis, and even though the accumulated budget deficit may affect the private sector for years to come, in 2009, our participants were not significantly affected by the crisis.

4.5 Conclusions

By using an age-cohort representative sample of longitudinal data, the present study suggested a positive, though limited, relationship between income and well-being in middle age. We found positive bivariate associations between income and mental well-being and its dimensions: emotional, psychological, and social well-being and reversed depression (OLS). Following the inclusion of the personality traits and control variables, gross monthly income was statistically significantly associated with reversed depression, and experienced household finances were related to mental well-being and its emotional well-being dimension (OLS). Once the labor market and education controls were added, income no longer yielded a statistically significant coefficient in the case of reversed depression.

Based on our results, the marginal utility of income seemed to depend on personality traits (FE): agreeableness and extraversion negatively moderated the

gross monthly income–emotional well-being relationship, while openness positively moderated this relationship. In addition to emotional well-being, extraversion negatively moderated the relationship between gross monthly income and general mental well-being, while neuroticism negatively moderated the association between gross monthly income and the dimension of social well-being. An avenue for future research would be to examine why certain personality traits moderate specific dimensions of mental well-being.

Based on our results, income, particularly the experienced household financial situation, was most consistently associated with the dimension that the economics literature has concentrated on: emotional well-being. Similarly, in the interaction analyses, the personality traits mainly moderated the income and emotional well-being relationships. While it may be interesting to see how income associates with emotional well-being or one of its subcomponents (such as happiness or life satisfaction), we suggest further research into general mental well-being in relation to income. If only one dimension of mental well-being is studied, it should be noted that it would not describe the individual's mental well-being as a whole.

References

Angeles, L. (2011). A closer look at the Easterlin paradox. *The Journal of Socio-Economics, 40*(1), 67–73. https://doi.org/10.1016/j.socec.2010.06.017

Anger, S., Camehl, G., & Peter, F. (2017). Involuntary job loss and changes in personality traits. *Journal of Economic Psychology, 60,* 71-91. https://doi.org/10.1016/j.joep.2017.01.007

Boyce, C. J., & Wood, A. M. (2011). Personality and the marginal utility of income: Personality interacts with increases in household income to determine life satisfaction. *Journal of Economic Behavior & Organization, 78*(1), 183–191. https://doi.org/10.1016/j.jebo.2011.01.004

Boyce, C. J., Wood, A., & Powdthavee, N. (2013). Is personality fixed? Personality changes as much as "variable" economic factors and more strongly predicts changes to life satisfaction. *Social Indicators Research, 111*(1), 287-305. https://doi.org/10.1007/s11205-012-0006-z

Brown, S., & Gray, D. (2016). Household finances and well-being in Australia: An empirical analysis of comparison effects. *Journal of Economic Psychology, 53*, 17–36. https://doi.org/10.1016/j.joep.2015.12.006

Brown, S., Taylor, K., & Wheatley Price, S. (2005). Debt and distress: Evaluating the psychological cost of credit. *Journal of Economic Psychology, 26*(5), 642–663. https://doi.org/10.1016/j.joep.2005.01.002

Brown, S., & Taylor, K. (2014). Household finances and the 'Big Five' personality traits. *Journal of Economic Psychology, 45*, 197-212. https://doi.org/10.1016/j.joep.2014.10.006

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature, 46*(1), 95–144. https://doi.org/10.1257/jel.46.1.95

Costa, P. T., & McCrae, R. R. (1985). The NEO Personality Inventory. Odessa, FL: Psychological Assessment Resources.

Costa, P., & McCrae, R. (1989). The NEO/NEO-FFI Manual Supplement. Odessa, FL: Psychological Assessment Resources.

Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the Gallup World Poll. *The Journal of Economic Perspectives, 22*(2), 53–72. https://doi.org/10.1257/jep.22.2.53

Depue, R. (1987). General Behavior Inventory. Ithaca, NY: Department of Psychology, Cornell University.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*(3), 542–575. https://doi.org/10.1037/0033-2909.95.3.542

Diener, E., Suh, E., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*(2), 276–302. https://doi.org/10.1037/0033-2909.125.2.276

Diener, E., Hentzelman, S. J., Kushlev, K., Tay, L., Wirtz, D., Lutes, L. D., & Oishi, S. (2017). Findings all psychologists should know from the new

science on subjective well-being. *Canadian Psychology, 58*(2), 87-104. https://doi.org/10.1037/cap0000063
Drentea, P. (2000). Age, debt and anxiety. *Journal of Health and Social Behavior, 41*(4), 437–450. https://doi.org/10.2307/2676296
Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness–income paradox revisited. *Proceedings of the National Academy of Sciences, 107*(52), 22463–22468. https://doi.org/10.1073/pnas.1015962107
Elkins, R., Kassenboehmer, S., & Schurer, S. (2017). The stability of personality traits in adolescence and young adulthood. *Journal of Economic Psychology, 60*, 37-52. https://doi.org/10.1016/j.joep.2016.12.005
Feldman, L. A. (1995). Variations in the circumplex structure of mood. *Personality and Social Psychology Bulletin, 21*(8), 806–817. https://doi.org/10.1177/0146167295218003
Ferrer-i-Carbonell, A. (2005). Income and well-being: An empirical analysis of the comparison income effect. *Journal of Public Economics, 89*(5), 997–1019. https://doi.org/10.1016/j.jpubeco.2004.06.003
Frey, B. S. (2008). *Happiness: A revolution in economics*. Cambridge, Mass: MIT Press.
Freystätter, H., & Mattila, V. (2011). Finanssikriisin vaikutuksista Suomen talouteen [Effects of the financial crisis on economy of Finland]. *BoF Online, 2011*(1). http://urn.fi/URN:NBN:fi:bof-20140807703
Gallagher, M. W., Lopez, S. J., & Preacher, K. J. (2009). The hierarchical structure of well-being. *Journal of Personality, 77*(4), 1025–1050. https://doi.org/10.1111/j.1467-6494.2009.00573.x
Golsteyn, B., Schildberg-Hörisch, H. (2017). Challenges in research on preferences and personality traits: Measurement, stability, and inference. *Journal of Economic Psychology. 60,* 1-6. https://doi.org/10.1016/j.joep.2017.03.001
Headey, B., & Wooden, M. (2004). The effects of wealth and income on subjective well-being and ill-being. *Economic Record, 80*(1), 24–33. https://doi.org/10.1111/j.1475-4932.2004.00181.x
Helson, R., Jones, C. J., & Kwan, S. Y. (2002b). Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. *Journal of Personality & Social Psychology, 83*(3), 752–766. https://doi.org/10.1037/0022-3514.83.3.752
Helson, R., Kwan, V. S. Y., John, O. P., & Jones, C. (2002a). The growing evidence for personality change in adulthood: Findings from research with personality inventories. *Journal of Research in Personality, 36*(4), 287–306. https://doi.org/10.1016/S0092-6566(02)00010-7
Inglehart, R., Foa, R., Peterson, C., & Welzel, C. (2008). Development, freedom, and rising happiness: A global perspective (1981–2007). *Perspectives on Psychological Science, 3*(4), 264–285. https://doi.org/10.1111/j.1745-6924.2008.00078.x

Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly, 61*(2), 121–140. https://doi.org/10.2307/2787065

Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*(2), 207–222. https://doi.org/10.2307/3090197

Keyes, C. L. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology, 73*(3), 539–548. https://doi.org/10.1037/0022-006X.73.3.539

Keyes, C. L. (2006). Subjective well-being in mental health and human development research worldwide: An introduction. *Social Indicators Research, 77*(1), 1–10. https://doi.org/10.1007/s11205-005-5550-3

Keyes, C. L. M., & Waterman, M. B. (2003). Dimensions of well-being and mental health in adulthood. In M. H. Bornstein, L. Davidson, C. L. M. Keyes, K. A. Moore, & The Center for Child Well-Being (Eds.), *Well-being: Positive development across the life course* (pp. 477–497). Mahwah, NJ: Lawrence Erlbaum Associates.

Kinnunen, M.-L., Metsäpelto, R. L., Feldt, T., Kokko, K., Tolvanen, A., Kinnunen, U., et al. (2012). Personality profiles and health: Longitudinal evidence among Finnish adults. *Scandinavian Journal of Psychology, 53*(6), 512–522. https://doi.org/10.1111/j.1467-9450.2012.00969.x

Kokko, K., Korkalainen, A., Lyyra, A., & Feldt, T. (2013b). Structure and continuity of well-being in mid-adulthood: A longitudinal study. *Journal of Happiness Studies, 14*(1), 99–114. https://doi.org/10.1007/s10902-011-9318-y

Kokko, K., Rantanen, J., & Pulkkinen, L. (2015). Associations between mental well-being and personality from a life span perspective. In M. Blatný (Ed.), *Personality and well-being across the life-span* (pp. 134–159). London: Palgrave Macmillan.

Kokko, K., Tolvanen, A., & Pulkkinen, L. (2013a). Associations between personality traits and psychological well-being across time in middle adulthood. *Journal of Research in Personality, 47*(6), 748–756. https://doi.org/10.1016/j.jrp.2013.07.002

Kokkonen, M. (2001). *Emotion regulation and physical health in adulthood: A longitudinal, personality-oriented approach.* Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-8444-1

Korkalainen, A. (2007). Psykologian tutkimuksessa käytettyjen hyvinvoinnin osa-alueiden suhde toisiinsa, masentuneisuuteen, persoonallisuuden piirteisiin ja objektiivisiin tekijöihin [Association between different dimensions of well-being and their links to personality traits and objective factors]. Unpublished master's thesis, University of Jyväskylä, Finland.

Lachman, M., & Weaver S., L. (1998). The sense of control as a moderator of social class differences in health and well-being. *Journal of Personality and Social Psychology, 74*(3), 763–773. https://doi.org/10.1037//0022-3514.74.3.763

Luttmer, E. (2005). Neighbors as negatives: relative earnings and well-Being. *Quarterly Journal of Economics, 120 (3)*, 963–1002. https://doi.org/10.1093/qje/120.3.963

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology, 55*(1), 102–111. https://doi.org/10.1037/0022-3514.55.1.102

McInerney, M., Mellor, J. M., & Nicholas, L. H. (2013). Recession depression: Mental health effects of the 2008 stock market crash. *Journal of Health Economics, 32*(6), 1090–1104. https://doi.org/10.1016/j.jhealeco.2013.09.002

Mroczek, D. K., & Spiro, A. III. (2003). Modeling intraindividual change in personality traits: Findings from the normative aging study. *Journals of Gerontology B*: *Psychological Sciences and Social Sciences, 58*(3), 153–165. https://doi.org/10.1093/geronb/58.3.P153

Mueller, G., & Plug, E. (2006). Estimating the effect of personality on male and female earnings. *Industrial and Labour Relations Review, 60(1),* 3–22. https://doi.org/10.1177/001979390606000101

Official Statistics of Finland (OSF): Consumer price index [e-publication]. ISSN=1799-0254. Helsinki: Statistics Finland. http://www.stat.fi/til/khi/index_en.html. Accessed 13 June 2017.

Perho, H., & Korhonen, M. (1993). Elämänvaiheiden onnellisuus ja sisältö keski-iän kynnyksellä [Happiness of life stages in front of middle age]. *Gerontologia, 7*(4), 271–285.

Proto, E., & Rustichini, A. (2015). Life satisfaction, income and personality. *Journal of Economic Psychology, 48*, 17–32. https://doi.org/10.1016/j.joep.2015.02.001

Pulkkinen, L. (2017). *Human development from middle to childhood to middle adulthood: Growing up to the middle-aged* (in collaboration with Katja Kokko). London: Routledge.

Pulkkinen, L., & Kokko, K. (2010). Keski-ikä elämänvaiheena [Middle age as a stage of life]. Jyväskylä: University of Jyväskylä: Reports from the Department of Psychology, 352.

Pulkkinen, L., Kokko, K., & Rantanen, J. (2012). Paths from socioemotional behavior in middle childhood to personality in middle adulthood. *Developmental Psychology, 48*(5), 1283–1291. https://doi.org/10.1037/a0027463

Pulver, A., Allik, J., Pulkkinen, L., & Hämäläinen, M. (1995). A Big Five personality inventory in two non-Indo-European languages. *European Journal of Personality, 9*(2), 109–124. https://doi.org/10.1002/per.2410090205

Rantanen, J., Metsapelto, R.-L., Feldt, T., Pulkkinen, L., & Kokko, K. (2007). Long-term stability in the Big Five personality traits in adulthood. *Scandinavian Journal of Psychology, 48(6*), 511–518. https://doi.org/10.1111/j.1467-9450.2007.00609.x

Roberts, B. W., Caspi, A., & Moffitt, T. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81*(4), 670–683. https://doi.org/10.1037/0022-3514.81.4.670

Roberts, B. W., Caspi, A., & Moffitt, T. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology, 84*(3), 582–593. https://doi.org/10.1037/0022-3514.84.3.582

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*(1), 31–35. https://doi.org/10.1111/j.1467-8721.2008.00543.x

Roberts, B. W., Walton, K., Bogg, T., & Caspi, A. (2006a). De-investment in work and non-normative personality trait change in young adulthood. *European Journal of Personality, 20*(6), 461–474. https://doi.org/10.1002/per.607

Roberts, B. W., Walton, K., & Viechtbauer, W. (2006b). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Robitschek, C., & Keyes, C. L. (2009). Keyes's model of mental health with personal growth initiative as a parsimonious predictor. *Journal of Counseling Psychology, 56*(2), 321–329. https://doi.org/10.1037/a0013954

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125*(1), 3–30. https://doi.org/10.1037/0033-2909.125.1.3

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*(6), 1069–1081. https://doi.org/10.1037/0022-3514.57.6.1069

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*(4), 719–727. https://doi.org/10.1037/0022-3514.69.4.719

Scollon, C. N., & Diener, E. (2006). Love, work, and changes in extraversion and neuroticism over time. *Journal of Personality and Social Psychology, 91*, 1152–1165. https://doi.org/10.1037/0022-3514.91.6.1152

Sen, A. (1973). Behaviour and the concept of preference. *Economica, 40*(159), 241–259. https://doi.org/10.2307/2552796

Sen, A. (1999). *Development as freedom*. New York: Random House.

Small, B. J., Hertzog, C., Hultsch, D. F., & Dixon, R. A. (2003). Stability and change in adult personality over 6 years: Findings from the Victoria Longitudinal Study. *Journals of Gerontology B: Psychological Sciences & Social Sciences, 58*(3), 166–176. https://doi.org/10.1093/geronb/58.3.P166

Soto, C. J., & Luhmann, M. (2012). Who can buy happiness? Personality traits moderate the effects of stable income differences and income fluctuations on life satisfaction. *Social Psychological and Personality Science, 4*(1), 46-53. https://doi.org/10.1177/1948550612444139

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin, 134*, 138–161. https://doi.org/10.1037/0033-2909.134.1.138

Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox. *Brookings Papers on Economic Activity, 39*(1), 1–87. http://www.jstor.org/stable/27561613

Tay, L., Batz, C., Parrigon, S., & Kuykendall, L. (2017). Debt and subjective well-being: the other side of the income-happiness coin. *Journal of happiness studies, 18*(3), 903-937. *https://doi.org/10.1007/s10902-016-9758-5*

Veenhoven, R. (1991). Questions on happiness: Classical topic, modern answers, blind spots. In M. Argyle, N. Schwarz, & F. Strack (Eds.), *Subjective well-being*: *An interdisciplinary perspective* (pp. 7–26). Oxford: Pergamon.

Viinikainen, J., Kokko, K., Pulkkinen, L., & Pehkonen, J. (2010). Personality and labour market income: Evidence from longitudinal data. *Labour, 24*(2), 201–220. https://doi.org/10.1111/j.1467-9914.2010.00477.x

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology, 64*(4), 678–691. https://doi.org/10.1037/0022-3514.64.4.678

Zimmerman, F., & Katon, W. (2005) Socioeconomic status, depression disparities, and financial strain: What lies behind the income-depression relationship? *Health Economics, 14*(12), 1197–1215. https://doi.org/10.1002/hec.1011

Appendix

TABLE 6 Means and standard deviations of the control variables (N = 163); p-values from paired t-tests

	Age 42 (2001)		Age 50 (2009)		
Employment situation	*Mean*	*SD*	*Mean*	*SD*	*p-value*
Unemployed	0.09	0.28	0.08	0.27	0.836
Part-time employee	0.08	0.27	0.05	0.22	0.198
Full-time employee	0.83	0.37	0.87	0.346	0.290
Stability of career line					
Unstable career	0.09	0.28	0.07	0.26	0.565
Changeable career	0.08	0.27	0.09	0.29	0.696
Stable career	0.83	0.37	0.83	0.37	1.000
Occupational status					
Blue-collar	0.26	0.44	0.22	0.42	0.052
Lower white-collar	0.46	0.50	0.49	0.50	0.286
Upper white-collar	0.28	0.45	0.29	0.46	0.319
State of health					
Very good health	0.16	0.37	0.19	0.39	0.386
Fairly good health	0.64	0.48	0.53	0.50	0.029
Moderate health	0.18	0.38	0.25	0.44	0.064
Fairly bad health	0.03	0.16	0.03	0.16	1.000
Very bad health	0.00	0.00	0.01	0.08	0.319
Relationship status					
Single	0.07	0.26	.06	.23	0.319
Married	0.69	0.46	.67	.47	0.451
Cohabitation	0.12	0.32	.09	.29	0.319
Divorced or widowed	0.12	0.32	.18	.39	0.027
Size of the household					
Household size	3.65	1.34	3.03	1.26	0.000
Education					
Course	0.14	0.35	0.14	0.35	-
Vocational school	0.36	0.48	0.36	0.48	-
Vocational college	0.34	0.48	0.34	0.48	-
University	0.15	0.36	0.15	0.36	-
Gender					
Female	0.55	0.50	0.55	0.50	-
Follow-up year					
Year 2009	0.00	0.00	1.00	0.00	-

Notes: R = reversed score; education and gender recorded no variation between ages 42 and 50.

TABLE 7 Pooled OLS and FE regressions: Gross monthly income, household finances, and mental well-being and its different dimensions

	Mental well-being (S)			Emotional well-being (S)			Psychological well-being (S)			Social well-being (S)			Depression (S, R)		
	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)	OLS (1)	OLS (2)	FE (3)
Log gross monthly income	.516**	-.036	-.150	.339**	-.201	-.301	.363**	-.079	-.211	.307*	.050	.012	.564**	.121	.043
	(.124)	(.106)	(.138)	(.122)	(.155)	(.218)	(.115)	(.114)	(.158)	(.132)	(.148)	(.148)	(.134)	(.115)	(.167)
R^2	.096	.674	.450	.041	.416	.331	.046	.586	.286	.061	.439	.333	.096	.469	.233
Household finances	.343**	.122*	-.014	.345**	.194*	.218	.268*	.076	-.117	.156	.025	-.081	.279**	.077	-.062
	(.103)	(.061)	(.074)	(.103)	(.082)	(.116)	(.104)	(.065)	(.110)	(.084)	(.072)	(.079)	(.099)	(.073)	(.083)
R^2	.072	.679	.445	.061	.425	.339	.040	.587	.287	.045	.439	.337	.039	.469	.235
Personality traits and controls	-	√	√	-	√	√	-	√	√	-	√	√	-	√	√
N	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326

Notes: **$p < 0.01$; *$p < 0.05$. S = standardized score, R = reversed score. Controls: gender, relationship status, employment situation, stability of career line, occupational status, education, state of health, household size, and a dummy for the second survey year. Specifications 1 and 2 were analyzed using OLS. Specification 1 controlled for the second survey year, and specification 2 added all other controls. Specification 3 was analyzed using FE and included controls other than education and gender, which were excluded because of a lack of within variation between the given time periods. Standard errors are in parentheses.

TABLE 8 FE regressions: Gross monthly income, household finances, and personality trait interaction effects on mental well-being and its different dimensions

	Mental well-being (S)		Emotional well-being (S)		Psychological well-being (S)		Social well-being (S)		Depression (S, R)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Income										
Log gross monthly income	-0.136		-0.171		-0.179		-0.029		-0.037	
	(0.143)		(0.232)		(0.162)		(0.147)		(0.176)	
Household finances		-0.035		0.206*		-0.167		-0.096		-0.048
		(0.068)		(0.103)		(0.091)		(0.084)		(0.090)
Personality trait interactions with income										
Agreeableness (S)	-.137	-.162	-.416*	-.231	-.164	-.155	-.032	-.047	.194	-.062
	(.132)	(.083)	(.196)	(.130)	(.148)	(.082)	(.138)	(.060)	(.136)	(.096)
Conscientiousness (S)	.131	.023	.287	.016	.151	.043	.017	.093	-.054	-.082
	(.094)	(.072)	(.162)	(.120)	(.103)	(.086)	(.095)	(.061)	(.137)	(.079)
Extraversion (S)	-.260*	-.157	-.364*	-.085	-.294*	-.189	-.213	-.103	.077	-.103
	(.109)	(.088)	(.179)	(.138)	(.146)	(.104)	(.131)	(.080)	(.164)	(.097)
Neuroticism (S)	-.138	-.141*	-.027	-.007	-.228	-.092	-.264*	-.123	.098	-.208*
	(.104)	(.071)	(.185)	(.107)	(.129)	(.092)	(.111)	(.077)	(.127)	(.092)
Openness to new experiences (S)	.203*	.006	.288*	.176	.185	-.125	.137	-.070	.011	.038
	(.102)	(.070)	(.141)	(.094)	(.120)	(.089)	(.104)	(.081)	(.175)	(.100)
Personality traits										
Agreeableness (S)	1.111	0.550*	3.143*	0.583	1.388	0.622*	0.376	0.305	-1.519	0.169
	(1.048)	(0.276)	(1.544)	(0.410)	(1.173)	(0.274)	(1.045)	(0.173)	(1.085)	(0.294)
Conscientiousness (S)	-0.844	0.070	-1.965	0.143	-1.050	-0.054	-0.147	-0.292	0.587	0.416
	(0.736)	(0.204)	(1.251)	(0.341)	(0.789)	(0.239)	(0.719)	(0.174)	(1.066)	(0.251)
Extraversion (S)	2.170*	0.623*	2.873*	0.292	2.483*	0.774*	1.767	0.440	-0.504	0.395
	(0.871)	(0.294)	(1.411)	(0.443)	(1.158)	(0.354)	(1.014)	(0.251)	(1.294)	(0.315)

continues

TABLE 8 continues

	Mental well-being (S)		Emotional well-being (S)		Psychological well-being (S)		Social well-being (S)		Depression (S, R)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Neuroticism (S)	0.737	0.071	-0.196	-0.366	1.496	-0.010	1.884*	0.201	-0.936	0.392
	(0.827)	(0.201)	(1.435)	(0.296)	(1.009)	(0.276)	(0.854)	(0.215)	(0.979)	(0.296)
Openness to new experiences (S)	-1.335	0.187	-1.649	0.079	-1.248	0.494*	-1.070	0.144	-0.103	-0.147
	(0.789)	(0.209)	(1.072)	(0.272)	(0.903)	(0.248)	(0.783)	(0.246)	(1.349)	(0.284)
Personality and Controls	√	√	√	√	√	√	√	√	√	√
R^2										
N	326	326	326	326	326	326	326	326	326	326

Notes: **$p < 0.01$; *$p < 0.05$. S = standardized score, R = reversed score. Specification 1 was analyzed using gross monthly income as an independent variable and specification 2 with household finances as an independent variable. Both specifications were estimated using FE, and controls (relationship status, employment situation, stability of career line, occupational status, state of health, household size, and a dummy for the second survey year) were included in all regressions. Education and gender were excluded because of a lack of within variation between the given time periods. Standard errors are in parentheses.

YHTEENVETO (SUMMARY)

Tämä väitöskirja tutkii koulutusta, työmarkkinatulemia ja henkistä hyvinvointia alkaen alakouluikäisistä lapsista, siirtyen nuoriin aikuisiin ja lopulta keski-ikään. Väitöskirjan alussa on kokonaisuutta esittelevä johdantoluku (Luku 1), jota seuraa kolme empiiristä artikkelia (Luvut 2–4). Artikkeleissa hyödynnetään soveltavia ekonometrisia menetelmiä, kuten kiinteiden vaikutusten mallia, instrumenttimuuttuja menetelmää ja regressioepäjatkuvuus asetelmaa. Tutkimuksissa käytettävät aineistot yhdistävät laadullisia ja määrällisiä muuttujia. Aineistoina toimivat ProKoulu-tutkimuksen aineisto, Lapsesta aikuiseksi-pitkittäistutkimusaineisto, kauppatieteellisen alan yhteisvalinta-aineisto sekä Tilastokeskuksen rekisteriaineistot.

Ensimmäinen artikkeli Luvussa 2 tutkii luokkakoon vaikutuksia työrauhaan ja koulumenestykseen ala-asteikäisillä. Tutkimus osoittaa, että isommissa luokissa on enemmän työrauhahäiriöitä ja antaa myös viitteitä luokkakokoon ja koulumenestyksen välisestä negatiivisesta yhteydestä.

Toinen artikkeli Luvussa 3 tarkastelee kauppakorkeakoulujen laatua ja työmarkkinatulemia. Artikkeli osoittaa, että hakijoilla, jotka hyväksyttiin tai jotka aloittivat selektiivisemmässä kauppakorkeakoulussa vähemmän selektiivisen sijaan, oli vähemmän työttömyyspäiviä työuran alussa ja heillä oli korkeampi valmistumistodennäköisyys. Kertoimet tuloihin liittyen olivat positiivisia, mutta eivät tilastollisesti eronneet nollasta.

Viimeinen väitöskirjan artikkeli (Luku 4) tarkastelee tulojen yhteyttä henkiseen hyvinvointiin, sekä sitä, miten persoonallisuus mahdollisesti muokkaa tätä yhteyttä. Tutkimuksen perusteella tulotasolla on rajallinen merkitys osana keskiikäisten henkistä hyvinvointia. Lisäksi persoonallisuuden piirteet näyttivät liittyvän lähimmin yhteyteen tulotason ja tietyn henkisen hyvinvoinnin ulottuvuuden, emotionaalisen hyvinvoinnin välillä.